DOCUMENT RESUME

ED 164 517 95 SP 013 682

AUTHOR Turner, Richard L.; And Others
TITLE Interpreting Research Results on Teaching.
INSTITUTION ERIC Clearinghouse on Teacher Education, Washington, D.C.
SPONS AGENCY National Inst. of Education (DHEW), Washington, D.C.
PUB DATE Jan 79
NOTE 52p.

EDRS PRICE MF-$0.83 HC-$3.50 Plus Postage.
DESCRIPTORS *Analytical Criticism; *Educational Research; *Effective Teaching; *Research Design; Research Methodology; Research Reviews (Publications); Research Utilization; Teacher Behavior; *Test Interpretation

ABSTRACT

This publication is a collection of commentaries on a published monograph entitled "Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research" by Donald M. Medley (1977). Medley's stated purpose was to provide teacher educators with access to the meaningful findings of research in teacher effectiveness. The papers here comment on Medley's methodology and procedures; consider the inferences/implications for teacher education of the findings reported in Medley's review; and predict what Medley's approach portends for the future of research and selection of topics for further investigation in teacher effectiveness. (MM)

Commentary on Medley's
Teacher Competence and Teacher Effectiveness:
A Review of Process-Product Research

INTERPRETING RESEARCH RESULTS ON TEACHING

Reactive Papers by
Richard L. Turner - Robert M. W. Travers - Judy Lanier -
Barak Rosenshine - Marianne Murphy - George E. Dickson - Lee Morris

Response by
Donald M. Medley

Published by the
ERIC Clearinghouse on Teacher Education
One Dupont Circle, Suite 616
Washington, D.C. 20036

January 1979 SP 013 682

The ERIC Clearinghouse on Teacher Education is funded by the National Institute of Education, in cooperation with the following associations:

American Alliance for Health, Physical Education,
and Recreation
American Association of Colleges for Teacher Education
Association of Teacher Educators
National Education Association

CONTENTS

INTRODUCTION

In 1977 the American Association of Colleges for Teacher Education (AACTE) published a monograph entitled Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research.[1] The monograph was commissioned by the AACTE Committee on Performance-Based Teacher Education for the stated purpose of presenting "interim findings that can be used to improve teacher education now while we wait for the researchers to produce definitive results" (p. iii). The author, Donald M. Medley, worked with a researcher's panel created by the Committee to analyze and synthesize results of the voluminous, sometimes contradictory, body of research studies about teacher competence and effectiveness.

Medley began with a bibliography of 289 studies reporting thousands of relationships between teacher behaviors and pupil learnings. By applying four stringent criteria, he reduced the number of relationships to 613, all of which were derived from just 14 of the 289 studies. Medley then devised a unique format for providing his readers with direct access to the statistically most significant and dependable findings: he categorized the correlations in a series of 42 tables so as to leave the interpretation of those findings to the reader. (In illustration, Table 1 appears on p. 4; readers are referred to the monograph itself for the other tables as well as Medley's commentary.)

Thus the major portion of the Medley report consists in these correlational tables, preceded by a description of the criteria and methodology employed and followed by a section on future directions. As an example, the author also drew several inferences from the findings; he stressed, however, that readers should not confuse his interpretations with the facts on which they were based.

The critical importance of interpreting research results prompted the Clearinghouse on Teacher Education to solicit critiques of this report from a number of education researchers and teacher educators. Those responding were asked to comment briefly on Medley's methodology and procedures; to consider inferences/implications for teacher education of the findings reported; and to predict what this approach portends for the design of research and the selection of topics for investigation in teacher effectiveness. These commentaries are included in the present publication, as are a summary of the introduction to the Medley report and subsequent reappraisal by Medley.

It is to be hoped that this attempt to stimulate discussion of promising approaches to the translation of research into practical knowledge for teachers and teacher educators will encourage readers of the present publication to seek out and study the original Medley report, and will provoke further comment and reaction.

1 Donald M. Medley. Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research. Washington, D.C.: American Association of Colleges for Teacher Education, August 1977. ED 143 629

THE MEDLEY REPORT

These excerpts from the Medley report summarize its major points, as a context for the individual commentaries that follow. The interested reader is exhorted, however, to make reference to the original document for a complete understanding of the issues debated.

It is the primary purpose of this report to provide the teacher educator with access to the meaningful findings of research in teacher effectiveness. . . .

Efforts to develop performance-based programs both for educating and certifying teachers have made it painfully clear just how inadequate the base is for what we know today about the dynamics of teacher effectiveness. These efforts have also demonstrated how weak the connection is between research in teacher effectiveness and the teacher education curriculum. . . .

A number of sound scholarly reviews of this literature have appeared in recent years. . . . These are invaluable; but the reader comes away with the feeling (not really justified) that there has been access, not to the research itself, but to a synthesis or interpretation of that research. . . . In this project we have made a strong effort to put the reader in direct contact with the research. We have canvassed the literature and culled the most significant findings from it, without attempting to select or interpret them or to reconcile them with each other. We have then presented them in a series of tables in a particularly simple format. Readers of this report are invited to examine the process-product correlations presented in the tables and draw their own conclusions. These tables were designed to communicate the most clearly established facts about effective teaching and only those facts.

It is important to bear in mind certain limitations. . . . First, we have presented only the strongest and most dependable findings, ignoring both small correlations that are statistically significant, and larger correlations that are not. The fact that a relationship is not reported should not be taken as evidence that it does not exist, or even as indicating that there is no research evidence that it exists. Absence of a relationship . . . means only that its existence has not been clearly established as far as we can discover. . . . The second limitation has to do with generalizability. For reasons mainly connected with the funding strategy of the U.S. Office of Education, most of the research summarized here was done in one segment of the school population--in classes of Grade III or below in which most of the pupils come from homes of low socioeconomic status. To what extent these findings apply to pupils with other backgrounds or in other grades is not known. What evidence we have about pupils of high socioeconomic status and pupils in the higher grades indicates that results from one group do not always apply to another.

(Medley then discusses "Dangers of Misinterpretation," and considers the strategy, popular in some quarters, of "Waiting for Definitive Results Before Doing Anything." The author rejects this stance, saying, "Very few decisions worth making can be put off until there is adequate information to base them on. . . . Educators must make decisions every day, regardless of the availability of hard evidence on which to base them. With this need in mind, we have proceeded."

PROCEDURE

The basic bibliography of this study consisted of 289 studies [referenced in the Appendix] which purported to shed light on the question, "How does the behavior of effective teachers differ from that of ineffective teachers?" These studies were the survivors of a weeding-out process from an original list of 732 items. Most of the 445 rejected items were rejected because they reported no original research; some were reviews of research; others theoretical, philosophical, or opinionated discussions (from the armchair) of what a good teacher ought to do.

The remaining 289 items were examined for empirically obtained relationships between how a teacher behaves and how much the pupils learn from him or her, commonly called process-product relationships. Four criteria were used in deciding whether or not a reported relationship should be included in this review. Only those which met all four criteria were included. Briefly, the criteria were:

1. The study from which a relationship came had to be designed so that the relationship was generalizable to some population of teachers larger than the sample studied.

2. The relationship had to be both reliable enough to be statistically significant and large enough to be practically significant.

3. The measure of teacher effectiveness had to be based on long-term pupil gains in achievement areas recognized as important goals of education.

4. The process measure had to specify the behaviors exhibited in such a way that they could be reproduced as desired.

By the time we had applied these criteria to the thousands of reported relationships between teacher behaviors and pupil learning reported in the literature, the number of relationships which survived was 613; and these 613 correlations all came from just 14 of the 289 studies. . . .

Rationale of the Study

All four criteria proceed logically from a point of view adopted in this study which is at variance with that underlying most of the research reviewed. This viewpoint may be described briefly as follows:

The ultimate base of teacher education curriculum must be a thorough understanding of the dynamics of effective teaching--of what a teacher must know, and be, and do, in order to provide the greatest possible assistance to pupils in their efforts to achieve the goals of education. Such understanding depends on the establishment of cause-and-effect relationships between teacher behavior and pupil learning. Only when we know why a teacher is effective--as well as how--can we decide how best to train teachers. The recognized purpose of research in teacher effectiveness is to develop such an understanding by discovering the cause-and-effect relationships from which this understanding may be derived. . . .

What is the proper course of action for the teacher educator to follow while waiting for the researcher to develop this knowledge bit by bit? Are there interim findings of the research that can help teacher educators do a

Table 1
PUPIL INITIATIONS

BEHAVIOR [illegible]	GRADE	LOW SES PUPILS			HIGH SES PUPILS			SOURCE SYMBOL
		reading gains complexity low high	arithmetic gains complexity low high	affective gains [illegible]	reading gains complexity low high	arithmetic gains complexity low high	affective gains [illegible]	
Pupil-initiated vs. teacher-initiated interchanges	I		L^1			L^1		WGC OScAR
Pupil-initiated interaction vs. response to teacher	I-II	L^1	L^1					S73 RCS1
Pupil initiates substantive interchange	II		L^1			L^1		WGC OScAR
Pupils speak freely	II	L^1			L^1			WGC OScAR
Pupil task-related comments to adults	III		L					SK 388a
Pupil questions, requests, commands--non-academic	III		L L					SK 477c, 346a
All non-responsive pupil utterances to adults	III		L L					SK 343a
Pupil initiates substantive interchange	III-VIII		H^1 H^2			H^1 H^2		WGC OScAR
Pupil volunteers information vs. pupil asks for information	III-VIII	H^1			H^1			WGC OScAR
Total pupil-initiated contacts	IV					H H		GG
Pupil-initiated vs. teacher-initiated substantive interchange	IX-XII	H^1			H^1			WGC OScAR

better job--or must we wait until the researcher is satisfied that the findings are definitive before touching them? This project was undertaken under the assumption that interim results can be useful. The nature of the research is such that it generates information that is currently useful to teacher educators as is. . . . This is the point of view we have used in this survey of the literature--that process-product research can tell us quite a lot about how competent and less competent teachers differ in their classroom behavior, even though we may not know exactly why.

A strong relationship between a behavior variable and a measure of teacher effectiveness need not be regarded as evidence that the observed behavior caused the measured effect. Instead we shall use the measure of effectiveness as an indicator of teacher competence, inferring that teachers who are effective are more competent on the average than teachers who are ineffective. The distinction between competent and effective implied in this statement is important and yet easy to forget. Competence has to do with how a teacher teaches and is measured in terms of the teacher's behavior; how effective a teacher is is measured in terms of pupil learning. In other words, an effective teacher is always competent, but a competent teacher may not always be effective, for a multitude of reasons. . . .

(The author next explains the rationale for each of the four criteria applied to the process-product relationships considerd for inclusion in the report. He then describes the format of the tables.)

Notes on Interpreting the Tables

Table 1 illustrates the format in which the 613 relationships are displayed. . . .

The table title . . . is meant to identify a common element in the process measures listed at the left under the heading "Behavior Item." These are identified where possible by the actual item or category name used in the study; or when the name was not descriptive, a brief descriptive phrase is employed.

At the right of the list of behaviors is a column indicating the grade level or levels of the classes in which the behaviors were observed. At the extreme right, under the heading "Source Symbol," are codes identifying the studies from which the relationships reported for each behavior item came; when available, the number assigned to the item in the actual instrument is also included. . . . The studies are listed by code in the Appendix, with details about sample, instrumentation, and the like.

Each letter in the body--L, M, or H--identifies a strong relationship; and the location of the letter identifies the two variables related: to the left is the behavior or process variable; above is the teacher effectiveness or product measure.

The Structure of Teacher Competence

When independent relationships between a single behavior and two distinct kinds of teacher effectiveness are reported in a study, we have what will be called a pair of relationships. . . . If the two relationships in a pair match--if, for instance, both are L's . . .--the implication is that teachers competent in the two different ways tend to behave alike. If they do not match--if, for example, one is reported L and the other H--the

Table 2
PERCENTS OF PAIRS OF PROCESS-PRODUCT RELATIONSHIPS
BETWEEN THE SAME BEHAVIOR AND TWO TYPES OF OUTCOME MEASURES
THAT MATCH (I.E., HH, LL, OR MM)

RELATIONSHIPS PAIRED	Number of Pairs	Percent Matching
Attitude Toward School vs. Achievement Gains (same SES level)	54	72
Gains in Self-Concept vs. Achievement Gains (same SES level)	36	75
Reading Gains vs. Arithmetic Gains (same level of complexity and same SES level)	80	73
High Complexity vs. Low Complexity Gains (same SES level)	158	91
Gains in High SES Classes vs. Gains in Low SES Classes (same subject and level of complexity)	84	38

indication is that teachers competent in one way behave in an opposite manner from teachers competent in the other way. By examining all such pairs of relationships in [the tables], we can get some idea about the structure of competent teacher behavior. In other words, we can learn something about which behaviors are generic, in the sense that they are equally effective for different objectives and with different kinds of pupils.

(Medley comments on the findings from the pairing of relationships as shown in Table 3.)

In summary, the evidence is that, with early grade pupils of the same SES level, the teacher who produces maximum achievement gains in either reading or arithmetic is quite likely to produce high gains in both subjects and at both levels of complexity (as defined in this paper), and in the pupils' self-concept and attitudes toward school as well. There is also evidence (from one study) that competent teachers of low SES pupils behave quite differently from competent teachers of high SES pupils.

