DOCUMENT RESUME

ED 055 991 24 SP 005 355

AUTHOR Ryans, David G.
TITLE Teacher Evaluation Research, Part I: Consideration of Critical Issues, Feasibility of Collaborative Research, and Overall Design. Final Report.
INSTITUTION Hawaii Univ., Honolulu. Education Research and Development Center.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-9-8020
PUB DATE Jul 71
GRANT OEG-9-9-218020-0061(010)
NOTE 37p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS Cultural Pluralism; *Educational Research; *Effective Teaching; *Evaluation Criteria; *Evaluation Methods; Measurement Instruments; Teacher Behavior; *Teacher Evaluation

ABSTRACT

This report briefly reviews research on teacher effectiveness and presents the conclusions of a 2-day conference held in October 1969 to consider issues related to the assessment of teacher effectiveness and make recommendations to the U.S. Office of Education (USOE) regarding the encouragement of research in this area. The committee concluded that USOE should encourage and support a large-scale systematic approach to research on teacher effectiveness, including the development and testing of measurement instruments which allow for cultural diversity and varying criteria. These instruments would be used in descriptive research to identify teacher and student characteristics and behaviors and longitudinal experimental research which would assess effects, changes, and relationships that may be inferred to result from identifiable teacher education experiences. (The report includes an 89-item bibliography.) (RT)

FINAL REPORT

Project No. 9-8020
Grant No. 9-9-218020-0061(010)

# TEACHER EVALUATION RESEARCH, PART I: CONSIDERATION OF CRITICAL ISSUES, FEASIBILITY OF COLLABORATIVE RESEARCH, AND OVERALL DESIGN

David G. Ryans

University of Hawaii

Honolulu, Hawaii

July, 1971

The research reported herein was performed pursuant to a grant with the Office of Education, U. S. Department of Health, Education, and Welfare. Contractors undertaking such projects under government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U. S. DEPARTMENT OF
HEALTH, EDUCATION, AND WELFARE

Office of Education
Bureau of Research

Education Research and Development Center
University of Hawaii
Honolulu, Hawaii 96822

Contents

## INTRODUCTION

This report represents an effort to present some general points of view regarding the status of research on teacher effectiveness and make recommendations to the U. S. Office of Education regarding the encouragement of research in this area.

Consideration of the extensive investigation of teacher competency leads to the conclusions that many facets of the area have been given attention by researchers. Approaches employed have been diverse and the findings capable of little generalizability. Considerations of methodological problems and conduct of unified, systematic studies have been few. And, although a number of potentially important clues are available, differences in points of view relative to instructional processes, guiding philosophies of teacher education, and methodology, together with the difficulties involved in obtaining large scale cooperation and replication, have resulted in a state of confusion about teacher effectiveness and, indeed, a state of distrust. The opinion held by many educators, therefore, is that teacher evaluation, or the assessment of teacher effectiveness, is "intangible" and incapable of empirical study and/or attainment in practice.

To the contrary, the committee reports that although the difficulties are real, teacher effectiveness is capable of assessment.

It is the opinion of the committee that procedures can be applied to estimate teacher effectiveness or dimensions of teacher effectiveness. And it is the committee's recommendation that the U. S. Office of Education encourage and support research, particularly research along lines to be noted in this report, to develop, validate, and disseminate instruments and procedures for the assessment of teacher effectiveness.

## I. NATURE OF THE REPORT

A. The chairman of the committee charged with preparing this report takes full responsibility for its contents.

Members of the committee shared generously during a two-day meeting ideas and suggestions that were as stimulating as they were diverse. They ranged from complete agreement upon the need for theory-based approaches to teacher effectiveness and the need for continued empirical study of many facets of the problem, to views that research properly should be based upon the continued study of teachers in the roles they currently play in the educating process, and an almost diametrically opposed position stating that such study (of what teachers do now in manifesting certain dimensions of teacher behavior and/or in producing certain pupil skills) is destined to be sterile, at least in its contribution to the future, and that teacher effectiveness research should instead direct its attention to the role of the teacher in an envisioned computerized, completely individualized, media-oriented educational environment--a future when the teacher's "motivating, set-providing, and reinforcing" behaviors, "presenting/explaining" behaviors, "evaluating" behaviors, "counseling/advising" behaviors, and to a large degree, "planning and organizing" behaviors may be reduced to a minimum and new responsibilities are taken on as technician, "man-machine" manager of pupil learning. (Proponents of the latter view felt research should more properly be directed to the experimental study of pupil-machine relationships and to the most effective ways the teacher could intervene, complement, and control pupil learning in such an educational environment.) The contributions of the committee membership, thus, were many.

The chairman hopes omissions, inadvertently made, are not numerous. He regrets any misrepresentations of viewpoints [illegible] may be reflected. He does candidly admit that the present report represents his own best summarization of the state of teacher effectiveness research, taking into account the committee deliberations as well as the abundance of literature that has accumulated, some preceding and much important work appearing subsequent to the committee meeting (e.g., the critiques of Rosenshine).

B. The report which follows is brief. Extensive and often competent reviews of theoretical positions relative to teacher effectiveness (including the need for theory in guiding such research), of methodology (including the all-important "criterion" problem), and of research, good and poor, are omitted for the most part; the literature is replete with references to teacher effectiveness, teacher characteristics, criteria of teacher competency, descriptions of teachers (and "most liked" and "least liked" teacher traits), and suggested means for improving instruction. References to some of the major and more recent studies and reviews are listed in the appendix devoted to "References." The bibliography is a highly selected one, but a reader will find the references noted,

plus those that accompany the individual studies or reviews, include most of the important topics relating to teacher effectiveness research. Redundance is avoided in this report by frequent reference to the bibliographical appendix.

C. Essentially the point of view taken by this report is that the study of teacher effectiveness is neither an area that already has exhausted means of approach nor one that is incapable of empirical investigation. Practical problems are recognized. Differences of opinions among educators are recognized regarding the direction of and approaches to research on teacher behavior (e.g., study of the teacher of the present in his roles of planner, motivator/reinforcer, presenter/explainer, evaluator, and counselor--an interacting aide and guide to the pupil--is contrasted with study of the possible teacher of several decades hence, with a major role of technician/organizer/adapter of learning environments (including computers and other media) to assist pupil learning by introduction to appropriate steps of largely mechanized programs that are preplanned to fit varying stages of development reached by pupils; or possibly middle-of-the-road type study directed at identifying and emphasizing "teacher-dependent" classroom activities as distinct from "teacher-independent" activities of the future).

Perhaps research in a number of such directions should be supported; they well may complement one another. But the opinion of the committee definitely embraces the position that teacher effectiveness is capable of empirical research and that such research be encouraged and substantially supported with funding from the U. S. Office of Education.

## II. BACKGROUND AND OBJECTIVES

In the spring of 1969 a proposal was submitted to the U. S. Office of Education for the conduct of a small group two-day conference to consider issues germane to the assessment and evaluation of teacher effectiveness.

It was proposed that the participants be researchers and scholars in pertinent areas and that the broad objectives of the conference would be to (1) consider critical issues in the identification, validation, and practicality of techniques of assessing teacher effectiveness, and (2) if a major effort appeared feasible, formulate recommendations for relevant research, including instrumentation, that might be disseminated for use by teacher educators and by public schools.

It was proposed that the initiator would serve as coordinator of the conference and that a small group of selected educators would review the feasibility of research addressed to salient objectives. Prior to the conference, the proposer/coordinator was to send to each of the collaborators a statement delineating some of the major facets of the problem. Conferees would be asked to make modifications of the agenda that seemed desirable, particularly in areas in which a conferee was best qualified by virtue of his training and experience. Each collaborator would be requested to prepare a short outline of "guidelines" for research and development in the teacher effectiveness domain. The submitted outlines would be distributed to participants prior to the conference.

Following the conference the proposer was to prepare a report representing the considerations of the group. Participants were to review this report and suggest revisions which might be incorporated in the final report to the Office of Education. If a major program were to appear feasible to the conferees, it was hoped the report might serve as a basis for subsequent proposals for complementary studies, collaborative in nature and involving frequent exchange of information among investigators. Such a report, it was hoped, might provide a general design by suggesting major facets of a problem toward which research and development in the area of effectiveness of teachers should be directed.

Plans were originally made to hold the conference in Chicago on May 2 and 3, 1969. Due to commitments of the conferees it became necessary to postpone the meeting until early October, 1969.