(The author continues with illustrative interpretation of the tables, under the headings "The Competent Teacher of Low SES Pupils in the Primary Grades" and "Teacher Competence and Pupil SES in the Primary Grades." He stresses, however, that "the facts shown in the tables are the primary product of this study, and we would prefer that the value of the study be judged according to their usefulness rather than on the merits of our interpretations of them.")

FUTURE DIRECTIONS FOR PROCESS-PRODUCT RESEARCH

. . . Under the assumption that the goal of research in teacher effectiveness is to strengthen the knowledge base for teacher education, there are at least two steps that should be taken. One involves a change in priorities--in the way in which process variables are chosen for study; the other involves a change in strategy--a modification in the model itself. And the implementation of these two steps implies closer collaboration between the teacher educator and the researcher than we have seen in the past. The knowledge and resources that each possesses must be brought together in a unified effort.

New Priorities for Research in Teacher Effectiveness

One consequence of the lack of such collaboration in the past can be seen by comparing a list of the teacher behavior (process) variables studied by the researchers with the list of competencies that define the objectives of a competency-based teacher education program. By and large, the researchers do not seem to be studying the teacher behaviors that the educators regard as important. There is overlap, but the lists are far from congruent.

This has two implications, both of them bad. First, we lose the important contribution the teacher educator could make to selecting for study those teacher behaviors likely to characterize effective teachers. And second, the results the researcher gets would have much more direct implications for teacher education if they involved the competencies

teacher education programs seek to develop directly. Their relevance would be obvious; and negative findings (which, alas, are far more common than positive ones) would be almost as useful as positive findings. . . .

The way out of this situation is for future process-product research to use as process variables the same competencies that the teacher education programs are trying to help teachers acquire. Researchers and teacher educators should get together to investigate the validity of the latter's program goals. If the information thus developed were used as the basis for program revision, and if the researcher continued to study the revised objectives in the same way, the result would amount to a large-scale, continuous experiment in teacher education. Such an experiment--or better yet, a number of them linked to several teacher education programs--would have at least two important effects. First, it would directly improve the effectiveness of the program studied. And second, it would add as much to our understanding of the dynamics of effective teaching--or more--than any amount of the one-shot, process-product research that is the present norm. . . .

(A discussion of "Levels of Assessment in Teacher Education" follows.)

Future Strategy for Process-Product Research

. . . It is perhaps time to consider whether the process-product model as we know it may not have outlived its usefulness, if for no other reason than that it ignores two critical variables almost completely: the intent of the teacher, and the behavior of the individual pupil. . . . Somehow, in future research in teacher effectiveness, we must find and use a model in which the teacher's intent or purpose and the behavior of the individual pupil both play a part.

Perhaps the answer will involve the description or assessment of teacher purpose, meaning the learning experiences teachers intend their pupils to have; and instead of correlating teacher behaviors with outcomes, we will correlate teacher behaviors with pupil behaviors. The competent teacher would be the teacher who can behave in such a way that pupils have the learning experiences the teacher intends them to have--prescribes for them, if you will.

There is a second component in teacher competence, of course: the prescribed learning experiences must be those that maximize pupil learning outcomes; the competent teacher must, then, be able to diagnose pupil needs and recognize what each pupil needs to do in order to learn.

Teacher competence thus involves a knowledge component--knowledge of relationships between pupil behaviors and learning outcomes; and a performance component--the ability to act, to behave, in ways that will help pupils exhibit these behaviors, have these learning experiences. Research in teacher effectiveness might split, then, into two phases: the study of teacher behavior in relation to pupil behavior, and the study of pupil behavior in relation to pupil learning outcomes. It would seem much more productive if the principal focus of future research were on correlations between teacher behaviors and pupil behaviors, that is, between competencies and learning experiences, rather than on correlations between teacher behaviors and outcomes. . . . At the same time, the secondary focus should be on correlations between pupil behaviors and pupil outcomes. Our understanding of the dynamics of effective teaching should increase much more rapidly if this strategy were to be adopted. . . .

RICHARD L. TURNER

Dean, School of Education
University of Colorado

The question a reader of any research report must ultimately answer is the amount of confidence he or she is willing to place in it. When confidence is high, the reader is willing to act in accord with study outcomes. When confidence is low, action is withheld and doubts seriously entertained.

In the Medley review, two questions involving one's confidence can be raised. The first is whether the results are trustworthy or reliable; the second is which actions one would be willing to take on the basis of the data or outcomes provided. These questions are examined here.

HOW TRUSTWORTHY ARE THE FINDINGS?

Confidence in the trustworthiness of the procedures or methodology used in the study rests on whether or not one believes the results can be replicated. The question of replicability enters first in the selection of studies to be reviewed. About five percent of the studies (14 out of 289) reviewed by Medley entered the report; the remainder were rejected against the criteria stated. Initially, this selection process did not bother me, but later I became uneasy over the possibility that the statistically significant correlations cited by Medley in his sample of studies appear in other studies as nonsignificant correlations. If they do, an issue of the replicability of the data cited by Medley as "reliable" is raised.

The second point at which replicability as the central indicator of reliability enters is in the between-study replication of teacher behavior items. A notable aspect of the tables is that each teacher behavior item entered is identified with a single study and typically with a single observation instrument; there appears to be no replication of the same behavior item between studies. This places the reader in the position of having to base confidence on the statistical significance of a correlation in a single study. The evidence presented for replicability of significant behavior items between studies on the same instrument is virtually nonexistent.

Theoretically, Medley's procedure of grouping items into categories and presenting each category (such as group size, seatwork, number of teacher questions) as a table could ease the question of the replicability of behavior items between studies. This would happen if each category presented were interpreted as a homogeneous factor, and each behavior item within a category were interpreted as an individual item loading significantly on the factor. The fact is, however, that this interpretation cannot legitimately be made. One simply does not know how items placed in the same category but drawn from different studies and different instruments would correlate, hence provide the basis of a homogeneous factor and serve as evidence of the replicability of similar-appearing items within categories.

The whole point above seems technical, but I find it bothersome to my confidence in the data presented in the review. The significant data in this study lie in the correlations (the L's, the M's, and the H's) in the body of the tables. These tell which relationships between teacher

behavior and pupil achievement or affect are supposed to be trustworthy. But if one does not know how the items within and between the categories presented relate to each other, it is virtually impossible to make a coherent interpretation of the data. The reason is this: The behavior items shown in any one category may in fact be correlates of other items shown in other categories. If this is true, the correlations distributed among the tables cannot be shown to be independent. Thus what appears to be a significant relationship between one behavior item and achievement may simply be a "rider" on another behavior item which is also a correlate. Moreover, if two dissimilar-appearing items are in fact correlates, the suspicion arises that some other behavior, not observed in the studies, is the key or basic behavior producing the correlation.

In reviewing the evidence available within the report bearing on the question "How trustworthy are the findings," I regret to say that I am considerably less optimistic than Professor Medley about "strong and reliable relationships" being reported in the study. My belief is that only a few of the behavior items given will hold up against a criterion of replicability and, therefore, that treating the set of items presented as entirely trustworthy is largely an act of faith. It is, however, no greater an act of faith than treating the teacher competencies in most performance-based teacher education programs as having any demonstrable validity.

WHAT ACTIONS MIGHT BE TAKEN ON THE BASIS OF THE FINDINGS?

Three main classes of action might be taken by teacher educators relative to the findings of the review: (a) ignore them, (b) utilize the behavior items in teacher training programs, (c) utilize them in assessment of teacher classroom actions. I am going to eliminate the first class since the findings generate sufficient confidence to make it unwise to assign them to the limbo of forgotten library shelves.

Training

There are a number of problems in utilizing the individual behavior items presented in the report in training programs. A key set of these problems lie in the joint relationships between the context in which teaching occurs and the standards or norms applicable to specific items. It is obvious that each research study reported by Medley dealt with distributions of teacher-pupil "process" behaviors. These distributions vary from some low frequency to some high frequency for each behavior. When the correlation between teacher-pupil and pupil product measures are positive, the greater the frequency or amount of the teacher process behavior and the greater the pupil product.

The problem of norms or standards in training teachers on items of specific behavior is to fix the range of permissible frequency in a teacher behavior to assure the claimed pupil product will occur. For example, in one of the tables the item "Teacher lectures, pupils bored" appears as a strong positive correlate of pupil gains in reading at both SES levels and for both levels of reading complexity. The question is, how much teacher lecturing and how much pupil boredom must be present to assure the optimum learning outcomes? If a literal interpretation is given to the reported linear relationship between lecture/boredom and pupil gains, teachers who lecture all of the time and bore pupils into perpetual sleep would get the

best learning. The idea sounds attractive but, intuitively speaking, one suspects it to be absurd.

If training is to be exact, training to the empirically sound or "right" norm on specific teacher behaviors is a critical matter. If the training is inexact and the preparatory or inservice teacher learns to emit the wrong frequencies of a particular behavior, ineffective teachers will presumably be produced. They will have the right behaviors but the wrong frequencies of these behaviors.

The key issue here is how exact training should be. Should it be exact to the point of the teacher's having to deliver a behavior within a particular range of frequency? I think not. What one seeks, rather, is to have preparatory teachers know about what appear to be significant classes of teacher behavior. These classes would be analogous to the types of categories utilized by Medley. The problem for the teacher is then to learn how to control the frequency of behaviors within these classes according to the context in which teaching occurs.

That context is a controlling factor may be readily observed in certain tables of the report, in which what are presumably similar teacher/pupil behaviors with similar frequencies have opposite effects in the primary and intermediate (and middle school) grades. As is made clear in the report, the socioeconomic status of pupils as well as the subject taught must also be taken into account as a contextual variable.

In sum, if one regards the classes or categories of teacher/pupil behavior presented by Medley as trustworthy (and their trustworthiness is not demonstrated in the report), a key issue to be dealt with in teacher training is to establish, in preparatory teachers behavioral control over the frequency with which the specific behaviors are engaged in according to particular teaching contexts. The standards against which such behaviors are to be judged are not only their trueness to the categories but also the appropriateness of the frequency of the behavior to its teaching context. Establishing a fixed or very exact standard for the frequency of a behavior would not, in this interpretation, be a useful idea.

To use the data in the report along the lines suggested here, the best psychological assumption is probably that the type of training involved is skill acquisition. Under this assumption, the first training step is for the student to learn the concepts underlying the categories. These concepts need to be learned in such a way that the concrete instances of them are very clear. Thus, if the concept is "pupil initiation" (the first table in the report), concrete instances of the behavioral elements in the concept class--as represented, for example, in the individual behavior items shown in the table--need to be seen by the student, either by means of "protocol" materials or directly in live classrooms. To repeat an earlier point, it is in part for training reasons that the categories or concept classes used by Medley, and the items in them, need to be homogeneous and reliable (or true and trustworthy). Otherwise, mistraining can occur.

The second training step is of course to practice the behavior in the category. In at least some categories, a critical consideration is to practice varying the frequency of particular behaviors, preferably in varying contexts. To learn from this practice, students should be obtaining not only feedback cues from their pupils, but systematic feedback from the trainers. To give systematic feedback, trainers need (a) a systematic means of observing behavior, (b) norms for what constitutes high, middle, and low frequencies of particular teacher behaviors, and (c) knowledge of

the specific relationships between these levels of frequencies and pupil learning in the different classroom contexts, so that trainees can be alerted to those frequencies which best promote achievement in a particular context.

Assessment

The action that most needs to be taken on the basis of the current findings is instrument development. As suggested earlier, the categories of behavior used by Medley, and the behavior items assigned to them, have an ad hoc quality. This quality arises from having to generate categories to cover items in a variety of studies using different instruments and somewhat different data-gathering procedures. A useful action, from both a training and a research/evaluation perspective, would be to consolidate the behavioral items currently showing significant correlations with pupil achievement into a reduced and more manageable set of observational instruments.

Such an action would have certain dangers, as will be noted later, but it could also produce several benefits. First, if the instruments were used in subsequent research, a stable set of categories and intercorrelations among category items could be produced. This would help considerably in increasing the trustworthiness of subsequent findings. Second, use of a more nearly standard set of instruments would make possible some norming of the behaviors observed by means of the instruments. Additionally, experimental studies of the effects of clear deviations from normative teacher behaviors would be facilitated. Third, supervisory personnel would have a means of providing systematic feedback to preparatory teachers and interns, as well as feedback to programs about the efficacy of training. Evaluation studies within and between institutions would also be facilitated, as would school system studies aimed at diagnosing teacher inservice needs. The benefits thus seem to be several.

The dangers in developing a more nearly standard set of instruments based on currently significant findings is the omission of all other teacher behaviors from the sample taken by means of the standard instrument set. It is transparent in the Medley report that the significance of a particular behavior item is contingent on the context in which it is observed and the specific pupil product criterion with which it is correlated. An expansion of criterion variables or a shift in contextual variables can thus change which behavioral items are to be regarded as significant. An attendant danger is that a standard set of instruments based on current findings would tend to converge teacher training programs on the teacher behaviors which seem to produce the most convergent "basic skill" pupil achievements.

Although I believe these dangers to be real, they can be avoided to some degree by the development and use of instruments which sample new dimensions of teacher behavior and new aspects of pupil products. The Medley report clearly presents the best currently available set of hypotheses about which teacher actions are correlates of pupil learning. I do not regard these hypotheses to be "facts" in the usual sense, but they are excellent guides to the arena in which greater accomplishments in research on teaching and in teacher training can be attained.