The conferees constituting the "task force" consisted of Dr. Lindley J. Stiles, Professor of Education for Interdisciplinary Studies, Sociology, and Political Science, Northwestern University; Dr. Fred M. Kerlinger, Professor of Educational Research, New York University; Dr. Richard L. Turner, Associate Dean for Research, Indiana University; Dr. Harold E. Mitzel, Assistant Dean for Research,

Pennsylvania State University; Dr. Robert Beezer, Program Director, Basic Studies Branch, representing the U. S. Office of Education; and Dr. David G. Ryans, Director, Education Research and Development Center, University of Hawaii. David G. Ryans served as chairman and coordinator of the group.

In keeping with the proposal, the chairman sent to each collaborator a tentative statement suggesting possible issues and questions in the area of teacher effectiveness research as topics that might be considered at the meeting. Some members of the task force responded either by submitting additional items prior to the meeting or at the time of the conference in Chicago.

At the conference it was agreed that a report would be prepared by the chairman. The report would be transmitted to each member of the group and to the U. S. Office of Education, with recommendations for subsequent attention by the Office of Education staff regarding the sponsorship of a research program to better come to grasp with the problems of assessment and evaluation of teacher effectiveness.

The procedure followed was essentially that suggested in the proposal--with one major exception. Due to the extreme tardiness of the chairman in preparing the report (a circumstance contributed to, in part, by (a) the chairman's own participation in an updating and continuation of extensive research in the area of teacher characteristics and (b) the increasing appearance in the literature of reviews--particularly that of Rosenshine (48) and Rosenshine and Furst (55)--and other studies that updated thinking on the topic beyond that which prevailed at the time of the committee meeting) the chairman has prepared this long overdue report just prior to an overseas assignment and has sent it simultaneously to the U. S. Office of Education and the committee members, neglecting the planned intermediate step which originally called for a preliminary version subject to review and amendment by other committee members. For this the chairman offers apologies to his fellow committee members and urges any member who feels so inclined to write an appending statement which the U. S. Office of Education may add to the report.

## III. MAJOR THESES

The major theses of this report are seven: (1) that instruments for the assessment of teacher effectiveness can be developed; (2) that such instrumentation necessarily must take into account the diversity of expectations of different school systems and communities with regard to behaviors or characteristics desired of teachers, i.e., that the cross-cultural nature of the problem cannot be ignored; (3) that a large scale, integrated, systematic attack need be made on the problem rather than further studies representative of the blunderbuss approach most frequently observed in the past when (a) a randomly selected teacher trait or sample of traits was correlated with some vague, contaminated, deficient, and unreliable criterion, or (b) a selected and sometimes operationally defined criterion, e.g., pupil performance in a subject-matter skill was correlated with some specific sample of teacher performance that was hypothesized with doubtful rationale to be related to the desired pupil performance; (4) that to avoid pitfalls that have deterred teacher effectiveness research in the past, (a) careful attention must be given to the identification, cataloging, and development of measures reflecting characteristics of teachers at any given time or place, (b) the similarities and differences among such teacher characteristics in cultures and subcultures, i.e., communities and schools that differ in educational objectives, expectations, and value systems espoused, need be researched, (c) "master instruments" that may be used in all identifiable subcultures (instruments that embrace scales which reflect both common characteristics and unique characteristics desired in the variety of communities included in the master sample of schools and school systems, i.e., instruments that will permit a particular school system to identify teachers possessing the qualities desired in that community, while ignoring qualities that may be desired in other communities) must be compiled; (5) that with such instruments (instruments that obviously must be subject to minor modification, or possibly major revision, as objectives of schools may change and/or responsibilities of teachers may change with introduction of new curricula, teaching programs, media, etc.) teacher education programs could and should take a "hard look" at the extent to which the objectives of their respective teacher education programs, innovative practices, etc., actually selected and/or produced teachers who possessed the characteristics expected from such a teacher education program or practice, e.g., effects of micro teaching, teaching by discovery, teaching for desired personal or social attitudes on the part of pupils, etc., as they are manifested by teacher education graduates; (6) that important as it is in teacher effectiveness research and in all research, the criterion against which the efficacy of a teacher education course or program is ultimately a matter of "reasoned decision" of those responsible for an educational program; whether such efficacy, or teacher effectiveness, be judged from quality of teacher performance per se (i.e., whether the teacher performs in the required manner under specified circumstances) or from hypothesized products of teacher performance (i.e., pupil change in specified behaviors) is

a matter of choice or decision. (A teacher, or a surgeon, may with his respective repertoire of responses perform either capably or ineptly with similar results--pupils may or may not learn the desired responses due to conditions over which the teacher has little control, or the surgeon's patients may respond to surgery favorably or unfavorably quite apart from the surgeon's own skill; quality of "behavior in process" often is a useful criterion in itself. On the other hand, few persons would fail to acknowledge that if (a) adequate controls over non-teacher or non-surgeon conditions could be maintained and (b) satisfactory estimates of student change (e.g., in cognitive or motor skills, attitudes, adjustment, etc.) and of patient improvement or recovery could be effected, and, of course, provided there was sufficient replication to permit generalization, the "products" of the teacher or surgeon would weigh heavily in judgement of his effectiveness or ineffectiveness); (7) that regardless of criterion selected, the assessment of teacher effectiveness is premised upon an empirical, inductive approach, which looks to (a) descriptive research for identifying and measuring dimensions of teacher characteristics, including changing characteristics, and for identifying/defining, clarifying and measuring dimensions of pupil characteristics that a teacher and teaching program is expected to cultivate, and (b) longitudinal experimental research, with adequate "before-after" measures of teacher characteristics and/or pupil characteristics to assess effects, changes, and relationships that may be inferred to result from identifiable teacher education and in-service experiences.

## IV. SOME BASIC ISSUES

In view of the extensive writing about teacher effectiveness by educational researchers, philosophers, practicing educators, teacher educators and others, this section of the report, like following sections, will make slight effort to summarize the extended discussions of topics and investigations that have been noted in the literature. Instead, attention will be called to some major issues and the reader referred to the statements and reviews cited in the accompanying bibliography.

For convenience the major issues faced in teacher effectiveness research will be categorized and briefly noted under the following rubrics: feasibility of the assessment of teacher effectiveness in research and practice; practicality of assessing teacher effectiveness; assumptions that are necessary to the assessment of teacher effectiveness; theory in the assessment of teacher effectiveness; and methodological and procedural problems in assessing teacher effectiveness.

In introducing the consideration of basic issues it seems appropriate to comment briefly on attitudes of educators and others toward the attainment of teacher effectiveness in teacher education and educational practice.

Most of the adult population recall times when there was essentially general consensus on the desirability of competent performance in a person's chosen occupational field, indeed at whatever task one undertook. Today it sometimes appears that values have changed dramatically and many people are interested less in how effectively they can perform and more in how little they can "get by" with--accompanied, of course, with maximum material reward.

Individuals who are sincerely interested in education and learning--whether pupils, teacher trainees, teacher educators, school personnel, or others--do not (it is devoutly hoped) subscribe to such a viewpoint that minimizes the importance of competency. Not that one's eyes and ears should be closed to its manifestations; it cannot be denied that as the fourth quarter of the twentieth century is approached, our lives are complicated with myriad paradoxes, often it seems, dilemmas. We come face to face with them daily in school, in economics, in churches, in government, and in the home. Whether we attribute them to the rapid expansion of knowledge, the relative affluence of the recent past, effects of wars, developments in communication and transportation, release from a sheltered and limited acquaintance with mores different from our own, a "generation gap," or other conditions, they do represent new sets of ideals and goals on the part of very vocal and sometimes sizeable groups. Whatever the underlying conditions may be, a deep involvement of personal values and apparent changes in values do seem indicated.

One set of values we sometimes feel is changing, one that is particularly closely tied to our interests as educators, is the sort of thing John Gardner alludes to in his book, Excellence; Can We Be Equal and Excellent Too? (18). A brief quotation from Gardner's book is pinned to the writer's office wall, a quotation that reads: "The society which scorns excellence in plumbing because plumbing is a humble activity, and tolerates shodiness in philosophy because it is an exalted activity, will have neither good plumbing nor good philosophy. Neither its pipes nor its theories will hold water." This allegorical comment may be quite applicable to many of our concerns in education and to the topic of effective teaching.

The area of work that has consumed a great deal of many professional lives has to do with this aspect of excellence--maximal effectiveness of teaching. Actually, the more immediate goal toward which many research efforts, including the writer's, have been directed is the description of different kinds of important teacher behavior or teacher charcteristics and the determination of conditions to which they seem to be related. But in such cases the ultimate interest obviously (as with teacher education and the teaching profession in general) is with effective teacher behaviors.