ROBERT M. W. TRAVERS

Distinguished Professor
Western Michigan University

The Medley report is a strange combination of modern scientific sophistication, folklore, and Madison Avenue methods of selling--with the three often intertwined and tangled to the point where they cannot be separated. The testimonial letters reproduced at the end of the report represent a strange contrast with the attempt found in Medley's own contribution to remain close to the data. Are we supposed to think more of the findings because they are endorsed by luminaries, some of whose findings are included in the report? The inclusion of the testimonial letters suggest that we should. Are we supposed to accept without question that teachers who do things that produce high test scores in reading and arithmetic are really working for the good of the children? Modern folklore asserts that whatever produces a high test score represents good teaching, though careful reflection and analysis related to the whole problem of what teaching is and how it should be evaluated have long abandoned such an idea. Does this latter view imply that the kinds of studies that have been included in the review are worthless? No, not at all, but it does imply that the results have no direct implication for practice, even though they may have value for the construction of a theory of teaching.

I cannot possibly agree with Medley that one of his major conclusions "fairly leaps from these pages." Indeed, the conclusion didn't even seem to wiggle like a worm out of the pages, and should probably not have been drawn at all. The conclusion in question is that "where sufficient effort and resources have been applied to the study of teacher effectiveness, useful and dependable findings emerge." The conclusion raises serious questions which will have to be considered in the pages that follow.

A "SYNTHESIS" OF RESEARCH?

Any document that claims to be a synthesis of research must surely aim to put together research into concepts that present a simplified picture of the nature of phenomena. The term "synthesis" implies such a putting together of facts into a simplified structure. All the great syntheses of research do this. The ancient Egyptians put together a mass of facts and came to the surprising simple conclusion that the Earth was round. Lamarck put together the established facts of animal and plant structures and concluded that complex species must have evolved from simpler species. All significant syntheses of knowledge that have had any utility have started with a large body of quite unorganized but well-established facts, and have put them together to provide a relatively simple picture of some corner of nature. Yet the Medley attempt to synthesize research on teacher effectiveness achieves no simple picture of some corner of nature. The Medley attempt to synthesize research on teacher effectiveness achieves no simple understandable picture of the nature of teaching. To a great extent the report starts with a large mass of quite disorganized facts and ends with a large body of quite disorganized facts. One cannot necessarily criticize the research worker for this failure to achieve a useful synthesis, but one wonders why a synthesis was not, or could not be, achieved.

The main difficulty in synthesizing the kind of research included in the review derives from the fact that each research study, in itself, represents little more than an assemblage of motley facts, and sometimes a search for motley facts. The latter is a part of a recent atheoretical trend which has tended to view effective teaching as a motley collection of acts, referred to as competencies. Little theory is available concerning how these acts are integrated into a system. Indeed, the implication to be derived from much that is said and written in this area is that the competencies are independent and distinct and can be learned one by one. This is an antitheoretical view of behavior, in that behavioral scientists have already evolved quite useful theories of how skill elements are integrated and the kinds of integrations that can occur. The teacher competency model is basically an atheoretical model of teacher behavior. Pupil learning in terms of specific components of a prescribed subject matter also reflects a similar atheoretical stance which operant psychologists have so much admired.

The lack of any significant or discernible theoretical position with respect to teacher behavior or teaching has made the task of developing a useful synthesis a difficult, if not impossible, task. Some kind of synthesis with theoretical implications may be possible of the collection of facts, but not a synthesis that can have practical implications, such as other great syntheses of scientific facts have been able to effect in other areas of science.

A FOLKLORE CONCEPTION OF TEACHING

Let us examine further the ideational base of the studies included in the report and attempted synthesis. The studies derive from what can be properly called a folklore conception of teaching, which cannot be considered in any way a scientific conception. It involves the idea that the core of education is the teaching of "basics" to pupils and that the success of the teaching process can be measured by giving periodic tests. Education is viewed as a simple routine of teaching and testing, a conception of teaching that has long dominated the public view of education. During those periods when it has been taken most seriously by teachers, their efforts have become more and more directed toward training the children to obtain good scores on tests.

The teachers of Victorian England did just that; they were paid in terms of the achievement of pupils as measured by tests administered by Her Majesty's Inspectors of Schools. For each pupil who scored above the grade norm on each test of basic skill, the teacher's salary was raised by an increment above the subsistence level. Teachers did, of course, teach to produce the best test scores they could possibly obtain from their students, and probably manifested many of the behaviors listed as those of effective teachers in the Medley report. But the system was a complete disaster and after nearly half a century reduced education in England to ruin. Eventually, after the situation had been examined and condemned by more than one Royal Commission, education authorities came to realize that there is more to education than instilling basics and then testing for them. By the turn of the century, the system was completely abandoned. Nevertheless, the damage had been done, and the English elementary education system remained the worst in Europe, if not in the world, for another half-century.

The studies on which the Medley report are based seem to be saturated with the same philosophy of education that brought ruin to education in Victorian England. It is hardly surprising that the description of the effective teacher that emerges fits quite well with the 19th century popular description of what the effective teacher should do. A reader of the studies or the report is surprised to find that there is no mention of any category of teacher behavior such as teaching children how to answer test questions. It would surely have turned out to be one of the vital teacher competencies identified by the researchers. This should not be viewed as a facetious suggestion. A teacher interested in the objective of helping pupils to become avid consumers of printed materials would not be criticized for making every reading lesson a fun-filled exploration of some book, and for encouraging children to engage in voluntary reading. A teacher interested in helping children obtain good test scores should not be criticized for helping children to learn to answer test questions. However, there is a contrast. The one objective is trivial, while the other could be of considerable significance.

The results of the Medley report would have been helpful to the teachers in Victorian England, since they were vitally interested in obtaining good test scores from their children. Of course, the report comes a century late to help the Victorians, but it does come at a time when the traditional folklore about the nature of teaching is showing a great revival in many American communities. The Medley report will be viewed as providing a substantive basis for that revival. It will be viewed as research support for the idea that elementary education involves the instilling in children of certain skills called basics by traditional techniques. Parents concerned with test scores, and little else, will be demanding of administrators and boards of education that the teachers manifest the kinds of behaviors presented in the Medley report, for these will produce good test scores.

The situation would not be so filled with disaster if the conclusions of the report had been written in guarded terms; but the conclusions, as they are written, clearly imply that Medley believes that here are findings with very general immediate application both in education and in teacher education. The unguarded way in which the conclusions are stated in the report makes it a very dangerous document. The danger is doubled by the presentation of an array of Madison Avenue-type endorsements that appear at the end of the report. Indeed, the total effect of such a presentation appears more as a power play on the part of conservatives in education than as a serious scientific effort to achieve some understanding of teaching.

CONFLICTING THEORIES OF EDUCATION

The Medley report is an example of the long conflict that has existed between theory of education as presented by the research worker and theory of education as viewed by the analytic thinker in the field of pedagogy. In a previous generation this conflict was represented by the conflict between the followers of Edward L. Thorndike and his connectionist views of education and the followers of John Dewey and what was then termed progressivism. For Thorndike the important outcome of instruction was for the pupil to be able to solve the problems on a test. In mathematics, for example, if pupils could show that they could make the right connections between the problems and their solutions, then instruction had been a

success. On the other hand, the progressivist viewpoint of learning mathematics was that the important outcome was to be able to understand the logic of mathematics; computational skill was secondary. The difficulty of undertaking research within the progressivist conception of education was that it is hard to measure understanding of the logic of mathematics, so little research emerged. In contrast, studies within a Thorndikian framework proliferated, even though they were based on the common folklore conception of education.

Although research of the latter kind increased, it had very little to do with what experimentally minded educators and innovative teachers were attempting to do in the classroom. Such research has had a long history of proliferating, yet having no impact simply because it is out of tune with professional educational thought. Indeed, the main impact of such research has been to give support to the most reactionary community elements that would willingly move education back a full century if they could.

Educational research workers do not have to be tied to a folklore conception of education; they remain so tied largely because they have had no contact with the history of education. Many have also had no contact with modern innovative thought, such as that of Jean Piaget. Yet the fact is that from the first survey of educational achievement using written tests, undertaken in 1845, the educational research worker has typically reflected the most reactionary views of education.

There are, of course, other reasons for this besides a failure to understand history and the social context of education. One of these is that the folklore theory of education is extremely simple and easy to incorporate into a research design. It can also be elaborated to give the illusion of sophistication without too much mental effort. Competency-based theory of teaching is a simple and naive elaboration. A similar conception of medicine in which the physician has only to match symptoms with drugs is equally worthless. The great reformers of education such as Johann Pestalozzi, John Dewey, and more recently Jean Piaget have all pointed out the enormous damage that the perpetration of the folklore conception of education has probably produced in terms of socially alienated children and achievement-alienated learners.

Nevertheless, there are real difficulties in developing research programs related to innovative conceptions of education. Research that deals with innovative forms of classroom teaching is unlikely to flourish. Throughout the world, bureaucracies that sponsor educational research are typically tied to a folklore conception of teaching. These bureaucracies tend to embrace the popular ideas of the public of the moment, and to support research related to these ideas, which are nearly always reactionary. The National Institute of Education from its inception has fallen into this pattern; if it has become anything at all, it is an institution that supports prosaic research projects related to prosaic conceptions of education. Indeed, much of the research covered by the Medley report reflects the dull and unenterprising work sponsored both by the National Institute of Education and by its parent body, the U.S. Office of Education.

Educational research does not have to be undertaken within the current framework of folklore. Here and there throughout history have appeared small islands of research related to a more professional conceptualization of education. The famous Eight Year Study, being a child of the Progressive Education movement, was one such enterprise. Although the analysis of thinking skills undertaken by the staff of the study would now be regarded as completely out of date, in its day it was the best analysis ever

undertaken. Some of the results of that study and related studies in the 1930s are of considerable interest in relation to the Medley report. Such studies found that schools with programs classified as progressive did a rather poor job of instilling mechanical skills, such as computation and spelling, but quite a good job of developing skills at more complex levels.

One suspects that there is a tradeoff. Computation skills are likely to be traded for mathematical understanding. Enjoyment of reading may also be traded for skilled mechanics. Every objective cannot be achieved by everybody. Some objectives may even be incompatible with one another. Fun with books may not be compatible with some of the drill-oriented procedures that are now returning to schools through the combined efforts of community reactionaries and educational research workers.

SUPPORT FOR REACTIONARY TRENDS

Finally, objections must be raised to the implication which seems to be found in many places through the Medley report that the teacher educator and the teacher should look at the results of the research studies and modify their practices accordingly. If my premise is correct--namely, that the results of research have already been misinterpreted by research workers--then surely there is even greater danger that those less in contact with research will be even more likely to misinterpret the findings. In the hands of those teacher educators whose position is reactionary, the Medley report will provide a license to move teacher education back to the beginning of the century. Teachers, overawed by the prestige of the research worker, will view the report as legitimizing the practices of the past century, and as condemning attempts to raise the level of thought in schools to new levels.

If the conclusions had been drawn from the report with some of the reservations that have been stated here, the potential harm that the report can produce would not have been so great. As the report stands, it can do great harm. It will do much to perpetuate the myth that there is a single variable called teacher effectiveness when, in fact, the term represents a complex set of variables. The practice of teaching to help children pass a spelling test is obviously quite different from the way in which a teacher goes about helping children understand how to design a simple experiment. The Medley report confirms folklore about how to achieve simple goals, but says almost nothing about how to achieve a very great deal that has to be accomplished in the classroom, and by techniques entirely other than those discussed in the report. The report describes data that are statistically sophisticated in design but, nevertheless, naive in terms of the psychology and pedagogy they involve.

If I had been writing the conclusions of the report, there is one that would have jumped out of the pages for me--namely, that there is something very wrong with research. I don't really expect that what is so obviously wrong is going to be changed very much, for what is wrong lies at the very source of research support, that is to say the federal government. It is also obvious that the universities need to develop sophistication in other areas of educational research besides statistics. Research workers need to become sophisticated in their understanding of behavior theory, and no educational research workers are worth their salt unless they have knowledge of the history of education. The studies reviewed in the Medley report indicate the present lack of training in both of these areas.

JUDY LANIER

Associate Dean for Program Development
Co-Director, Institute for Research on Teaching
Michigan State University

Donald M. Medley must be complimented for assuming the arduous task of evaluating nearly 300 studies of teaching. His commitment to facilitating access to and understanding of this major set of results is both admirable and appreciated; too often researchers lack concern for dissemination of findings to relevant and interested practitioner groups.

SIGNIFICANCE OF RESEARCH REVIEWS

My concern in making these introductory comments, however, is that they might be interpreted as ritualistic courtesy and, in that sense, substantively meaningless. To the contrary; for the task of reviewing research is becoming an increasingly significant endeavor. As Gene Glass (1977, p. 353) noted, "The integration of research studies requires the best minds. It should be valued more highly than many forms of original research." Such is the case for the scholarly document Medley has produced. It is especially valuable, I think, since it provokes critical thought on issues important to research on teaching and the relationship of these findings to teacher education.