It is not entirely clear that all members of the teaching profession are equally concerned about teaching effectiveness today. Even when some feeble steps are made in the direction of describing and identifying teacher behaviors that may be effective, the profession may still be faced with the problem of inculcation in teachers of the motivation to be effective--to be as excellent teachers as they possibly can.

And we can try to train teachers to be effective and motivate them to achieve highly effective performance only to the extent that the behaviors that contribute to effective teaching under different conditions can be identified. So the matter of "effective teaching," one long discussed, seems likely to remain one of foremost importance; certainly one that needs much further research and understanding.

As we attempt to discuss research on teacher effectiveness, we are faced with a number of issues and questions that must be considered. These are the issues listed at the beginning of this section having to do with feasibility, practicality, assumptions we are willing to make, theoretical models that will guide the research, and methodological problems and procedural issues.

(1) Is it feasible to undertake research on teacher effectiveness and to try to assess teacher effectiveness?

Are the components of teaching behavior tangible and sufficiently describable to permit identification and analysis?

Are aspects of teacher effectiveness today and in the recent past in the same behavioral domain as those that may be required a

generation hence? As noted earlier, some educators predict the teacher will become more of a technical manager of the adaptation of mechanical aids to student requirements and less of a personal motivator and reinforcer, presenter and demonstrator, evaluator, and counselor/advisor. Should this be an accurate description of teacher roles in the future, is it feasible to research those behaviors that reflect the teacher's present functions?

Furthermore, does the state of current knowledge and knowledge that may emerge in the foreseeable future justify attempts at evaluating teacher behaviors or judging teacher effectiveness?

Turning to a second set of issues, (2) is it practical to pursue research on teacher effectiveness?

Even if it is possible to identify certain effective teacher behaviors, do administrative conditions in school practice preclude practical application that would permit placement of teachers where their competencies would be most useful?

A further question here--are demands of school systems, and communities within systems, so varied that no universal teacher education program would really be practical? Would the prediction of school needs and the necessary establishment of many varieties of training programs be too costly and demanding for teacher education to support?

Still further regarding practicality, in practice is it really possible to upgrade and reward effective teachers and downgrade or eliminate teachers who are clearly ineffective; can results of teacher effectiveness research be useful for purposes of supervision with teachers who already have attained tenure in a school system?

Will school systems, for political and other reasons, be averse to applying, or slow to apply, procedures that might be developed?

In connection with practicality we also must always remember that research yields results which must always be stated and interpreted in terms of "probabilities"--never can the findings be accepted with complete assurance; research on teacher effectiveness cannot yield infallible guidelines. Will employing schools accept this "principle of uncertainty" that characterizes all scientific endeavor?

Furthermore, supporters and potential users of research on teacher effectiveness must recognize that research efforts, except under unusually auspicious circumstances or an occasional stroke of serendipity, seldom yield "instant results" that can be quickly adapted to practice. Still further, it must be recognized that research on teacher effectiveness is likely to yield relatively few "universals," i.e., common means of assessing behaviors that will be

applicable to all teachers regardless of grade level, the values espoused by the community in which the teaching is accomplished, subject matter content, and the like. (It may be anticipated, however, and does seem very probable that "classes" of teacher behaviors may be identified that are agreed to be effective in certain "classes" of conditions.)

(3) The matter of assumptions is always an important issue in research. The assumptions one is willing to make determines what kinds of results he is willing to accept and what kinds of findings will not be acceptable. What are some assumptions necessary for teacher effectiveness research--and, a very important "issue" type question--are such assumptions as must be made acceptable to the education profession?

Some such assumptions involved in teacher effectiveness research and teacher effectiveness assessment include such as the following: (a) Teacher behaviors fall into identifiable "sets" or classes that are characterized by some degree of internal consistency and some degree of stability over time and across somewhat differing classroom situations. (b) Empirical study and inductive inference provide the valid approach to teacher effectiveness research. (c) Teacher behaviors can be observed--i.e., distinguishing features of classes of teacher behavior can be noted and can be recorded at some level of quantification. The behavioral aggregates that contribute to teaching performance are capable of being identified and described in operational terms. (d) Teachers differ among themselves with regard to the observable classes of behaviors they exhibit, i.e., teachers vary in the degree to which certain characteristics are represented, and a teacher who possesses a high level of "task-orientation" may possess a quite different level of some characteristic as "motivating teacher behavior." (e) Teacher behavior is social in nature--the teacher reacts to and interacts with students. (f) The ultimate goal of teacher behavior is the development of specifiable sets of pupil behaviors. (g) Teacher behavior is relative to the context in which it is performed--what may be effective in one situation may be less effective in another.

Those who are engaged in teacher education assume, of course, that behaviors of students can be shaped or molded by educational and classroom experiences--at least, to some extent. (If teacher behavior can be materially influenced by teacher education courses and programs, we still need a great deal of information, to be sure, about the more appropriate teacher behaviors to educate for, about which teacher behaviors are more, and which less, susceptible to modification, and about the kinds of educational experiences that maximize the shaping of effective teacher behaviors.)

And, of course, in dealing with the assumptions made in researching and assessing teacher effectiveness, those must be considered important issues that have to do with the kind of teacher

behavior with which we are concerned, i.e., assumptions referred to earlier which were related to whether the teacher is considered a presenter/motivator/planner, or a technician/organizer/adaptor of educational media.

(4) Still another issue the researcher faces is that of the theoretical model by which his research will be guided. Of course, one may have no particular theoretical model at all in mind and may pursue a sort of blunderbuss, or perhaps catch-as-catch-can, naturalistic approach. There is something to be said for this naturalistic kind of undertaking. (Darwin employed it to great advantage.) But theoretical models do offer certain inducements in narrowing the range of facts, serving a heuristic function, making explicit existing gaps in knowledge, and emphasizing systematic rather than haphazard gathering of information. Some current models for research on teachers are outgrowths of role theory, some of social interaction theory, some of linguistic-related theory, and many are varieties of behavioral information theory involving readiness, presentation, channeling and information flow, feedback and reinforcement, decision-making, etc. The theoretical position from which one proceeds is an important issue in conducting research on teacher effectiveness, and indeed often underlies other of the major issues referred to here.

(5) Still another very large set of issues that affect research on teacher effectiveness has to do with methodology to be employed. Definition of the dependent variable, i.e., criterion behavior, is a matter usually accorded far too little serious attention in much behavioral research. Here we become concerned with distinctions between "assessment" and "evaluation," with process and product criteria, with what Rosenshine and Furst (55) and Gage (Rosenshine's early collaborator) refer to as "high inference" and "low inference" type variables, with descriptive, correlational, and antecedent-consequent types of research (the present writer, for one, feels we have a great deal of work still ahead in satisfactorily describing and identifying the important behavioral variables involved in teaching, the determining of how behaviors aggregate and form relevant behavioral compounds, and the distinguishing of relevant from insignificant or inapposite teacher behaviors), with correlational research versus controlled experimental research, with the reliability of measures of independent and dependent variables, with sampling and control (both of populations of teachers and of classes of behaviors), with multivariate analysis procedures--to mention some of the many methodological issues.

Although only tangentially concerned with "issues," comment is appropriate here on research dealing with "teachers in training" in relation to research dealing with "teachers in service." When we undertake research on teacher effectiveness we usually elect to pursue our study with either one or the other population; but seldom both teachers-in-training and teachers-in-service. The two are, of course, inseparable in many respects. But research efforts

usually are directed at one or the other; and, to date, very few investigations that have been begun in teacher education have followed through with study of the same teachers on-the-job. At least one reason for this, it may be presumed, is that some greater degree of constancy can be maintained in "teacher training research," permitting greater control over measures of the dependent variables and even allowing variation in experimental treatment of different groups. In school practice, actual experimentation with adequate control over experimental, concurrent independent, and dependent variables is difficult--particularly with variation in the students, objectives, and expected teacher behaviors in different schools and communities. For this reason it may be noted that most of the in-service studies (as well, of course, as many of those of teachers-in-training) have been correlational in nature, and of the few experimental researches that have been undertaken those that appear to be most meaningful have been conducted in teacher education institutions. (The writer hastens to add that even these have been woefully few.)

## V. TEACHER EFFECTIVENESS RESEARCH IN PERSPECTIVE

An effort has been made to avoid the usual redundancy of reviews by providing a fairly complete, albeit short (some eighty-nine titles), set of references with this report, accompanied by relatively little comment. There may be a few conspicuous omissions and some relatively inconsequential titles may have been included. But most of the important discussions of issues and research on teacher effectiveness, particularly the relevant reviews of literature, have been included.

The present section attempts very little summarization, generalization, or editorializing on the literature discussed. The only exception will be brief reference to some of the most recent reviews, e.g., the capable critiques of Rosenshine, with a few additional comments on current efforts in progress.