The charge given to those of us reviewing Medley's work included discussion of his methodology and procedures as well as the implications of his findings for teacher educators and researchers. Though asked to focus major attention on inferences and implications for teacher education, I feel it is imperative to emphasize both issues. But before discussing Medley's methods and procedures for summarizing the research on Teacher Competence and Teacher Effectiveness and its implications for teacher education, I find it necessary to come to terms with his "choice of terms." I find the explanation for choosing the words "teacher competence and effectiveness" both confusing and obfuscating. Note, for example, the following passage:

> We shall use the measure of effectiveness as an indicator of teacher competence, inferring that teachers who are effective are more competent on the average than teachers who are ineffective. The distinction between competent and effective implied in this statement is important and yet easy to forget. Competence has to do with how a teacher teaches and is measured in terms of the teacher's behavior; how effective a teacher is is measured in terms of pupil learning. In other words, an effective teacher is always competent, but a competent teacher may not always be effective. . . .
>
> We shall view the behavior of the teacher as an effect rather than a cause, assuming that the competent teacher behaves in a certain way because he or she is competent. A strong relationship between teacher effectiveness and a particular behavior will be interpreted as indicating that such a behavior characterizes competent teachers, and therefore may deserve to be called a competency (Medley, 1977, pp. 6-7).

Choosing value-laden terms such as "teacher competence and teacher effectiveness" to describe a set of results from research on teaching seems especially unfortunate for practitioners. It deemphasizes the particular variables studied, the unique nature of the inquiries undertaken, and the findings that were accumulated. It does this by lumping the critical and specific attributes of the research into abstract and general constructs (that is, teacher competence and teacher effectiveness) and, in so doing, leads the practitioner to forget or overlook the idiosyncratic set of variables that came to define the "inherently" good qualities of competence and effectiveness.

It would have been more appropriate, for example, to title the review Teacher Behavior and Pupil Gains on Standardized Achievement Tests. This is, after all, what the selected studies and reported findings are primarily about. Further, practitioners would have had a clearer and more immediate understanding of the particular research that is reviewed in this volume, as well as the values held by the reviewer.

Explicit descriptions of research in this area seem important because of the multiple meanings and beliefs about what constitutes teacher competence and teacher effectiveness. There is little reason for one view or one set of values to dominate summaries of empirical work on teaching. Since diverse judgments and beliefs on the subject do exist, I suggest that we recognize them openly and admit that one set of similar inquiries is in fact only one part (perhaps a small one at that) of the explanation for what might constitute teacher effectiveness. This recommendation is not meant to be nitpicking in any sense; I just think that practitioners are capable of understanding and should have the opportunity to know that there are pluralistic views among respectable researchers about what constitutes effective teaching and about how it should be systematically studied.

METHODOLOGY AND PROCEDURES

From both the descriptive rhetoric and the methods used in selecting studies for Teacher Competence and Teacher Effectiveness, it is clear that Medley was guided by the prevailing value of a particular form of methodological rigor.

The methods he used for identifying the research to be reviewed had the severe effect of discarding 275 of the 289 studies potentially capable of shedding light on the behavior of effective or ineffective teachers. This left only 14 studies from which Medley would describe "what could be trusted." The application of such stringent criteria must be questioned, particularly in light of the findings of Glass, a noted methodologist and authority on procedures for summarizing research findings.

While visiting with the staff at the Institute for Research on Teaching, Glass reported that it is common in reviews of research literature for important studies to be overlooked and less significant work to acquire undue recognition when selection criteria hold methodological purity as the sine qua non of worthy research. Glass has examined numerous studies which vary greatly in their quality and power dimensions of research design, sample size, and methods of gathering, accumulating, and aggregating data. Interestingly enough, results from the "lesser" studies have not conflicted significantly with the findings of the methodologically "superior" ones. Furthermore, some of the methodologically weak studies examine important

factors that are overlooked in the more controlled studies that would survive "rigorous" selection processes.

Glass suggested that studies should not be thrown out on methodological criteria alone, since methodology does not appear to be all-important. Additional factors, such as judgments about the practical significance of the questions under study or the authenticity of the research context, should be concomitantly considered as worthy criteria for study selection.

Thus, a major concern I have with Medley's work centers around his criteria and method of selecting the "significant 14" that came to be reviewed. For the most part, the criteria Medley employed to judge "worthy" studies of teaching emanate from the traditional research values of the natural sciences: replicability of research design, statistically significant findings, standardized outcome measures, and low inference process measures. Questions and debates about the appropriateness of applying the natural science paradigm to the study of human activities, such as teaching, continue to gain increased attention. The criticisms of the social and behavioral scientists' dependence on the natural science model should, I believe, be seriously considered and taken into account when doing and reviewing research. The criteria selected by Medley indicate that this growing trend is ignored, however--with the consequence that all studies based on alternative models were disregarded or discarded. Varenhorst (1978) summarized the problem of undue emphasis on the natural science model:

> Those who are critical point out the demanding exactness it requires, which does not allow for accommodating intangible, uncontrolled variables of human existence. They point to the changing, rather than stable object of our studies, the human being. Human phenomena are not comparable to natural phenomena. Therefore, the natural science model is inadequate to illuminate much pertinent information we need (p. 4).

Bronfenbrenner (1977) stated the point even more emphatically:

> Human environments are so complex in their basic organization that they are not likely to be captured, let alone comprehended through simplistic, unidimensional research models that make no provision for ecological structure and variation (p. 516).

Thus, by rigorously adhering to the natural science paradigm, Medley ignored research on teachers and teaching that has important implications for understanding and improving teacher effectiveness. I cannot agree, for example, that nonstatistical studies fall into the categories of "no original research," "reviews of research," or "theoretical, philosophical, or opinionated discussions (from the armchair) of what a good teacher ought to do" (Medley, 1977, p. 5). Cusick's Inside High School (1973) is a contrasting case in point; as is Jackson's Life in Classrooms (1968), Lortie's Schoolteacher (1975), and Wolcott's Teacher Versus Technocrat (1977). It must be noted, however, that I am using the term "teacher effectiveness" in a broader sense than Medley used it. Rather than limiting the concept to teacher behaviors that correlate with standardized achievement gains, I am referring to teaching thoughts, actions, and conditions that can be shown to relate systematically to valued outcomes.

Thus, I find myself in a quandary. Clearly Medley couldn't review all research on teaching that relates to someone's conception of effectiveness, and it follows that he should not be criticized for limiting his domain and excluding some particular types of research. But this brings me back full swing to the original point. He should have described his narrow focus. He claimed to be examining ". . . studies which purported to shed light on the question, 'How does the behavior of effective teachers differ from that of ineffective teachers?'" (1977, p. 5). Determining what knowledge the research enterprise has produced on this question is a genuinely important and scholarly endeavor. Systematic biases should be acknowledged so that sins of omission are as readily visible as sins of commission.

In summary, I would say that the criteria Medley employed for selecting studies and findings that qualified as "worthy" of consideration warrant serious questioning. The first criterion having to do with a study's generalizability forced the exclusion of some potentially valuable research on teaching effectiveness, particularly descriptive field studies. The second criterion required statistically significant findings, which may not be as important as replication of findings across studies. The third criterion required long-term pupil gains in achievement and is open to all sorts of criticism for its narrowness (as occurred in the letters and comments in Appendix C of the Medley publication). The fourth and final criterion, a clearly specified and reproducible process measure of exhibited behavior, was eventually questioned by Medley himself, when he suggested that perhaps in future research teacher intentions should be considered as well (p. 69). In "Integrating Findings: The Meta-Analysis of Research," Glass aptly described my position regarding improved methods and procedures for reviewing research findings:

> Research criticism has taken an unhealthy turn. It has become confused with research design. The critic often reads a published study and second guesses the aspects of measurement and analysis that should have been anticipated by the researcher. If a study "fails" on a sufficient number of these criteria--or if it fails to meet conditions of which the critic is particularly fond--the study is discounted or eliminated completely from consideration. Research design has a logic of its own, but it is not a logic appropriate to research integration. The researcher does not want to perform a study deficient in some aspect of measurement or analysis, but it hardly follows that after a less-than-perfect study has been done, its findings should not be considered. A logic of research integration could lead to a description of design and analysis features and study of their covariance with research findings. [If, for example, the covariance is quite small between the size of an experimental effect and whether or not subjects were volunteers, then the force of the criticism that some experiments used volunteers is clearly diminished.] Obviously, studying the covariation between design characteristics and findings can lead to better designs (1977, p. 355).

INFERENCES AND IMPLICATIONS FOR TEACHER EDUCATION

I am puzzled about Medley's report in terms of its inferences and implications for teacher education. Part of the puzzlement might result from the ambiguous tenor communicated about the research findings

themselves. On the one hand, I hear Medley pointing to the collected findings from the process-product studies with pride and confidence in their work. He stated, for example,

> . . . after reading this report and studying the findings presented, the reader will agree that no serious student of teaching can afford to be ignorant of the findings produced by research in teacher effectiveness (p. 4).

On the other hand, I hear serious doubt expressed:

> . . . it is perhaps time to consider whether the process-product model has outlived its usefulness, if for no other reason than that it ignores two critical variables almost completely: the intent of the teacher, and the behavior of the individual pupil (p. 69).

As a teacher educator this leaves me doubtful about the credibility that should be placed in the findings that have been so rigorously distilled. Teaching is intentional and individual pupil behavior should be considered, as must the characteristics of students and the complexities of instructional context. As Medley himself pointed out, these factors are critical. Thus, as a teacher educator, I am perplexed as to how to interpret findings from research that has, for the most part, ignored these factors.

Judicious reservations are expressed in the letters and comments in Appendix C of Teacher Competence and Teacher Effectiveness. Concerns about application of findings from those whose own studies are included carry particularly urgent and eloquent requests for caution. I share their reticence to modify teacher education curricula based on these findings. Furthermore, most of the process-product studies summarized by Medley lack adequate contextual consideration to provide sufficient direction for teachers. Teachers must decide when, with whom, under what conditions, and toward what ends it is important that certain behaviors be enacted in teaching. They are humans, to be educated in exercising professional judgments under complex conditions of uncertainty; they are not machines to be programmed or animals to be trained. Unfortunately, results of significance are not yet available.

While I am eager for teacher education to have a more scientific basis, I am not eager for it to have one built on a weak foundation. Teaching is a very complex phenomenon and its study has just begun. We are breaking important ground, to be sure, but we are a long way from being ready to inform practice in any significant manner. I worry, however, about questions about "where all this investment in research is going, if not into teacher education."

My first response to this concern is to urge a better awareness of history. Research on teaching has been undertaken only "rather recently" (for about 50 years). In terms of the history of natural science, that is not even a baby step. Patience is clearly in order. Further, we are making significant progress in our research methodologies and questions.

The increased use of diverse research methodologies and the increased involvement of practitioners in identifying important areas for study are greatly helping to advance the field. The wisdom of practicing teachers and teacher educators is seldom referred to any longer as "lore which does not lend itself to further refinement and study." This "lore" is rich and

is now being seriously studied. In the section, "Future Directions for Process-Product Research" (pp. 66ff.), Medley recognized this need and urged movement toward improved collaboration.

Medley's review provides a benchmark for those of us interested in teacher education and research on teaching. It marks the promising beginning of systematic review of K-12 research findings for teacher education. It is to be hoped that such reviews will continue, though they are by no means sufficient. The translation of K-12 research results into practice can best be expedited by supplementing research on teaching with research and development in three other areas: (a) development work for systematically adapting the knowledge accumulated through K-12 research efforts; (b) research for producing widely applicable knowledge regarding efficacious means of educating professional personnel, particularly teachers; and (c) research and development work for producing tenable specification of goals for K-12 teaching and teacher education (Lanier and Floden, 1977).

Respect and support for these research and development endeavors appear to be growing. There is no doubt that such efforts will improve the profession and its purpose: improved learning for children, youth, and adults.

REFERENCES

Bronfenbrenner, U. "Towards an Experimental Ecology of Human Development." American Psychologist 32 (7): 513-31; 1977.

Cusick, Philip. Inside High School. New York: Holt, Rinehart, and Winston, 1973.

Glass, Gene. "Integrating Findings: The Meta-Analysis of Research." In: Review of Research in Education 5: 353, 355; 1977.

Jackson, Philip. Life in Classrooms. New York: Holt, Rinehart, and Winston, 1968.

Lanier, Judith E., and Robert E. Floden. "Research and Development Needs for the Advancement of Teacher Education." East Lansing: Michigan State University College of Education, Institute for Research on Teaching, October 1977.

Lortie, Dan. Schoolteacher. Chicago, Ill.: University of Chicago Press, 1975.

Medley, D. M. Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research. Washington, D.C.: American Association of Colleges for Teacher Education, 1977.

Varenhorst, Barbara. "In Search of the Illusive View of Man." Educational Researcher 7 (9): 4; October 1978.

Wolcott, Harry. Teacher Versus Technocrat. Eugene, Ore.: Center for Educational Policy and Management, 1977.

BARAK ROSENSHINE

Professor, Department of Education
University of Illinois

The Medley report is an excellent review which will be extremely useful for students of teaching, developers of future correlational and experimental studies, and developers of future reviews of research. It is a reference that everyone developing correlational or experimental studies should study.

The implications and inferences for teacher education programs, however, are limited by the fact that the review is based on only 14 studies and most of the conclusions are supported by only one, two, or three studies. I inspected the review to identify the number of conclusions that are supported by five studies and found only three:

--Effective teachers of low SES students in primary grades engage their students in more lesson-related activities than less effective teachers do.

--Effective teachers of low SES students ask more questions classifiable in the lower levels of Bloom's taxonomy.

--There is less deviant student behavior in classes taught by effective teachers.

I would endorse the use of these three findings in teacher education programs. Teacher educators may wish to use the remainder of the findings in their programs--they seem reasonable and sensible--but they should do so with the knowledge that these findings are based on only one, two, or three studies.

There is a general model of teacher directed, small group, small step, academically focused instruction which emerges across the results summarized by Medley and other reviewers, and I would suggest that this model--and the contributions which Medley's review makes to this model--be the basis for future experimental studies in teaching.

MARIANNE MURPHY

Chairperson, Department of Education
Mundelein College

"Very few decisions worth making can be put off until there is adequate information to base them on."[1]

Teacher training, as a formal discipline, has been practiced for too long a time with too little research to draw from and assist it. This may seem an odd statement when libraries are filled with research--but not so odd when we reflect that in ancient cultures, to sit at the feet of the master and observe was considered sufficient training. In the early, and not so early, days of America mere completion of an academic level qualified one to teach at that level in our common schools. Dedication and devotion to duty were allowed to compensate for a lack of formal training. We can no longer treat the professional training of educators as a mere imitative process.

Dr. Medley's study goes far in helping to analyze and synthesize some critical behaviors for teachers. By examining a number of studies and submitting them to his four criteria, he has offered a survey of behaviors found to be beneficial for a particular age and socioeconomic class of children. The document as presented is easily understood and should be of considerable assistance in planning standards for teacher education programs in the future.

One concern that is immediately apparent is the restriction of this study to long-term outcomes only. In the typical common school classroom, short-term outcomes assume a somewhat more important role. The child's immediate--even if small--successes create instantaneous goodwill and positive feelings toward school. It is these feelings that go a long way toward making tomorrow possible. At the secondary level, adolescents make decisions on remaining in school, on attending classes, or on truancy based on the day-to-day, short-term experiences they are having.

This study is also restricted in two other ways: in grade level, it is limited for the most part to studies in the primary grades, while in socioeconomic status its findings are concentrated on subjects in the lower level. The outcomes, however, surely warrant sufficient compilations and analysis of studies to broaden the scope of process-product relationships to other populations and hence to be of greater value in the training of teachers. These limitations present a problem in trying to use the results of this study too extensively in the typical teacher training program. Most institutions of higher learning attract students from diverse backgrounds who have widely varying ambitions about where and with whom they wish to teach. To rely too heavily on research that was never intended to apply to everybody, but only to a limited class, would result in stressing inappropriate behaviors for certain classes of students. As individual teachers use a document such as Teacher Competence and Teacher Effectiveness, they must always be conscious of the real limit of its scope.

1 Donald M. Medley. Teacher Competence and Teacher Effectiveness. Washington, D.C.: American Association of Colleges for Teacher Education, 1977. p. 3.

Societal changes make the analysis of certain behaviors, especially those that deal with pupil-teacher interaction, almost ephemeral. In such a transitory society the reactions to specific behaviors will be in a state of constant change. Those persons who have done work with multicultural education will attest to the fact that one of the most important areas they face in dealing with children from the non-predominant cultures is to inculcate into teachers a thorough understanding of the child's relation to adults in general and to adults in positions of authority within the ethnic background. Only then can a teacher react in a positive manner to the child and know what to expect in the way of reaction from that child. For example, in a culture which demands complete acceptance by the young of what is said by an elder, it is apt to be nearly impossible to get a verbal interaction with the child at all.

Pupil-teacher interaction cannot be viewed in isolation either. There are a variety of outside influences on the child, and to ignore them appears to relegate them to a negligible position. The home and the community play a far greater role in the child's life than the school does and have to be considered in assessing the teacher's behaviors. The teacher's influence on the student's behavior and perhaps ultimately on the student's achievement is clearly subject to the influence of both the home and the community. These spheres have had the child during the more formative period; and even after that child starts to school they continue to have him or her for a greater period of time.

We also teach many more things than content subjects that can be measured by standardized test at the end of the term. In the school we teach social behaviors and mores by our own behavior as well as by formal instruction. These other, nontestable, areas contribute to the student's overall feelings toward school and the content areas. Anyone who has taught in the elementary or secondary schools knows countless examples of students who exhibited many positive behaviors and had very desirable relationships with the instructors but who failed to master the discipline of reading or arithmetic at the appropriate times. If the teacher is to be evaluated as effective on the results of these students' tests, the most effective teacher--in terms of establishing long-lasting relationships and having more permanent effects on their pupils' lives--may well turn out to be labeled ineffective.

Finally, there is the issue of where this study fits into the ultimate objective of certification and retention of teachers in the profession. Competencies and specific behaviors are far more difficult to assess than achievement in traditional areas. Though most teachers can be certified without displaying anything other than the completion of an approved program, others must endure the rigor of qualifying examinations, such as the National Teacher Examination. Traditionally the ability to demonstrate, by examination, proficiency in the content to be taught and the pedagogy of teaching was deemed sufficient to be awarded a certificate. Studies such as Dr. Medley's and those incorporated into it reinforce the view that content alone is insufficient and that, to produce the most effective teacher, affective factors must be considered as well.

The danger at this point lies in allowing the pendulum to swing too far to the behaviors side and substantially lessening the emphasis on the theory and content. What is necessary is that we produce a congenial marriage between the two so that students will have teachers competent in both aspects. Graduates of teacher education programs will be justifiably angered if the theory and content they must master in order to pass written

certifying examinations are overlooked to devote too much time to analysis of specific behaviors. In this era of accountability, school districts might well find themselves the defendants in lawsuits brought by dissatisfied parents; and their best defense will be the rigorous training in all disciplines of the teachers they hire.

Those responsible for teacher education would do well to examine Dr. Medley's study thoroughly and to use it within the limits he so clearly delineated. Also, research agencies ought to be encouraged to fund more studies of this type, in order to widen the scope and applicability of this kind of behavioral analysis.

GEORGE E. DICKSON

Dean, College of Education
University of Toledo

Those affiliated with competency-based or performance-based teacher education programs and those who continue with the more traditional experience-based model for preparing teachers should all recognize a present reality about teacher competence. It is one thing to develop programs to prepare, supposedly, competent teachers, but it is quite another matter to determine if the products of such programs are indeed competent professionals and effective in the cause-and-effect relationships between teacher behavior and pupil learning. The problem, basically, is that of determining the elusive relationship between teacher training experiences and pupil outcomes.

Teacher educators have not been much concerned about this problem because we have always somehow assumed that an individual experiencing a program of teacher education would be able, ipso facto, to handle children in the classroom with some degree of academic success. There was never really much of anything developed in the way of evidence to support such a belief; we just accepted the seemingly obvious and let it go at that. But with the coming of the competency-based teacher education (CBTE) movement, there has been a considerable focus on teacher performance and competency, and what people used to take for granted about successful teacher behavior and resultant pupil outcomes is now being challenged daily. Naturally, most of the concern has been aimed at the CBTE movement which has brought the situation to center stage in teacher education development. Those of us involved in CBTE accept the challenge, although all types of teacher educators must be subject to such concern. As Medley has said, "Only when we know why a teacher is effective--as well as how--can we decide how best to train teachers" (1977, p. 6).

A major criticism of the CBTE movement has been the amount of time and money spent in the development and implementation of a type of teacher education designed to promote measurable capabilities of teachers with relatively little knowledge of what kinds of teacher behaviors really are associated with pupil growth and learning. The critics point out a need for considerable research and assessment in the CBTE movement before it goes much further down the road of program development. Obviously, CBTE proponents must be about the business of engaging in the research and evaluation efforts that will provide us needed information about teacher performance and its relationship to teacher effectiveness.

Medley's publication, Teacher Competence and Teacher Effectiveness, is most important in such a context. In relatively simple terms, Medley has provided competency-based educators with a useful, thoughtful compilation and integration of research on teacher effects. He has isolated the important studies and analyzed their findings. We are provided with the research results which appear consistent and which, heretofore, have been available only in isolated reports or summaries that have not promoted total consideration of such efforts. Medley's publication highlights the distinction between competent and effective, in that the measure of teacher effectiveness (long-term pupil gain) is the indicator of teacher competence. Whether educators like it or not, the effectiveness of teachers must be considered in terms of pupil learning. Any teacher behavior that

results in effective learning characterizes a "competent teacher," as this term is interpreted by the general public. All types of educators need to learn as much as possible about the relationships existing between the levels of assessment of teacher competence in teacher education (see the paradigm in Figure 1), but teacher educators have the greatest concern and are the most vulnerable to the results of process-product, cause-and-effect relationships as revealed in research on teacher effectiveness.

The request for this commentary on Medley's work suggested that reaction be provided concerning his methodology and procedures, the inferences/implications for teacher education programs from his findings, and what can be said about the current state of the art in such research as well as new approaches in investigating teacher effectiveness. Because Medley has done his homework and his publication adequately speaks for itself, limited comment seems appropriate for the first two suggestions. The final item requires more consideration.

METHODOLOGY AND PROCEDURES

There is no doubt that Teacher Competence and Teacher Effectiveness is a major contribution to and model for any review of research on the subject. It provides the teacher educator with easy access to the meaningful, major studies of research in teacher effectiveness. As organized, the study is straightforward, and the best compilation of process-product relationships to date. The four criteria used for the selection of the research studies reported are sensible and defensible. Criteria had to be identified, and Medley took positions which resulted in identifying the truly important studies in the field, eliminating those that could not measure up to his stringent standard.

It is interesting to note that the procedures applied resulted in including the four major studies on teacher effectiveness (Brophy and Evertson, 1974; McDonald and Elias, 1976; Soar, 1973; and Stallings and Kaskowitz, 1974) which most major researchers agree are the most prominent in the field. Gage, for example, had stated that these four studies ". . . may be regarded as providing a capstone of the correlational approach to meeting the need for an improved scientific basis for the art of teaching

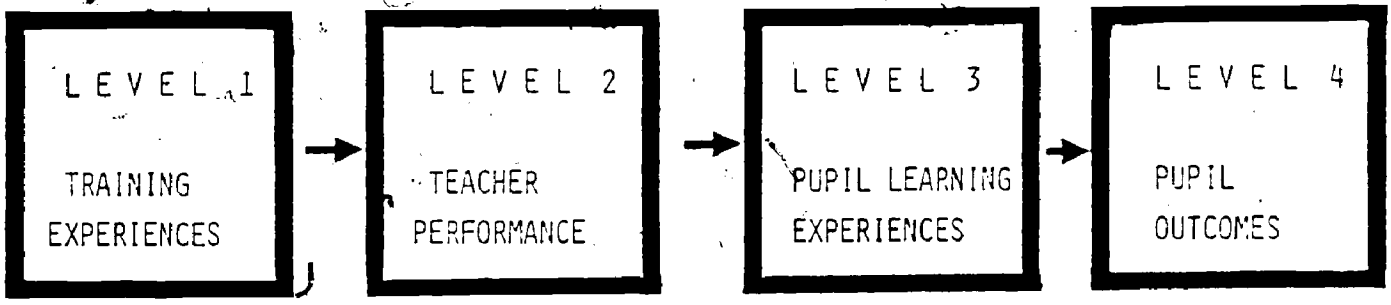

Figure 1
ASSESSMENT LEVELS IN TEACHER EDUCATION

From: Donald M. Medley, Robert S. Soar, and Ruth Soar. Assessment and Research in Teacher Education: Focus on PBTE. Washington, D.C.: American Association of Colleges for Teacher Education, June 1975. p. 2.

and the education of teachers" (1978, p. 1). One can quibble with the criteria and study procedure (and some research/statistical experts will), but their application has resulted in a quantity of important data for review.

Although subscribing to the use of all of the criteria, I take some exception to the elimination in Criterion III of studies involving short-term learning gains, in that the elimination may suggest that such studies should not be pursued or considered sufficiently in the future. Certainly, any studies involving student teachers and the teacher education programs which produce them, and yielding short-term gains, could be down-rated or even eliminated if this criterion were always stringently applied. As Brophy has stated:

> Short-term outcomes are convincing in their own right, whether or not they correlate with long-term outcomes. In fact, they provide the linkages to explain why teacher behavior influences long-term outcomes, especially test performance that is not related in any direct way to the teaching behavior of interest. Data on long-term outcomes are needed to show that the teacher behaviors have important effects, but short-term outcomes lead us toward explanation of how the processes work and provide evidence that correlational relationships reflect causal ones. Linkages between teaching behaviors and short-term outcomes are useful even in the absence of information about long-term outcomes, and linkages between teacher behaviors and long-term outcomes are incomplete (in Medley, 1977, p. 117).

INFERENCES/IMPLICATIONS

Much space could be spent in detailed comment on the inferences/implications for teacher education programs from the Medley findings. Any inferences made are necessarily limited by the fact that most of the studies reviewed dealt with the lower grades (1-3) and the teaching of reading or arithmetic. This limitation is understandable but points toward the need for more such research in other subject areas as well as at higher grade levels. Nevertheless, certain implications for teacher education programs are evident, with the cautious provision that the findings should not be overgeneralized.

An impressive finding is that apparently successful teacher behaviors for higher and lower socioeconomic status (SES) pupil groups differ so markedly. What turns out to be effective with one group proves ineffective with the other. Since many classrooms are heterogeneous in terms of pupil grouping, including SES, prospective teachers must learn more about effective grouping within the classroom and the use of appropriate teaching strategies for particular groups, in order to maximize pupil learning. The application of a general teaching methodology as provided in many teacher education courses does not prove useful.

The conduct of discussion or questioning by teachers of high and low SES groups is interesting. The effective teaching strategy is quite different for each group in terms of the way in which questions are asked, the level of question difficulty, the procedure for calling on pupils for answers, and the response to pupil answers. Evidence of this nature indicates some considerable changes in teacher education efforts.