For the reader concerned less with details and largely with the general state of knowledge today about teacher effectiveness research, attention is directed to (in alphabetical order): Flanders (12), Getzels and Jackson (19), Jackson (29), Mitzel (44), Pi Lambda Theta (46), Ryans (56, 57, 60, 61, 63), Stiles and Parker (74), among others.

Theoretical considerations have received increasing attention; for such the reader is referred to: Biddle (2, 3), Fattu (7), Gage (13, 15), Gage and Unruh (16), Gordon (21), Henderson (23), Hyman (27), Maccia (36), MacDonald and Leeper (37), Ryans (64), Siegel (66), Smith (69), Stolurow (75), Travers (77, 78), Turner (83), among others.

A special and particularly important problem that perhaps cuts across "theory" and "methodology" is that of the criterion, or dependent variable in teacher effectiveness research (actually more correctly noted as "criteria" rather than connoting any simple "criterion"). A number of references are devoted to this important issue upon which all interpretation of research hinges: Medley (39, 40), Mitzel (44), Pi Lambda Theta (46), Rosenshine and Furst (55), Ryans (57, 61), Stolurow (75), Turner and Fattu (87, 88), among others.

For many years little attention was given the complex methodological and procedural problems that perplex the researcher; fortunately, this issue is now widely recognized and an increasing number of persons are turning to attempts to reduce the error variance by improving the methodology of research: Biddle (2), Gage (14), Getzels and Jackson (19), Hodgson and Horst (25), Hughes et al. (26), Kerlinger (32), Kounin et al. (34), Mastin (38), Medley (39, 40), Rosenshine (48, 49, 50, 52), Rosenshine and Furst (55), Ryans (56, 57, 61), Skinner (67), Smith (70), Soar (71, 72), Wallen and Travers (89), among others.

A number of the reports appearing in the literature are essentially, or in part, substantive and have to do with descriptions

of teacher behaviors, teacher characteristics, or aspects of teacher-pupil interaction. These include, among many, many others, Bellack (1), Flanders (8, 9, 11), Gallagher (17), Getzels and Jackson (19), Hiller et al. (24), Hodgson and Horst (25), Kerlinger (30, 31, 32), Kerlinger and Pedhazur (33), Kounin et al. (34), Mastin (3 ), Medley (39), Medley and Mitzel (42), Rosenshine and Furst (55), Ryans (57, 58, 59, 61, 62, 63), Schmid et al. (65), Smith et al. (70), Torrance and Parent (76), Travers (77), Travers et al. (79), Turner (80, 81, 82, 85), Turner and Denny (86).

Some of the most searching critical delving into research on teacher behavior and teacher effectiveness that has come to the writer's attention has been the comprehensive reviews of Barak Rosenshine (48, 49, 50, 51, 52, 53, 54, 55).

Rosenshine has made a comprehensive search for recent studies--both correlational studies of teacher behavior and its association with "student products" and controlled experimental research with particular, although not exclusive, emphasis upon "low inference" teacher performance criteria.

Rosenshine and Furst (55) describe and critique process and product studies with regard to five teacher variables that appear to have strong support from correlational research of teacher performance as related to student achievement.

These five variables for which Rosenshine finds strong support when measured against student achievement are: clarity; variability/adaptability; enthusiasm/stimulation; business-like/task-oriented behavior; and provision of student opportunity to learn.

Six additional variables are noted which, although lacking in strong research support, Rosenshine feels may be worthy of future study: use of student ideas/teacher indirectness; use of criticism; use of structuring comments; use of multiple levels of discourse; probing; and student-perceived difficulty of the course.

A large number of what appear to be non-significant variables that have been researched by various methods also are noted by Rosenshine.

Rosenshine and Furst (55) also provide a number of suggestions for improved procedures in future process-product, i.e., teacher performance compared with student performance research on teaching. The criterion problem (stressed earlier in this report), of course, is emphasized, i.e., the importance of identifying and describing relevant teacher behaviors and relevant criteria of effective teaching.

Rosenshine suggests a paradigm he believes appropriate for experimental study of the effects in teacher education programs on training teachers to exhibit specified classroom behaviors or to elicit specified student behaviors; the paradigm requires that:

(1) some teachers be trained to teach a class of students in a certain manner (e.g., emphasis upon analysis, use of extensive student participation, encouragement of student enthusiasm, etc.); (2) observational measures be obtained to verify that teachers in the training groups really behaved as intended during the experiment; and (3) relevant end-of-experiment student measures be obtained. This is, of course, a very sketchy summary of the paradigm around which one would build an experimental design.

In even the "better-designed experiments" Rosenshine found that training teachers to use certain processes during teacher education programs has not usually resulted in measurable improvement in student achievement--at least, as measured. He observes that the inconclusive and insignificant findings in this area may suggest we have focused on the wrong teacher behaviors (see our earlier remarks on the need for identifying relevant aggregates of teacher behavior), that we have not observed the relevant and important processes that really affect learning, and that we should look with special care and in a systematic, objective, empirical manner for the relevant processes that are "critical" as opposed to the trivial.

The writer wholeheartedly agrees; while not forsaking controlled experimental research, for the time being (as earlier emphasized) there is still much to be accomplished in inventorying the observable aggregates of teacher behavior that it seems reasonable to hypothesize may contribute to teacher effectiveness.

Actually a number of distinguishable classes of teacher behavior have been identified to date. True, we have not determined their relationships to student achievement in most instances. Yet to seek and discover the characteristics of teacher behavior and their organization in different kinds of teachers, teaching situations, and subcultures is a very important end in itself.

In view of the importance of relevant teacher behaviors, a few of the rather widely confirmed behavior aggregates of teachers involving behavior styles and other characteristics of teachers that have emerged to date will be noted briefly.

Considerable evidence has accumulated suggesting a pattern of behavior identified by the Teacher Characteristics Study research some 20 years ago as "organized, systematic, responsible, business-like" teacher behavior, i.e., the behavior Rosenshine refers to as "task-oriented/business-like" behavior. (Incidentally Rosenshine did find evidence that this behavior seemed to be strongly related to student achievement.) Another of the major factors emerging from the Teacher Characteristics Study (a pattern which several other studies support) was "achievement-motivating, stimulating, imaginative" teacher behavior--which Rosenshine refers to as "enthusiastic" teacher behavior. This, too, Rosenshine finds related to student achievement.

Still another of the major factors indicated, by the Teacher Characteristics Study, had to do with "considerate, understanding, warm, supportive" teacher behavior. (Rosenshine's reviews do not reveal much evidence of this kind of behavior actually being related to currently measured student outcomes.)

Still another factor the Teacher Characteristics Study results suggested, but which was not pursued to any great extent, seemed to be associated with "expressive, attractive, clear, personally and academically impressive" teacher behavior; a number of studies seem to support the existence of such a pattern.

And another facet of teacher behavior that has received considerable attention is the dimension referred to in the Teacher Characteristics Study as "traditional, teacher-centered, academic viewpoints" versus "permissive, child-centered viewpoints." Kerlinger (30, 31, 32) and Kerlinger and Pedhazur (33) have studied such attitudinal responses of teachers in depth and find three factors, the two more prominent of which are (a) one that seems to reflect "progressivism" and which Kerlinger says "seems related to Ryans' $X_o$ teacher behavior pattern" (the warm, sympathetic, understanding pattern), and (b) a factor judged to portray the "traditional teacher" which Kerlinger states "seems to be closely related to Ryans' $Y_o$ teacher behavior pattern (organized, systematic, task-oriented, business-like). Flanders has employed his rather complex "interaction analysis" observational procedures to the study of these "indirect" (student centered) and "direct" (teacher centered) behaviors and has developed a detailed procedure for systematizing observations. Rosenshine concludes that the evidence of relationships between (a) assessments of facets of teaching behavior reflected in "interaction analysis" and (b) student achievement is not at present convincing but probably warrants further study.

In some recent factor analytic studies (not yet reported in the literature) the writer finds major patterns of teacher characteristics identified in the Teacher Characteristics Study (at that time, 1948-54, employing a different approach involving extensive systematic classroom observation by trained observers, prior to development of the Teacher Characteristics Schedule) to be supported by large scale factor analyses of teacher's self-report inventory responses. (Two new factors from preliminary and partial analysis of the 1970-71 data also appear.) The factors thus far identified in the updated (1970-71) Teacher Characteristics Study are:

1. --warm, sympathizing teacher behavior
2. --business-like, task-oriented teacher behavior
3. --original, motivating teacher behavior
4. --attitude toward pupils and other persons contacted in schools
5. --academic focused educational viewpoints

6. --permissive, pupil-centered educational viewpoints
7. --verbal/semantic facility in language in which teaching is accomplished
8. --social adjustment
9. --commitment/dedication to teaching as a profession
10. --teaching-associated activities, i.e., participation in cultural, community, and similar activities frequently expected of, and often attractive to, teacher.