Another intriguing SES inference is that the teacher who gives the most individual attention to pupils in high SES classes becomes the least effective in producing cognitive gains, and that the opposite is true in low SES classes. Instructional strategies in which pupils do individual assignments and work in small group activities will need attention. It appears quite obvious that pupils from different backgrounds demand different teaching strategies to maximize learning opportunities.

The evidence that effective teachers use more praise, reinforcement, and encouragement than do ineffective teachers has implications for teacher educators. Further, the fact that effective teachers tend to be more authoritarian says something about teacher planning, lesson organization, and classroom operation. The effective teacher is the better classroom manager, is less permissive, is more supportive, and maintains a learning environment freer from disruptive pupil behavior. A related instructive finding is that effective teachers of low SES pupils in the primary grades engage their children in more lesson-related activities than less effective teachers do. Time on task is an important criterion of teacher effectiveness at particular SES and age levels. Teacher educators can also be instructed from the evidence that teachers who produce maximum achievement gains also are likely to improve, considerably, pupils' self-concept and to develop in pupils positive attitudes toward school. The popular notion of classrooms exhibiting excessive pupil freedom and license apparently is not consistent with effective teaching as measured by long-term pupil subject matter gains.

Some of the research results from Medley's publication are not surprising, but others challenge previously held beliefs or assumptions about effective teaching. Teacher educators need to identify the substantive, consistent results and incorporate them into teacher education program development. An analysis is required of how programs can be defined to obtain the desired results.

NEW AND FUTURE APPROACHES

The knowledge base for teacher education has been extended by the Medley study of process-product research. Although teacher effectiveness research is still in its infancy, both methodologically and quantitatively, we apparently have some consistent indicators of how effective teachers behave. Those involved with teacher education activities not only need to learn how to use such research results but should conduct similar research and, if possible, improve upon what has been done. As was noted earlier, CBTE has focused increased attention on the competencies teacher education programs should be helping prospective teachers acquire. Interestingly, there seems to be evidence that the competencies promoted in CBTE programs are far from congruent with the competencies revealed by researchers in process-product research.

As Medley properly pointed out, teacher educators and researchers must get together on this matter to understand better the dynamics of effective teaching and then to promote the development, operation, and evaluation of competency-based teacher education programs. The researchers and teacher educators need to agree on the teaching competencies or acquisition of teacher behaviors which are likely to result in the preparation of effective teachers. Medley put this situation well (and all CBTE program operators should be in agreement) when he indicated the proper way to go is for

. . . future process-product research to use as process variables the same competencies that the teacher education programs are trying to help teachers acquire. Researchers and teacher educators should get together to investigate the validity of the latter's program goals. . . . The result would amount to a large-scale, continuous experiment in teacher education. Such an experiment--or better yet, a number of them linked to several teacher education programs--would have at least two important effects. First, it would directly improve the effectiveness of the program studied. And second, it would add as much to our understanding of the dynamics of effective teaching--or more--than any amount of the one-shot, process-product research that is the present norm (1977, p. 67).

Medley asserted that such a research shift and development ought to be an integral part of any CBTE program.

THE TOLEDO CBTE PROGRAM

The developers of the CBTE program operated at the University of Toledo are completely in agreement with Medley's position. We have been thinking this way for the past five years and have been attempting to develop for our CBTE program a research/evaluation effort which will result in its evaluation and validation. In 1974 and 1975, we developed a general model of effects in CBTE (Dickson, 1975, p. 106) (see Figure 2) which parallels the levels of assessment of teacher competence and teacher education as developed by Medley, Soar, and Soar (1975) (see also Medley, 1977, p. 68). Our principal goal was to establish the connective links as represented in

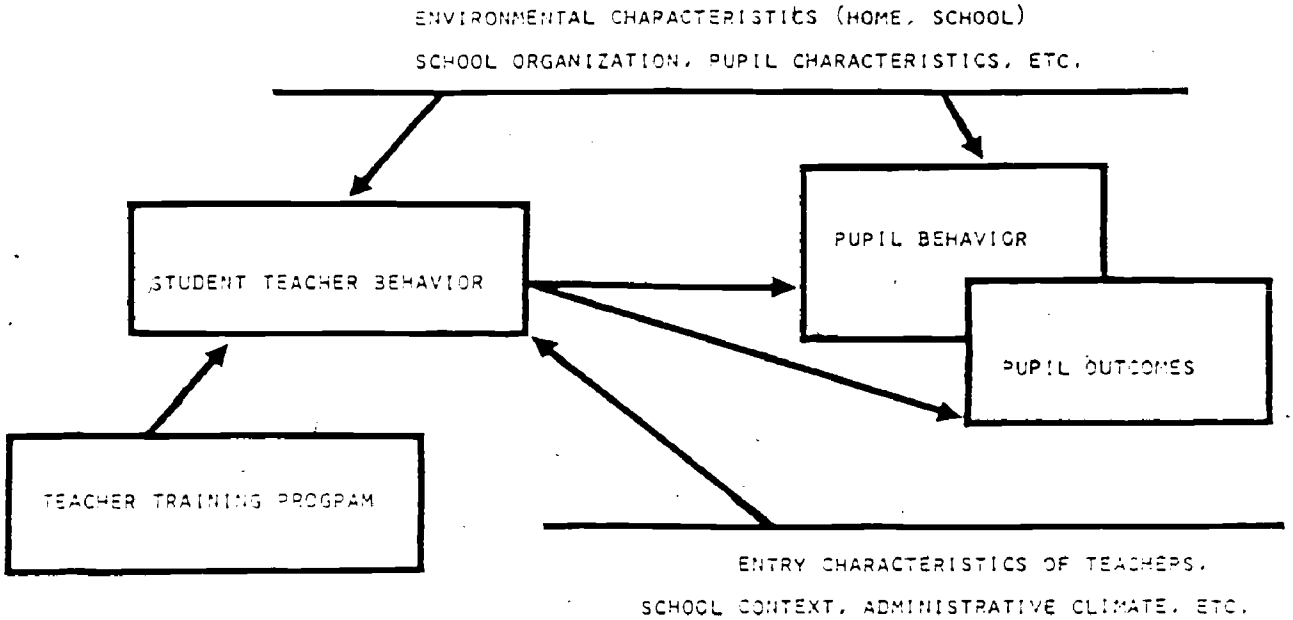

Figure 2
GENERAL MODEL OF EFFECTS IN CBTE
(Medley, Soar, and University of Toledo Research Staff)

our paradigm, and we planned to do that through a coordinated series of correlational and experimental studies which would consider specific teaching strategies through observations of classroom behavior, scores on teacher-made and standardized tests, and scores on affective inventories. The independent variables would be the teaching strategies, while the dependent variables would be pupil behavior and pupil outcomes in the basic content areas of reading and mathematics. We wanted to assess the links between the CBTE program and student teacher behavior, between student teacher behavior and pupil behavior, and between pupil behavior and pupil outcomes. Program evaluation would focus on the relationships between teacher training program efforts and student teacher behavior. Program validation would involve the relationships between student teacher behavior and pupil behavior and pupil outcomes.

Unfortunately, the resources to conduct such activity have not been made available to us. We cannot mount such effort with the normal funds provided in the regular university budget for teacher education instruction, and we have been unable to secure sufficient resources for such effort through federal government agencies or private foundations. Without sufficient research/evaluation funds, we have turned to the development of a statewide planning effort for teacher education research and evaluation in Ohio; this conceivably could mount the expertise and clout necessary for obtaining adequate research/evaluation funds for teacher education programs. This effort has progressed well and has resulted in the establishment of a Study Council for Research and Evaluation in Teacher Education in Ohio. The Council becomes operative during the 1978-79 academic year.

The entire planning effort has been funded mainly by the teacher education division of the State of Ohio Department of Education, in conjunction with statewide efforts to establish and promote a new set of Standards for Colleges or Universities Preparing Teachers in Ohio--commonly referred to as Project 419, the Redesign of Teacher Education. A full account is found in a planning report, Planning Teacher Education Research and Evaluation in Ohio (Dickson, Wiersma, and Gibney, 1977). Once the new Council becomes fully operable, we anticipate that funds will be found for the type of teacher education process-product research advocated by Medley. Ohio will then be a place where teacher education programs and their products can be examined to understand better the dynamics of effective teaching and teacher education.

The evaluation and validation of the University of Toledo CBTE program must take place. A beginning will be made in the 1978-79 academic year with $35,452 of special funding provided by the State of Ohio Department of Education for program evaluation and followup. With these funds we will design and conduct a modest process-product study to ascertain some of the short-term effects of student teaching behavior in relation to pupil behavior and outcomes in reading and mathematics. The effort will be limited to elementary student teachers at the second and fifth grade levels. The research/evaluation design will be developed in such a form that additional research effort can be incorporated at a later date when increased funding becomes available. Our teacher education faculty members are interested in such research, and some are thinking about research possibilities that go beyond current studies of teacher effectiveness. Their concerns involve the factors that impinge on teachers in their classroom planning and decision making, or what is meant by the expression "the teacher as a decision-maker" (see Dewitz and Hecht, 1978).

If real progress is to be made in future process-product research and if teacher educators are to be caught up completely in such work (as they should be), there must be close and constant collaboration between researchers and teacher educators which can only eventuate in significant improvement in the effectiveness of teacher education programs. Again, agreeing with Medley, what is needed is a focus on studies of the relationship between teacher behavior and pupil behavior (between competencies and learning experiences) and then continued study of the relationships between pupil behavior and pupil learning outcomes. These research efforts can certainly utilize the newly minted products of teacher education programs, especially CBTE programs. Studies can be organized to examine the short-term results of student teacher behavior and resultant pupil behavior. Long-term studies can be developed to follow up first-year teacher education program graduates as they practice in schools in their first teaching positions.

The best possible chance for meaningful research/evaluation lies in the real world of CBTE programs and their cooperating schools. It is one thing for CBTE critics to state that such programs have not been developed on a solid research base; it is quite another matter to pronounce such criticisms and then deny program developers the opportunity to create the research base needed. Teacher effectiveness research should be organized from both preservice and inservice standpoints. Such experimentation with linkage to teacher education programs is not yet evident.

To enable teacher educators to get on with the research/evaluation task before them will require the allocation of funds by federal agencies and other sources which have not been provided to them in the past. Medley's conclusions on future directions for process-product research show teacher educators the way to go. Whether they will be able to make the journey depends on the resources made available to them.

REFERENCES

Brophy, Jere E., and Carolyn M. Evertson. Process-Product Correlations in the Texas Teacher Effectiveness Study: Final Report. Austin: University of Texas at Austin, 1974. ED 091 394

Dewitz, Peter, and Lawrence W. Hecht, eds. Teacher Planning and Decision-Making, 1978 Educational Comment. Toledo, Ohio: University of Toledo College of Education, July 1978.

Dickson, George E., William Wiersma, Jr., and Thomas C. Gibney. Planning Teacher Education Research and Evaluation in Ohio. Toledo, Ohio: University of Toledo College of Education, Center for Educational Research and Services, Summer 1977.

Dickson, George E., ed. Research and Evaluation in Operational Competency-Based Teacher Education Programs, Educational Comment/I, 1975. Toledo, Ohio: University of Toledo College of Education, 1975.

Gage, Nathaniel L. "An Experiment on Teacher Effectiveness and Parent-Assisted Instruction in the Third Grade: Background and Rationale." In: An Experiment on Teacher Effectiveness and Parent-Assisted

Instruction in the Third Grade, Program on Teaching Effectiveness. Stanford, Calif.: Center for Educational Research at Stanford, May 1978.

McDonald, Frederick J., and Patricia Elias. The Effects of Teaching Performance on Pupil Learning. Beginning Teacher Evaluation Study: Phase II, 1973-74. Final Report: Vol. I. Princeton, N.J.: Educational Testing Service, 1976.

Medley, Donald M. Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research. Washington, D.C.: American Association of Colleges for Teacher Education, August 1977.

Medley, Donald M., Robert S. Soar, and Ruth Soar. Assessment and Research in Teacher Education: Focus on PBTE. Washington, D.C.: American Association of Colleges for Teacher Education, June 1975.

Soar, Robert S. Follow Through Classroom Process Measurement and Pupil Growth (1970-71). Final Report. Gainesville: Institute for Development of Human Resources, University of Florida College of Education, June 1973.

Stallings, Jane, and D. Kaskowitz. Follow Through Classroom Observation Evaluation, 1972-73. A Study of Implementation. Menlo Park, Calif.: Stanford Research Institute, 1974.

LEE MORRIS

Executive Secretary, Southwest Teacher Corps Network
University of Oklahoma

In his process-product study, Teacher Competence and Teacher Effectiveness, Medley summarized data of 14 investigations in order to examine relationships between teacher behavior and pupils' achievement in reading and mathematics. Most samples were taken from the primary grades, with approximately half of them using pupil attitude as one of the outcome measures.

This paper has three purposes. First, there will be a brief response to the Medley study. Next, the conclusions drawn by Medley will be discussed from the perspective of other research on teacher education. Finally, some current educational research activities will be considered, with implications for future research.

For years teacher educators and practitioners have debated the issue of which of two or more teaching methods are most effective in the teaching-learning process. In many respects, educational researchers have provided substance for the debate by investigating these methods, using a variety of research designs and statistical techniques.