It is interesting to note that Turner has mentioned to the writer evidence that he, too, has found what seems to be a "professional involvement," or "career motivation" pattern of teacher characteristics. (See No. 9 in foregoing list.)

These studies are being extended to a number of teacher subcultures in the United States and plans are also underway for cross-national and cross-cultural studies in other countries. In the cross-cultural studies there are a number of difficult methodological problems to overcome, but the results may throw light on important problems of differences in teacher behavior that accompany differences in culture, as well as common patterns across different cultures. These may provide important aids to teacher education. Studies in Hawaii (of teachers of different ethnic background) have been completed, as has similar research on Australians from Victoria.

Closely related to research the writer is pursuing on teacher characteristics in general and on their cross-cultural similarities and differences is somewhat similar work dealing with the "life views" or value patterns of teachers who also have responded to the Teacher Characteristics Schedule. Unfortunately the writer's own attempts in this area have not proven as meaningful as the similar large scale factor analyses of Teacher Characteristics Schedule responses. Admittedly this "values" or "life views" area is a difficult one to dissect and estimate. Perhaps because of the extreme uniqueness of each person's "values," as well as our proneness to inconsistency of specific behaviors within "values" categories, the writer's data thus far have yielded little more than what may be an acquiescence factor--essentially a clustering (intercorrelation) of responses representing a number of different value domains (at least, as those domains are considered from the viewpoint of rational validity). The only interpretation the writer can make of the complex of values that load on a large general factor in these studies is that the factor may reflect the teacher's indoctrination during his teacher education days and points of view he has learned are expected of him as a member of the teaching profession.

## VI. SOME RECOMMENDATIONS

A variety of recommendations was made by members of the committee. Many were of a very specific nature regarding methodology. Others were related to such topics as grade levels, locales, and similar conditions that might make for greater convenience, as well as better experimental control, in undertaking the study of teacher effectiveness. Still others were in terms of broad objectives, e.g., study of teacher activities "today" and in the near future as contrasted with an engineering approach to understanding the predictors of effectiveness that might be fitting with respect to teachers fifty years hence. Some were philosophically oriented. Almost all recommendations encompassed the need for systematic theory. Some were in terms of workable school-university relationships and collaborative study. Many of the recommendations were in no sense antithetical with regard to others, but it was difficult to fit the lot of them into a single pattern that could be recommended to the U. S. Office of Education for the encouragement of a major undertaking to encourage teacher evaluation among the schools.

Several large scale programs, which may be considered as either alternative plans or parallel plans, are here suggested.

One study recommended to the Office of Education is an outgrowth of at least two major research projects conducted by members of the committee. It is here recommended by the Chairman of the Committee and preparer of this report; it is hoped concurrence of other members may also support it.

Results of the Teacher Characteristics Study led to conviction of the study director (the current report writer) that behaviors of teachers that are judged "effective" vary with a number of conditions, including educational viewpoints (i.e., values held regarding education by different communities), grade levels or subject matters involved, geographical area, socio-economic conditions, etc.; and that the proper development of means for identifying teacher effectiveness, therefore, must take into account the so-called "teacher effectiveness research paradigm" frequently referred to by the writer over the past 15 years.

To accomplish the goal of the paradigm, the expectations of persons to whom teachers are responsible must be ascertained with regard to their conception of characteristics contributing most significantly to teacher effectiveness. These conceptions appear to vary so greatly from one community to another that no pretense should be made at trying to determine any "universals" that relate to teacher effectiveness. This was the position taken by the Teacher Characteristics Study and by the current research on the Teacher Characteristics Schedule. The Teacher Characteristics Schedule provides estimates of some of the dimensions of behavioral and attitudinal responses of teachers that may relate to objectives sometimes adopted by teacher education programs and by the schools--objectives that persons belonging to one or another

group may associate with effective teaching in their own minds. If a teacher education program espouses objectives that place a premium upon "pupil-centered viewpoints," or "warm, sympathizing teacher behavior," some of the scales provide estimates of such characteristics. If another teaching program places emphasis upon an "academic-centered set of educational viewpoints," upon "task-oriented teaching behavior," etc., some of the scales tap these areas.

From rather long experience, the writer has learned that although some educators give lip service to universal marks of teacher effectiveness, there actually is far from complete agreement among communities, as well as among teacher education programs, regarding just what the individuals concerned want teachers to be like. As long as diverse opinions exist, the best we can do in determining whether a teacher education program (or the teacher objectives of a school) may be achieving what it intends to achieve, is to get the program planners to spell out their objectives as behaviorally as possible and then develop research designs that will provide information about the extent to which a specified program is nurturing teacher characteristics that relate to the chosen objectives.

It should be observed, also, that other researchers share this view.

Based upon a series of researches he has conducted, F. N. Kerlinger (30, 31), a member of this committee, urges, that in trying to determine what is considered effective teaching, one must recognize that any person's opinion about good or poor teaching is a reflection of his basic educational orientation and the underlying criteria operating to predispose his opinion must be known for effective teaching to be identified.

Similarly R. L. Turner (81, 82), another member of this committee, has reported notable variations in what school districts with different supervisory organizations expect and value with respect to primary compared with elementary teachers; also, that the emphasis of supervisory personnel on teacher "task performance" and "ability to organize" (compared with person-oriented, permissive teacher behavior) materially increases as the proportion of working-class to middle-class schools in a school system increases.

In view of ample research evidence, as well as a logic, supporting the premise that characteristics of the teacher that may be judged "good" by one person in one community or at one time, may not be similarly viewed as "good" by another person, another community, or at another time, it is proposed that a first major undertaking, preceding any substantive teacher effectiveness research projects directed at instrumentation and experimentation, should be concerned with identification of the major sets of teacher-related expectations and values represented in American education or a specified segment of American education, today.

I. Very briefly, one recommendation is that such a basic study might be conducted which would involve: (1) generation of a limited number of statements relative to school-related values or expectations and roles of teachers, reflecting some of the basic issues noted in this report and including others that may seem relevant; (2) preparation of sets of cards, each card bearing one of the teacher/school-related statements; (3) selection of a stratified random sample of teachers, superintendents and principals, school board members, and parents representing selected and identifiable geographical areas, kinds of communities, etc.; (4) participation of a sample of such respondents, each respondent performing Q-sorts (one with respect to elementary teachers and elementary school practices and the other with respect to the secondary teacher and secondary school practices), arranging the cards in piles representing what the responding individual feels to be the relative importance of each statement with respect to effective teaching; (5) conduct of factor analyses of the Q-sorts (with referent data included) to determine major patterns or factors; (6) identification of the factors generated with respect to expectations, roles, and values that define different teacher effectiveness emphases in different situations; (7) determination of which patterns of teacher effectiveness are associated with which kinds of communities, which kinds of school settings, which kinds of persons (e.g., parents, teachers), etc.

With this type of basic information it would be possible for different school systems to more readily identify "teaching value contexts" that characterize them, and to proceed through the "teacher effectiveness research paradigm" that the writer has frequently proposed. It also would make it possible for researchers, independent of the complete paradigm, to develop instruments for estimating behaviors related to different teacher effectiveness patterns, as well as to investigate the myriad of interrelationships of conditions and factors suggested in the list of "basic issues." Still further, teacher education institutions undoubtedly would find the baseline information of substantial utility value both in planning their teacher education programs and also as a guide to obtaining criterion data for validation of teacher education efforts.

The "teacher effectiveness research paradigm" referred to in the last paragraph is a rather straight-forward, step-by-step approach to teaching in any, or many, specified situations. The approach probably has seldom been employed because of the "research" required in each particular school system employing it, but it is one the writer has long felt could do much to enlighten us in the field of education--e.g., for purposes of teacher selection, placement, supervision, and promotion in situations where it would be possible to put applications of findings into practice. Above all it would provide individual guides to teacher education if its results were observed and put into practice. One version of the paradigm (it has differed from time to time only in the number of steps included) lists some 10 phases, as follows:

(1) Selection and designation of general aspects of the value system framework of the school/community as those aspects relate to teacher behavior. The agreed-upon qualities that are desired, or expected, of teachers in a particular place and in particular kinds of teaching situations are referred to. (Note that this process of arriving at criteria necessarily is subjective and a matter of the values that individuals or groups of individuals may possess in common. When we designate criteria we proceed from a context of an accepted value system. We view teacher behavior in light of a set of attitudes, opinions, and viewpoints that reflect the sorts of teacher behavior we approve and prefer, and also the kinds of behavior we disapprove and find unacceptable. To the extent any group of persons share in common certain expectancies, preferences, or biases about teachers and teaching, criteria of teacher behavior may be defined for that particular group.) When the objectives emphasize (as many of us believe they should) cognitive learning, the criteria may include student products or student achievement--"low inference" criterion variables. But there are also numerous confounding conditions that may make it difficult to isolate the effects of a particular teacher on student performance--and we recognize these methodological problems as well as the "face-validity" of student product criteria. In any event, we must agree upon what general behaviors we expect of teachers in specified situations.

(2) Identification of observable properties of teacher classroom behavior that may be related to the specified operationally described criteria (i.e., the descriptive cataloguing of teacher characteristics and behaviors and significant aggregates of teacher behavior that occur in the classroom).

(3) Identification of kinds of situations in which the agreed-upon "valued teacher behaviors" may occur--and in which they may be observed and assessed.

(4) Operational description (i.e., description in terms of actual teacher behaviors) of the agreed-upon "valued behaviors" that are to comprise the criteria of teacher behavior and reflect effective teaching.

(5) Selection of methods of estimating the operationally (i.e., behaviorally) described "valued behaviors." This is the problem of instrumentation relative to the criterion behavior and obtaining assessments of the criterion behaviors. In assessing some aspect or characteristic of the criterion behavior of teachers we are trying to estimate the extent to which that defined characteristic is manifested by some teacher.

(6) Development of instruments and procedures that are hypothesized to yield estimates that will reflect the operationally described teacher behaviors (criterion behaviors)--which, in turn, are assumed to reflect the value framework of the school and community served.

(7) Assembly of data yielded by the teacher assessing instruments and procedures noted in Step 6 above.

(8) Assembly of data yielded by the procedures used to estimate the criterion behaviors--Step 5 above.

(9) Analysis of relationships between estimates of the behaviorally defined criterion behavior and the estimates of teacher characteristics used in the teacher assessment procedure.

(10) Finally, evaluation of teacher effectiveness by drawing of inferences about the validity of those procedures that were employed to measure teacher behavior for predicting the criterion behaviors designated in Steps 1 through 3 above.

When these steps have been taken and the paradigm followed through to its conclusion so that knowledge of the extent to which reliable relationships exist between "valued behaviors" and "observable teacher behaviors," then the necessary groundwork has been laid for evaluation and in-service evaluation and teacher education evaluation can be accomplished objectively and successfully.

II. A second procedure, rather closely related to the one just mentioned is also highly recommended. This recommended procedure admittedly has grown from the writer's interest in cross-cultural comparisons of teachers and indeed considers different component groups of teachers of a single national or ethnic grouping as simply representing subcultures. Thus secondary mathematics teachers represent a subculture and are not directly comparable with respect to effective teacher characteristics to, say, teachers of grades one and two. Similarly, male and female sex groups represent different subcultures. And certainly different communities within a large school district, different districts throughout a state, different sections of the country, and different ethnic-heritage groups represent often quite different subcultures of teachers.

Drawing upon a still broader cross-cultural context the writer has noted the many difficulties involved in comparing teachers of different nations, customs, etc. In the cross-national context, he has observed that we may think of at least two major classes of variables that must be taken into account in any cross-cultural comparisons:

(1) those variables which may be thought of as manifest or "surface" variables--simply those effects that are easily distinguished--and (2) those variables that may be thought of as "latent" or "generic" and underlying psychological variables that influence behavior but which are often not readily observable.

Within the major category of "manifest" variables, "structural language" variables are concerned with symbols employed, combinations of symbols representing objects and processes, syntax, and the like; these are identifiable and very clearly must be dealt with in cross-cultural study. (Certainly we have some of these same problems when we are studying different subcultures within the same nation.) When words employed in different languages have common roots the problem of translation, and back-translation, often is not too difficult to overcome. But unfortunately for cross-cultural research many verbal expressions evade direct and straight-forward translation from one language to another and are either unique to a particular language or perhaps set of colloquial expressions, or else vary in meaning of translatable forms across different groups. National and racial background characteristics of social and psychological (as well as political, economic, religious, etc.) nature also often represent reasonably easily identified variables (frequently represented to some degree by language forms) that might be similarly classified under the rubric of "surface" variables. Such variables include readily recognized and operationally definable customs, traditions, mores, ideals, value systems, procedural practices, etc., that differ from culture to culture and subculture to subculture. Like language, these variables also pose problems for researchers who attempt to determine common bases for comparing personality characteristics of different cultural groups.

It should be noted that while the context from which we are speaking is that of cross-national and cross-racial cultures, we face some of the same kinds of problems among subcultures of teachers within a given country. And, of course, another kind of "manifest" or "surface" variable is represented by subcultures, per se, and their attributes (e.g., in the case of teachers as members of an occupational subculture, still more specific distinguishable subcultures identified through age, experience, level of instruction, subject matter taught, sex, geographic area within a particular nation or culture, etc.). In comparing major cultures, these subcultural variables are important concerns both because of their significance as separate variables worthy of consideration as dependent variables and also as conditions or independent variables that must be taken into account in sampling for the comparison of major cultures of which they are components.

The second major grouping referred to was the "latent" variables or "generic" variables. These are less readily recognized variables, often insidious, yet representing extremely potent effects and sources of variation, and presenting particularly difficult problems for

research. Here the variables involve relativity of meanings, mores, values, etc., in different cultural and subcultural contexts. In this category could be included semantic variables, a sizeable array of characteristics that relate to the representativeness of the respondent or experimental subject as an individual typifying his subcultural group in comparisons made with members of other subcultural groups, and the like. Contaminating biases of this nature are difficult to determine and equate. Inevitably it seems cross-cultural personality research also is faced with respondent biases relating to "reservation" or "caution" in verbal response and/or overt behavior (as contrasted with "openness" and "frankness") when comparisons with other groups are known to be involved, with respondent concern about what he perceives as invasion of the researcher upon personal behavior of the individual, and with other respondent biases that may invalidate responses or lead to nonresponse. In a somewhat similar context are "researcher" or investigator biases that may confound cross-cultural studies. Variables of this sort involve the quality and extent of participation on the part of collaborating investigators in different subcultures or cultures; sometimes because of attitudes toward practices involved in the role of investigator in a particular subculture, or because of lack of assurance on the part of a collaborating investigator that he really "belongs to the team" and actually shares in participation in decisions and credits, or because of other similar conditions that bear upon maintenance of comparable conditions of data gathering in the subcultures being compared.

Thus far, reference has been made only to some of the more readily discernible variables that cross-cultural research must take into account and nothing has been said of methodology for attempting to equate, balance out, partial out, or otherwise control relevant variables in order to make possible the study of similarities and differences in personality dimensions across subcultures. Such methodology poses a major stumbling block. Applicable designs are difficult to conceive and if a viable rationale is formulated it is often difficult and costly in time and resources to achieve.

In cross-cultural research one seldom can start with any advance information of similarity of content; indeed there is recognition that similar appearing content may have dissimilar meanings in different subcultures and dissimilar appearing content may have similar meanings. The pollsters and market researchers are keenly aware of this. Attempts to define and designate criteria for determining responses that contribute to a particular domain across populations differing in expressions of "surface" and "latent" variables represent another example of this dilemma. And, how does one determine what behavior a response reflects, or how valid a single response or set of responses comprising a "score" may be in one or another culture or subculture? The crux of the question of cross-cultural equivalence or comparability seems to rest on the assumptions that can be made and procedures that can be employed relating to the concept of validity.

The difficulties involved in determining behavior and/or response equivalence across subcultures seems almost impossible to resolve; and these difficulties are multiplied for the behavioral scientist by the fact he is dealing with psychological constructs which seldom are invariant (with regard to meanings attached to a particular construct, e.g., persistence) within a given subculture. And, construct invariance necessarily increases as one moves from one subculture to another.

If there be procedures that are applicable to the determination of the personality patterns that are discernible across different cultures and the validity of response content in reflecting personality dimensions in differing cultures, they would seem to be those of the form of "construct validation" sometimes referred to as "intrinsic validation" or "factorial validation," i.e., validity determination by correlational and factor analyses. Here we attempt to establish meanings from contexts in which responses occur--a very necessary procedure for the understanding of psychological constructs and for determining the approximate equivalency of constructs and the equivalency of response content that reflects the constructs across cultures. In general, then, the procedures that may be used in studying the different subcultures of teachers in the United States involve large scale factor analyses applied to (a) judgments and (b) self-report inventory responses relating to preferred activities, typical behaviors, opinions, and the like.