The Medley summary of research findings represents one of the most important and interesting studies to date. By canvassing the literature, he succeeded in his effort of putting the reader in direct contact with 14 studies that stood the test of his criteria. The criteria used for selecting the studies were: (a) a relationship that is generalizable (to other teachers), (b) a strong and reliable relationship (between teacher competency and pupil learning), (c) a defensible measure of teacher effectiveness (pupil progress toward desired outcomes), and (d) an interpretable measure of teacher behavior (that is, to be able to reproduce the teacher behavior) (Medley, 1977, pp. 7-9).

Medley's study is laudable in the sense that his conclusions provide support for one theory (direct teaching), and it offers a design which may be useful for observing other variables associated with teacher behavior and pupil outcomes. He defines directed teaching as authoritative-type teacher behavior; stress on whole class or large group learning; maximal teacher talk and minimal pupil talk; simple factual questions; fewer pupil-initiated questions and comments; less amplification, discussion, or use of pupil answers; interaction at a low level of complexity and pupil initiative; infrequent use of pupil ideas and lack of inclusion of pupil ideas in discussion; and infrequent student-to-student interaction (pp. 12-21).

He made clear the limitations of his study and conclusions by acknowledging that most of the research was done in one segment of the school population (classes of Grade III or below where most pupils were from homes of low SES); he also described the procedures followed as efforts to reduce thousands of correlations to something both accurate and comprehensible.

The Medley study leaves a number of important questions unanswered. For example, why were so many studies rejected? Conversely, since only 14 of 289 studies reviewed were used, why should the education community believe the conclusions drawn from the few chosen? Admittedly, this may have been a function of the four criteria used in selecting the studies.

In light of Kennedy's (1978) analysis of Follow Through projects (one source from which Medley's data were drawn), one would have to question the credibility of Medley's criteria. Kennedy maintained that the Follow

Through study was not a true experiment, or even a well-designed quasi-experiment. She based this conclusion, in part, on the design complexity, political pressures that forced several modifications of the criteria, the process of selecting the control groups, and the inequality of the control groups. These and other factors led Kennedy to consider a variety of competing hypotheses which may also have implications for interpreting other measures of effects analyzed by Medley. She said an observed positive effect could be explained in any of four ways:

1. The treatment was superior;
2. The observed difference was simply a chance occurrence (a Type I error), and the treatment itself had no real effect;
3. The comparison group, selected non-randomly, was a poor match for the treatment group and the observed outcome difference is an artifact of initial group difference; or
4. There was a real treatment effect, but the effect was not due to the intended treatment. That is, perhaps the effect occurred because the model was not implemented (Kennedy, 1978, p. 6).

An alternative to rejecting so many studies and the possible information contained therein would have been to aggregate the findings using Glass' (1976) meta-analysis. While not condoning poor research designs, Glass contended that a study with several design and analysis flaws may still be valid. He maintained that the difference between well-designed and poorly-designed experiments is so small that to integrate research results by eliminating the "poorly done" studies is to discard a vast amount of important data. Use of meta-analysis would have allowed the size of effects to be observed in addition to the number of positive or negative effects. Additionally, variations in the size of treatment effects, as well as average treatment effects, could have been studied. Further, this analytic approach would have enabled the investigator to observe the consistency of effects.

RELATIONSHIPS BETWEEN TEACHER BEHAVIOR AND PUPIL READING/MATHEMATICS OUTCOMES

Whatever inferences are made as a result of the Medley findings should be limited strictly to a specific teacher behavior and pupil reading and mathematics achievement, as measured by standardized test scores. Also, inferences should be limited to the primary grades.

Whether or not one agrees with Medley's conclusions, they seem to be consistent and in agreement with many other findings. When the ambiguity and contradictions are filtered out, much of the research on teacher competence supports direct teaching methods for some children in lower elementary grades. This section will take the reader through a review of research on teacher competence and its relationship to reading and mathematics. No attempt will be made to cover the teacher competence/effectiveness literature comprehensively, because excellent reviews already exist (Rosenshine and Furst, 1971; Dunkin and Biddle, 1974; Brophy and Good, 1974; Brophy and Evertson, 1976).

Teacher behavior (competence or effectiveness) has been defined in the following research in terms of questioning techniques, grouping procedures, praise or criticism, use of class time, and the like, all of which are

believed to characterize either direct teaching or a discovery approach. In a sense, we are dealing with simple teaching techniques as contrasted with complex teaching techniques.

Brophy and Good (1974) argued that it is not appropriate as an unqualified general statement to claim that indirect or discovery teaching is more effective than direct teaching, and it is especially inappropriate for the early elementary grades. They maintained that one cannot, for example, teach first graders to read, write, and learn mathematics by conducting discussion. Advocates of direct teaching believe that at this level instruction must be heavily teacher-dominated, with frequent drill and repetition. The thought is that the educational objective is to get children to master fundamental skills to the point of their becoming automatic. Thus, Brophy and Good (1974) maintained that the "teacher talk is bad, student talk is good" dictum is inappropriate. Brophy and Evertson (1976) came to the same conclusions favoring direct teaching in early elementary grades. Basing their argument on Piaget's work, they suggested that children at this level are not cognitively ready for highly abstract material.

Ragosta, Soar, and Stebbins (1971) found that highly focused and concrete tasks were related to pupil growth in simple and concrete skills. They found also that emphasis on these "low level" tasks and skills maximized pupil growth on more abstract and complex skills. Teachers who were more successful at teaching simple skills were also more successful at teaching complex skills.

Hunt and Joyce (1967) found that student teachers who taught at a high conceptual level were more flexible, more capable of invoking alternate solutions and, in general, helped children think for themselves, in contrast with student teachers whose conceptual level, as measured by a special sentence completion instrument, was appreciably lower. Could pupil outcome measures, both affective and cognitive, then be more a function of the teacher's conceptual level rather than of the student's grade, race, SES, and other variables associated with teacher effectiveness and pupil growth?

In a five-year longitudinal study involving 154 children, Spaulding (1971) found that: (a) when the intent of the class was to pursue teacher-directed activities, the experimental children (as compared with control children) became more conforming and cooperative in their behavior; and (b) when the intent of the class was to have children operate productively on their own, there was an increase in independence, assertiveness, and productivity.

Flanders (1970) has conducted a number of studies supporting the theory that indirect teaching is positively related to student learning gains. He also has produced much evidence that indirect teaching is strongly related to positive student attitudes.

In a recent study (Rich and Bush, 1978) fourth, fifth, and sixth grade students were paired with students who were high and low in socioeconomic development (SED). Using direct and indirect teaching styles for 20 consecutive days to teach a series of reading lessons, the researchers found that indirect teaching related more closely to effective teaching, with the effect strongest for student affect, followed by achievement, then by time in attention to task.

The findings of Rich and Bush suggested that direct teachers who employed more structured learning environments appeared more facilitative for students low in SED than for students high in SED. Indirect teachers

who employed less structured environments appeared more facilitative for students high in SED than for students low in SED. These investigators concluded that ". . . an educational research commitment to search for the effective teacher regardless of context and type of student outcome appears to be an exercise in futility" (p. 456).

The findings and conclusions go on, seemingly contradictory but more toward a positive relationship between teacher effectiveness and direct teaching practices for primary grade children of low SES. Studies of relationships between teacher behavior and pupil mathematics achievement, while contradictory, seem not to point as strongly toward teacher directedness as does reading achievement.

For example, consider some of the findings on discovery versus directedness and inductive versus deductive modes of teaching mathematics. In a study with fifth and sixth grade children, Worthen (1968) concluded that discovery sequencing enhanced pupils' ability to retain and transfer mathematical concepts. Alternatively, he suggested that precise sequencing is better for immediate recall. Scandura and Wells (1964) reported that: (a) discovery enabled pupils to handle problem tasks better; (b) it took pupils longer to reach the desired skill using discovery; and (c) pupils of the precise method generally used the calculation taught, while discovery pupils varied in the calculation used. Armstrong (1969) suggested that the inductive mode aided in the learning of mathematical operations while the deductive mode resulted in greater learning of mathematical properties.

The results of many years of studying "meaningful" versus "mechanical" teaching have caused researchers to conclude that: (a) the meaning and rote or mechanical methods produce about the same results when immediate computational skill is used as a criterion; (b) when retention is used as a criterion, the meaning method is superior to the rote method; (c) greater transfer is realized by use of the meaning method; and (d) the meaning method produces greater understanding of mathematical principles and comprehension of complex analysis (Dawson and Ruddell, 1955; Miller, 1957; Rappaport, 1963; Kirch, 1964; Greathouse, 1966).

The reader is asked to treat these and other process-product studies with caution; for as Rosenshine and Furst (1971) warned, they are correlational, not experimental, studies. According to Rosenshine and Furst, the results of such studies can be deceptive in that they suggest causation, although the teacher behaviors which are related to student achievement may only be minor indicators of a complex set of behaviors that are yet to be identified.

RELATIONSHIPS BETWEEN TEACHER BEHAVIOR AND PUPIL AFFECTIVE OUTCOMES

There is an accelerating trend toward linking teacher behavior with the affective growth of pupils. This is occurring not only because of diverse racial composition of classrooms resulting from desegregation, but also because some researchers believe there is a high relationship between pupil attitudes and their cognitive growth.

In a study of 150 students in desegretated schools, Katz (1973) found that race, sex, and SES were all significantly related to the frequency of verbal initiation in the classroom. She concluded that desegregated schools simply reinforce rather than mitigate the racial differences

existing in the classrooms, as measured by a verbal initiation instrument. Whites initiated interactions much more frequently than blacks, and teachers either passively accepted or actively reinforced this trend rather than attempting to compensate for it.

Similarly, Yee (1968) found that teachers (both black and white) were most favorable toward middle class Anglo students, next favorable toward lower class Anglo students, next favorable toward lower class Mexican-American students, and least favorable toward lower class black students. Even among students of equal SES, according to Yee, whites were favored over Mexican-Americans, who were in turn favored over blacks, thus indicating the importance of considering race as a variable in studying relationships between teacher behavior and pupil affective and cognitive outcomes.

Williams, Whitehead, and Miller (1972) presented several studies showing that teachers tend to associate nonstandard English with negative attitudes and to develop low expectations for achievement in students who speak in this manner. They noted that teachers are likely to show negative attitudes, expectations, or behavior toward black children who speak in the form commonly referred to as "black dialect."

St. John (1971) found that warm and student-oriented teachers were more successful in producing achievement among low-income black pupils. Similarly, Brophy and Good (1974) concluded that teachers who are warm and student-oriented tend to be most successful when working with children who lack confidence, are anxious about school, and/or are members of rejected minority groups. The need for more research on relationships between teacher behavior and student attitude outcomes is clear.

It appears that the self-fulfilling prophecy hypothesis introduced by Merton (1948) and tested by Rosenthal and Jacobson (1968) continues to cause teachers to behave in ways that make their expectations more likely to come true. In fact, Brophy and Good (1974) did an extensive review of teacher expectation literature and established unequivocally that teachers' expectations of individual students can and do function as self-fulfilling prophecies.

CURRENT TEACHER EDUCATION ACTIVITIES AND IMPLICATIONS FOR FUTURE RESEARCH

Educational researchers have not provided those who train teachers with a variety of teaching skills which indicate that if behavior X is increased, and/or there is a decrease in behavior Y, then there will be a concomitant change in the cognitive and affective growth of pupils (Rosenshine and Furst, 1971). Aside from some of the competency-based teacher education (CBTE) programs, there also appears to be little done by teacher training institutions.

Heiss (1970) surveyed higher education to find out what is being done to improve college teaching. She concluded that practically no research of any kind is being done in this field. Peck and Tucker (1973) found the same results--nothing.

The picture may not be as bleak as it seems. The U.S. Office of Education sponsors a number of action programs under the auspices of the National Institute of Education (NIE), Follow Through, Teacher Corps, and other such programs that collaborate with institutions of higher education and local school districts in training and retraining teachers. These

programs could be a valuable data source for gathering information on relationships between teacher behavior and student outcomes.

For example, Teacher Corps has built-in components which allow decision makers to determine which programs work best with children from low socio-economic backgrounds. The demonstration thrust of some of these projects focuses on "validated research findings" for training and retraining teachers. Some of the training/retraining strategies are based on mastery teaching (Indiana University--Purdue University Teacher Corps Project; Youngstown State University Teacher Corps Project), individually guided instruction (West Virginia College of Graduate Studies Teacher Corps Project), children teaching children (CUNY--Queens College Teacher Corps Project), alternative reading approaches (Michigan State University Teacher Corps Project), and Joyce's models of teaching (San Jose State University), to name only a few.

Given action-oriented programs of this type, it would appear that researchers, teacher trainers, and practitioners together could investigate whether Approach A is better than Approach B. This would involve both primary and secondary analyses of a number of complex variables, such as content and teacher aptitudes, to determine the specified behavior that is most appropriate for each distinct set of subject area content and student population.

The growing interest in secondary analysis (Campbell and Erlebacher, 1970; Elashoff and Snow, 1971; Mosteller and Moynihan, 1972; Glass, 1976; Medley, 1977) seems particularly appropriate for answering new questions with old data. By studying the relationship between teacher effectiveness and--for example--race, SES, and age of students, it may be possible to determine the magnitude of effect on individual children, thus achieving Riedesel and Burns' (1973) hope; that is, for content material X, taught by teacher Q to pupil Z, the best strategy is A.