The second approach, therefore, that the writer is recommending as one that schools in the United States could well apply to the study of teacher effectiveness and the invariance, or lack of universality of meaning, of teacher effectiveness in one community as compared to another is essentially one of approaching each of a representative sample of communities as a subculture. The recommendation with this respect is that the program of research to be followed would be such as that noted below, with the phases as indicated:

(1) Determination in each community or subculture of those teacher characteristics that are generally valued in teachers in a particular subculture. This would be essentially a "judgmental" process, perhaps employing a "critical incident approach" (or possibly using a generalized "Q-sort" procedure, or some form of latent partition analysis--or possibly two or more such approaches in combination);

(2) Development in each community or subculture of operational definitions of those teacher characteristics that seemed most significant in the light of phase one;

(3) Development in each subculture or community of inventory type (self-report) items that could be hypothesized to estimate the major patterns of teacher characteristics revealed in phases 1 and 2;

(4) Incorporation of items from each subculture or community into a single inventory or "schedule" that could be administered in all subcultures or communities involved;

(5) Translation of the single overall inventory, insofar as possible, including different dialects and colloquialisms as well as common language structure (perhaps placing directly translatable items together and placing those items that were not universally translatable together in a subsection);

(6) Administration of the inventory to representative samples of teachers in each subculture or community;

(7) Factor analyses of the responses for (a) polled data for all participating subcultures or communities and (b) data from each separate subculture or community separately;

(8) "Factor matching" to determine (a) unique and (b) common factors;

(9) For common factors (a) development of a suitable procedure (and this introduces a most difficult problem) that might yield indices of "response equivalence" for differing responses that contribute to apparently similar factors in different subcultures or communities, and (b) development of suitable procedures (again, a most difficult problem) for determining norms comparability from one subculture or community to another with respect to common-appearing factors.

The writer, and the committee charged with the responsibility, respectfully submit this report to the U. S. Office of Education and urge serious consideration of the recommendations for study of the problem of teacher effectiveness, looking towards major study following one of the sets of recommendations suggested--a study that might provide answers to problems that, in the past, have appeared insoluble. It is firmly believed by the committee that the empirical study of teacher effectiveness is possible, that teacher effectiveness can be assessed, and that methods of analyses are available for making possible instruments that could provide assessments of teacher behavior which would conform to the qualities desired by a particular community or school. The relativity of "teacher effectiveness" is thus recognized, but procedures are suggested that might make possible the encouragement of the evaluation of teacher effectiveness without demanding that all schools and communities subscribe to a common set of values.

REFERENCES


1. Bellack, A. A., et al. The language of the classroom. New York: Columbia University Press, 1966.

2. Biddle, B. J. Methods and concepts in classroom research. Review of Educational Research, 1967, 37, 337-357.

3. Biddle, B. J. Teacher roles. In R. L. Ebel (Ed.), Encyclopedia of educational research. New York: Macmillan, 1969. Pp. 1437-1446.

4. Christensen, C. M. Relationships between pupil achievement, pupil affect-need, teacher warmth, and teacher permissiveness. Journal of Educational Psychology, 1960, 51, 169-174.

5. Cook, D. L., et al. A factor analysis of teacher trainee responses to selected personality inventories. Educational and Psychological Measurement, 1961, 21, 865-872.

6. Dickson, G. E., et al. The characteristics of teacher education students in the British Isles and the United States. U. S. O. E. Cooperative Research Project 2518. Toledo, Ohio: University of Toledo, 1965.

7. Fattu, N. A. A model of teaching as problem solving. In J. B. MacDonald & R. R. Leeper (Eds.), Theories of instruction. Washington, D. C.: Association for Supervision and Curriculum Development, 1965. Pp. 62-87.

8. Flanders, N. A. Analyzing teacher behavior. Educational Leadership, 1961, 19, 173-75, 178-180, 200.

9. Flanders, N. A. Analyzing teacher behavior. New York: Addison-Wesley Co., 1970.

10. Flanders, N. A. Interaction analysis in the classroom: A manual for observers. Minneapolis: Author (College of Education, University of Minnesota), 1960. (mimeo).

11. Flanders, N. A. Some relationships among teacher influence, pupil attitudes, and achievement. In B. J. Biddle & W. J. William (Eds.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart & Winston, 1964. Pp. 196-231.

12. Flanders, N. A. Teacher effectiveness. In R. L. Ebel (Ed.), Encyclopedia of educational research. New York: Macmillan, 1969. Pp. 1423-1437.

13. Gage, N. L. (Ed.) Paradigms for research on teaching. Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 94-141.

14. Gage, N. L. Teaching methods. In R. L. Ebel (Ed.), Encyclopedia of educational research. (4th ed.) New York: Macmillan, 1969. Pp. 1446-1458.

15. Gage, N. L. Theories of teaching. In Ernest R. Hilgard (Ed.), Theories of learning and instruction. Sixty-Third Yearbook, Part I, National Society for the Study of Education. Chicago: University of Chicago Press, 1964. Pp. 268-85.

16. Gage, N. L., & Unruh, W. R. Theoretical formulations of research on teaching. Review of Educational Research, 1967, 37, 358-370.

17. Gallagher, J. J. Analyses of teacher classroom strategies associated with student cognitive and affective performance. U. S. Department of Health, Education, and Welfare, Office of Education, Project No. 3325, Contract No. (OE-6-19-196). Urbana: University of Illinois, 1968.

18. Gardner, J. W. Excellence; can we be equal and excellent too? New York: Harper, 1961.

19. Getzels, J. W., & Jackson, P. W. The teacher's personality and characteristics. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 506-582.

20. Gilbert, H. B., & Lang, G. Teacher selection methods. Final report. U. S. Department of Health, Education, and Welfare, Office of Education, Project No. 6-1665. New York: Board of Examiners of the City of New York, 1967.

21. Gordon, I. J. (Ed.) Criteria for assessing the formal properties of theories of instruction. Washington, D. C.: Association for Supervision and Curriculum Development, 1968.

22. Gunnison, J. F. An experiment to determine the effects of changing teacher classroom behavior through training of student-teachers in the use of Flanders' interaction analysis system. Unpublished doctoral dissertation, Arizona State University, Tempe, Arizona, 1968. University Microfilms No. 68-14, 994.

23. Henderson, K. B. A theoretical model for teaching. School Review, 1965, 73, 384-391.

24. Heller, J. H., et al. A computer investigation of verbal characteristics of effective classroom learning. American Educational Research Journal, 1969, 6, 661-677.

25. Hodgson, T. F., & Horst, P. The general and primary factors in student evaluation of teaching ability. American Psychologist, 1959, 14, 374.

26. Hughes, M. M., et al. Development of the means for the assessment of the quality of teaching in elementary schools. U. S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 353. Salt Lake City: University of Utah, 1959.

27. Hyman, R. T. (Ed.) Contemporary thought or teaching. Englewood Cliffs, N. J.: Prentice-Hall, 1971.

28. Jackson, P. W. The teacher and the machine: Observations on the impact of educational technology. Mimeographed paper presented to the Committee on Economic Development, November, 1966. Chicago: Department of Education, University of Chicago.

29. Jackson, P. W. (Ed.) Teacher personality. Theory into Practice, 1963, 2, 57-112.

30. Kerlinger, F. N. Attitudes toward education and perceptions of teacher characteristics: A Q study. American Educational Research Journal, 1966, 3, 159-168.

31. Kerlinger, F. N. The factor structure and content of perceptions of desirable characteristics of teachers. Educational and Psychological Measurement, 1967, 27, 643-656.

32. Kerlinger, F. N. The first-and second-order factor structures of attitudes toward education. American Educational Research Journal, 1967, 4, 191-205.

33. Kerlinger, F. N., & Pedhazur, E. J. Educational attitudes and perceptions of desirable traits of teachers. American Educational Research Journal, 1968, 5, 543-559.

34. Kounin, J. S., et al. Explorations in classroom management. Journal of Teacher Education, 1961, 12, 235-246.

35. Lindvall, C. M., & Bolbin, J. O. "Programmed instruction in the schools: An application of programming principles in 'individually prescribed instruction.'" Programmed Instruction. Sixty-Sixth Yearbook, Part II, National Society for the Study of Education. Chicago: University of Chicago Press, 1967. Pp. 217-254.

36. Maccia, E. S., et al. Construction of educational theory models. Cooperative Research Project No. 1632. Columbus, Ohio: Ohio State University Research Foundation, 1963.

37. MacDonald, J. B., & Leeper, R. R. (Eds.) Theories of instruction. Washington, D. C.: Association for Supervision and Curriculum Development, 1965.

38. Mastin, V. E. Teacher enthusiasm. Journal of Educational Research, 1963, 56, 385-386.

39. Medley, D. M. The development of classroom behavior dimensions for student teachers and pupils. Mimeographed paper presented to the American Association for the Advancement of Science, December, 1962. New York: Division of Teacher Education, City University of New York, 1962.

40. Medley, D. M. The language of teacher behavior: Communicating the results of structured observations of teachers. Journal of Teacher Education, 1971, 22, 157-165.

41. Medley, D. M., & Mitzel, H. E. Measuring classroom behavior by systematic observation. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 247-328.

42. Medley, D. M., & Mitzel, H. E. Some behavioral correlates of teacher effectiveness. Journal of Educational Psychology, 1959, 50, 239-246.

43. Medley, D. M., & Mitzel, H. E. A tentative framework for the study of effective teacher behavior. Journal of Experimental Education, 1962, 30, 317-320.

44. Mitzel, H. E. Criteria of teacher effectiveness. In C. W. Harris (Ed.), Encyclopedia of educational research. (Rev. ed.) New York: Macmillan, 1960. Pp. 1481-1486.

45. Morsh, J. E. Systematic observation of instruction behavior. Development Report No. AFPTRC-TN-56-52. San Antonio: Air Force Personnel and Training Research Center, Lackland Air Force Base, 1965.

46. Pi Lambda Theta. The evaluation of teaching. Washington, D. C.: PLT, 1967.

47. Rosencranz, H. A., & Biddle, B. J. The role approach to teacher competence. In B. J. Biddle & W. J. Ellena (Eds.), Contemporary research on teacher effectiveness. New York: Holt, Rinehart & Winston, 1964.

48. Rosenshine, B. Enthusiastic teaching: A research review. School Review, 1970, 78, 499-515.

49. Rosenshine, B. Evaluation of instruction. Review of Educational Research, 1970, 40, 279-301.

50. Rosenshine, B. Experimental classroom studies of teacher training, teacher behavior, and student achievement. Mimeographed "draft copy" of paper presented at 1970 meeting of the National Council for the Social Studies. Urbana, Illinois: Barak Rosenshine, University of Illinois.

51. Rosenshine, B. Interpretative study of teacher behaviors related to student achievement. USOE Project No. 9B-010. Philadelphia: Temple University, 1970.

52. Rosenshine, B. The stability of teacher effects upon student achievement. Review of Educational Research. 1970, 40, 647-662.

53. Rosenshine, B. Teaching behaviors and student achievement. Stockton, Sweden: International Association for the Evaluation of Educational Achievement, 1970 (mimeo).

54. Rosenshine, B. To explain: A review of research. Educational Leadership, 1968, 26, 303-309.

55. Rosenshine, B., & Furst, N. Research on teacher performance criteria. In B. O. Smith (Ed.), Research on teacher education: An AERA symposium. New York: Prentice-Hall, 1971.

56. Ryans, D. G. Assessment of teacher behavior and instruction. Review of Educational Research, 1963, 33, 415-441.

57. Ryans, D. G. Characteristics of teachers. Washington, D. C.: American Council on Education, 1960.

58. Ryans, D. G. An information system approach to theory of instruction with special reference to the teacher. SP-1079. Santa Monica, Calif.: System Development Corporation, 1963.

59. Ryans, D. G. Inventory estimated teacher characteristics as covariants of observer assessed pupil behavior. Journal of Educational Psychology, 1961, 52, 91-97.

60. Ryans, D. G. Possible directions for teacher-behavior research. Theory into Practice, 1963, 2, 105-12.

61. Ryans, D. G. Prediction of teacher effectiveness. In C. W. Harris (Ed.), Encyclopedia of Educational Research. (Rev. ed.) New York: Macmillan, 1960. Pp. 1486-1491.

62. Ryans, D. G. Some relationships between pupil behavior and certain teacher characteristics. Journal of Educational Psychology, 1961, 52, 82-90.

63. Ryans, D. G. Teacher behavior can be evaluated. Educational Horizons, 1967, 45, 99-120.

64. Ryans, D. G. A theory of instruction with special reference to the teacher. Journal of Experimental Education, 1963, 32, 191-223.

65. Schmid, J., et al. An analysis of some necessary qualities of teachers. U. S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 453. Fayetteville: University of Arkansas, 1962.

66. Siegel, L. (Ed.) Instruction: Some contemporary viewpoints. San Francisco: Chandler Publishing Co., 1967.

67. Skinner, B. F. The technology of teaching. New York: Appleton-Century-Crofts, 1968.

68. Skinner, B. F. Why we need teaching machines. Harvard Educational Review, 1961, 31, 377-398.

69. Smith, B. O. A concept of teaching. Teachers College Record, 1960, 61, 229-241.

70. Smith, B. O., et al. A study of the strategies of teaching. Cooperative Research Project No. 1640. Urbana: University of Illinois, 1967.

71. Soar, R. S. Methodological problems in predicting teacher effectiveness. Journal of Experimental Education, 1964, 32, 287-291.

72. Soar, R. S. Multivariate statistical procedures in predicting teacher-pupil classroom behavior. U. S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 1170. Columbia: University of South Carolina, 1962.

73. Stiles, L. J. Revolution--in instruction. A decade of thought on teacher education. The Charles W. Hunt lectures, 1960-1969. Compiled by Esther Hemsing. Washington, D. C.: American Association of Colleges for Teacher Education, 1969. Pp. 33-53.

74. Stiles, L. J., & Parker, R. P. Teacher education programs. In R. L. Ebel (Ed.), Encyclopedia of educational research. (4th ed.) New York: Macmillan, 1969. Pp. 1414-1423.

75. Stolurow, L. M. Model the master teacher or master the teaching model. In J. D. Krumbolts (Ed.), Learning and the educational process. Chicago: Rand McNally, 1965. Pp. 223-247.

76. Torrance, E. P., & Parent, E. Characteristics of mathematics teachers that affect students' learning. U. S. Office of Education Cooperative Research Project No. 1020, Contract No. SAE-8993. Minneapolis: Minnesota School Mathematics and Science Center, Institute of Technology, University of Minnesota, 1966. ERIC ED: 010 378.

77. Travers, R. M. W. On the transmission of information to human receivers. In J. B. MacDonald & R. R. Leeper (Eds ), Theories of instruction. Washington, D. C.: Association for Supervision and Curriculum Development, 1965.

78. Travers, R. M. W. Towards taking the fun out of building a theory of instruction. Teachers College Record, 1966, 68, 49-60.

79. Travers, R. M. W., et al. Measured needs of teachers and their behavior in the classroom. U. S. Department of Health, Education, and Welfare, Office of Education, Cooperative Research Project No. 444. Salt Lake City: Department of Educational Psychology, University of Utah, 1961.

80. Turner, R. L. Beginning teacher characteristics and beginning teacher problems: Some predictive relationships. Paper presented at 1966 meeting of the American Education Research Association, Chicago, Ill.

81. Turner, R. L. Characteristics of beginning teachers: Their differential linkage with school system types. School Review, 1965, 73, 48-58.

82. Turner, R. L. Comparison of teacher characteristics of inner city (working class) and outer city (middle class) teachers falling in upper vs. lower quarters of principals' ratings. (Personal communication, 1966)

83. Turner, R. L. Conceptual formulations of research in teacher education. In B. O. Smith (Ed.), Research on teacher education: A symposium. New York: Prentice-Hall, 1971.

84. Turner, R. L., et al. Skill in teaching assessed on the criterion of problem-solving: Three studies. Bulletin of the School of Education (Indiana University), 1963, 39, 1-31.

85. Turner, R. L. Some predictors of problems of beginning teachers. Elementary School Journal, 1967, 67, 251-256.

86. Turner, R. L., & Denny, D. A. Teacher characteristics, teacher behavior, and changes in pupil creativity. Elementary School Journal, 1969, 69, 265-270.

87. Turner, R. L., & Fattu, N. A. Skill in teaching, a reappraisal of the concepts and strategies in teacher effectiveness research. Bulletin of the School of Education (Indiana University), 1960, 36, 1-40.

88. Turner, R. L., & Fattu, N. A. Skill in teaching assessed on the criterion of problem-solving. Bulletin of the School of Education (Indiana University), 1961, 37, 1-32.

89. Wallen, N. E., & Travers, R. M. W. Analysis and investigation of teaching methods. In N. L. Gage (Ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 448-505.