REFERENCES

Armstrong, J. R. "Curricular and Instructional Factors Which Facilitate Mathematical Learning." Journal of Experimental Education 38: 5-15; 1969.

Brophy, J. E., and C. M. Evertson. Learning from Teaching: A Developmental Perspective. Boston: Allyn and Bacon, 1976.

Brophy, J. E., and T. L. Good. Teacher-Student Relationships: Causes and Consequences. New York: Holt, Rinehart and Winston, 1974.

Campbell, D. T., and A. E. Erlebacher. "How Regression Artifacts in Quasi-Experimental Evaluations Can Mistakenly Make Compensatory Education Look Harmful." In: J. Hellmulth, ed. The Disadvantaged Child. New York: Brunner-Mazel, 1970.

Dawson, D. T., and A. K. Ruddell. "An Experimental Approach to the Division Ideas." The Arithmetic Teacher 2: 6-9; 1955.

Dunkin, M. J., and B. J. Biddle. The Study of Teaching. New York: Holt, Rinehart and Winston, 1974.

Elashoff, J. D., and R. E. Snow, eds. Pygmalion Reconsidered. Worthington, Ohio: Charles A. Jones, 1971.

Flanders, N. Analyzing Teacher Behavior. Reading, Mass.: Addison-Wesley, 1970.

Glass, G. V. "Primary, Secondary, and Meta-Analysis of Research." Educational Researcher 5: 3-8; 1976.

Greathouse, J. J. "An Experimental Investigation of Relative Effectiveness Among Three Different Arithmetic Teaching Methods." Dissertation Abstracts 26: 5913; 1966.

Heiss, A. M. The Preparation of College and University Teachers. Berkeley: University of California, Center for Research and Development in Higher Education, 1970.

Hunt, D. E., and B. R. Joyce. "Teacher Trainee Personality and Initial Teaching Style." American Educational Research Journal 4: 253-59; 1967.

Katz, M. "Attitudinal Modernity, Classroom Power and Status Characteristics: An Investigation." Paper presented at the Annual Meeting of the American Educational Research Association, 1973.

Kennedy, M. M. "Findings from the Follow-Through Planned Variation Study." Educational Researcher 7: 3-11; 1978.

Kirch, P. "Meaningful vs. Mechanical Method, Teaching Division of Fractions by Fractions." School Science and Mathematics 64: 697-708; 1964.

Medley, D. M. Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research. Washington, D.C.: American Association of Colleges for Teacher Education, 1977.

Merton, R. "The Self-Fulfilling Prophecy." Antioch Review 8: 193-210; 1948.

Miller, G. H. "How Effective Is the Meaning Method?" The Arithmetic Teacher 4: 45-49; 1957.

Mosteller, F. M., and D. P. Moynihan, eds. On Equality of Educational Opportunity. New York: Vintage Books, 1972.

Peck, R. F., and J. A. Tucker. "Research on Teacher Education." In: R. M. W. Travers, ed. Second Handbook of Research on Teaching. Rand McNally College Publishing Company, 1973.

Ragosta, M., R. Soar, and L. Stebbins. "Sign vs. Category: Two Instruments for Observing Levels of Thinking." Paper presented at the Annual Meeting of the American Educational Research Association, 1971.

Rappaport, D. "The Meanings Approach in Teaching Arithmetic." Chicago Schools Journal 44: 172-74; 1963.

Rich, H. L., and A. J. Bush. "The Effect of Congruent Teacher-Student Characteristics on Instructional Outcomes." American Educational Research Journal 15: 451-57; 1978.

Riedesel, C. A., and P. C. Burns. "Research on Teaching Elementary-School Mathematics." In: R. M. W. Travers, ed. Second Handbook of Research on Teaching. Rand McNally College Publishing Company, 1973. pp. 1153-54.

Rosenshine, B., and N. Furst. Research in Teacher Performance Criteria. In: B. O. Smith, ed. Research in Teacher Education. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1971.

Rosenthal, R., and L. Jacobson. Pygmalion in the Classroom: Teacher Expectation and Pupils' Intellectual Development. New York: Holt, Rinehart and Winston, 1968.

Scandura, J. M., and Wells, J. N. "Advance Organizers in Learning Abstract Mathematics." American Educational Research Journal 4; 295-301; 1964.

Spaulding, R. L. Educational Intervention in Early Childhood, Vols. I, II, and III. Durham, N.C.: Duke University, 1971.

St. John, N. "Thirty-Six Teachers: Their Characteristics and Outcomes for Black and White Pupils." American Educational Research Journal 8: 635-48; 1971.

Williams, F., J. Whitehead, and L. Miller. "Relationships Between Language Attitudes and Teacher Expectancy." American Educational Research Journal 9: 263-77; 1972.

Worthen, B. R. Discovery and Expository Task Presentation in Elementary Mathematics. Journal of Educational Psychology Monograph Supplement, Vol. 59, Part 2; 1968.

Yee, A. "Interpersonal Attitudes of Teachers and Advantaged and Disadvantaged Pupils." Journal of Human Resources 3: 327-45; 1968.

USING INTERIM RESEARCH RESULTS TO IMPROVE TEACHER EDUCATION

Donald M. Medley

Chairman and Professor
Department of Research Methodology
University of Virginia

When I prepared my monograph, Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research, for the American Association of Colleges for Teacher Education, I had one purpose in mind: to help close the gap between that research and the practice of teacher education. And so I suppressed any urge I felt to discuss how the findings might be used by teacher educators. I was delighted to learn of the plans for assembling a collection representing various points of view on the question, and am pleased at this opportunity to add some thoughts of my own.

A recent editorial in Science on the revolutionary impact that quantitative research has had on the practice of medicine since the beginning of this century seems to me to have strong implications for teacher education. The revolution began when the medical profession was forced to accept with great reluctance two conclusions, the research evidence for which was becoming impossible to ignore.

One conclusion was that almost none of the methods of treatment on which medical practice was then based was efficacious. Nineteenth-century physicians had at their command a tremendous armamentarium of therapeutic procedures which were well established but which research clearly showed were almost entirely useless.

The second conclusion was that a large proportion of patients recovered regardless of what treatment they received, or indeed whether they received any. Apparently, the profession was surviving by taking credit for those patients who recovered without accepting blame for those who failed to recover.

One cannot help wondering whether the practice of teaching may not now be just about where the practice of medicine was a century or so ago. The generally low correlations reported between teacher behaviors and pupil learning are consistent with the idea that many pupils may learn regardless of how they are taught. And we have no more evidence to support our repertory of teaching methods than the nineteenth-century physicians had to support their methods of treatment.

Once these conclusions were accepted, the best medical schools stopped teaching methods of treatment and concentrated on diagnosis and prognosis. In the meantime, medical research continued until, with the discovery of miracle drugs such as penicillin and cortisone, a scientific basis for treatment was laid, and medical practice became the powerful influence it is today. As far as I can tell, the confidence of the public in the competence of the medical profession seems never to have wavered during those decades in which the profession really had nothing to offer that would justify such confidence.

It is pretty clear that the public has no such confidence in the teaching profession. Unless research in teaching can build a knowledge base on which we can operate before the public discovers how little we know, stormy seas lie ahead of us. The experience of the medical profession and our own indicates that informal experimentation and clinical experience cannot turn

the trick. As my friend and one-time collaborator Harold Mitzel is fond of saying, the only way we can sort out the things we know that are so from those that ain't is by careful, sound research.

In his presidential address to Division 15 of the American Psychological Association, David Krathwohl pointed out that both the lay public and its elected representatives have been conditioned to expect research findings to be rather spectacular, to expect periodic major breakthroughs to be reported in the public press, and to view as unproductive--and, therefore, unworthy of support--research that fails to yield exciting new findings. The fact that the theory of relativity, the structure of DNA, the lunar landing, and similar headliners were all end products of many years of dull and costly research which gradually built up the empirical bases of the sciences concerned is a fact that tends to be ignored.

Educational practitioners often behave like the lay public in that they prefer to wait for a major theoretical development and to disregard interim findings--perhaps because the interim findings are "dull."

VALIDATING CURRENT CONCEPTS OF EFFECTIVE TEACHING

I am afraid the 600 relationships that turned up in my search must be regarded as contributions to an empirical base for the study of teaching rather than as the beginnings of a dramatic new breakthrough. And rather early contributions at that, in the sense that we need many, many more before it will be useful to attempt to synthesize them and evolve a theory of teaching. Not until that is possible can we expect any results exciting and important enough to produce even a modest breakthrough.

Most teacher educators feel a need for some general theory or model of teaching to serve as a basis for teacher education, but research in teacher effectiveness is as yet in too early a stage of development to provide such a theory. As a result, they tend to turn elsewhere--perhaps to a philosopher like Dewey or a psychologist like Rogers or Skinner--for a model of teaching to use. But anyone who does this needs to check the prescriptions of his or her model against the findings of the research in teacher effectiveness as these become available--that is, to validate the model.

This is the first way I see that the findings reported in the monograph might be used by teacher educators--to validate present conceptions of the nature of teacher effectiveness on which their programs are based. It is also the simplest.

Before discussing a second way, I should like to remind the reader of a point that the authors of some of the other papers in the collection seem either to have missed or to have forgotten. It is a very important point about the nature of process-product research.

It is not productive to conceive the purpose of such research as finding one-to-one correlations between teacher behaviors and pupil learning outcomes. To do so involves the implicit assumption that the learning is an effect of the behavior, and that if the novice teacher behaves in a certain way then the pupils will learn. If one conceives of process-product research in this manner, two consequences are likely to follow.

One is a tendency to move toward the use of short-term outcome measures rather than long-term ones. It is easier to detect correlations between variables measured close together in time; if you want to detect an effect, measure it promptly.

The other likely consequence is that you will conceive teacher preparation as primarily a matter of training teachers to behave in specific ways, ways shown by research to produce pupil learning. The trained teacher is then seen as one who knows what to do, and how often, in order to produce any desired effect on pupils. The implication is that (in theory at least) every situation a teacher will encounter can be anticipated and the optimum behavior prescribed beforehand; and that it is possible for a teacher to learn and remember all of this.

Such a model seems to me far too simplistic to be useful. It seems to equate the teacher with a skilled mechanic and the pupil with the machine being maintained. The job of the teacher is much more like that of any other professional in that a teacher must be prepared to deal with many problems which are novel and for which there is no prescribed solution. Teacher education involves at least two distinct phases. Teachers must learn to solve problems--to diagnose and prescribe, as it were. And they must also acquire a repertory of treatment skills, of behavior patterns they can call upon as needed in implementing their prescriptions.

IDENTIFYING BEST TEACHING PRACTICE

The number of skills that are potentially useful to a teacher is probably very large. If it is important for a teacher to know how to ask "higher-order questions," it is probably useful to know how to ask other kinds of questions, too. The problem for the teacher educator is to find out which skills are the most important ones to help students acquire. And the problem is compounded by the fact that this varies from teacher to teacher; that what works for one may not work for another. A second use for the findings in my monograph is to solve this dilemma.

I have already suggested in the monograph that what process-product research can tell us is which behavior patterns or practices are used more (or less) frequently by effective teachers than by ineffective ones. It seems reasonable to expect that these "practices" are the ones likely to prove most useful to the beginning teacher, as well as to the teacher in service who wishes to improve. If certification of teachers is to be based in part on demonstrated performance skills, it also seems reasonable to require that each candidate demonstrate mastery of a certain number of these "best" practices before being permitted to teach.

If the identification of best practice in this sense is conceived to be the purpose of process-product research, if as I have suggested these behavior patterns are regarded as the consequences rather than the causes of teacher effectiveness, then the use of short-term pupil gains as criteria becomes inappropriate. To assume that the teacher who produces the greatest gains on such a test is the most effective teacher is not justified. This is not to say that studies using such criteria are not useful for some purposes; it only means that such criteria are inappropriate for use in identifying effective teachers.

The study by Brophy and Evertson reviewed in the monograph was exemplary in that, in order to be identified as an effective teacher in this study, a teacher had to produce long-term gains well above average in three successive school years; and in order to be identified as an ineffective teacher, a teacher had to produce long-term gains well below average in three successive years. Differences in behaviors of teachers in these two groups in the classes they taught in the fourth year would certainly

reflect differences in practices of effective and ineffective teachers, even though no attempts were made to measure the effects of those behaviors on the students in those fourth-year classes.

Let me finally urge that if the findings of these studies are used in this second way--as a basis for deciding which performance skills/behavior patterns undergraduates should acquire in preservice training--they be presented not as sure-fire or even as the best available techniques. They should be presented as what they are; that is, as the practices the most effective teachers seem to prefer, and therefore the ones most likely to be useful. In the final analysis, each teacher must discover individually what practices are best for him or her.

IMPROVING THE CLASSROOM LEARNING ENVIRONMENT

I have discussed two ways in which the process-product correlations reported in the monograph might be used by teacher educators. One is to validate models for teacher education already in use; the other is to identify performance skills most likely to be useful to beginning teachers. A third possible application involves a somewhat different concept of the nature of effective teaching.

Much current thinking about the nature of effective teaching grows out of learning theory, out of what we know (or think we know) about how pupils learn. Learning is, after all, something that pupils do, and it has a much closer connection with gains in achievement than teaching does. This line of reasoning usually leads to a concept of teaching which gives the teacher a much more active role in pupil learning than seems feasible or even desirable in the typical classroom. Extreme examples are provided by some schemes in which the teacher is the principal source of reinforcement in a fairly elaborate reinforcement schedule tailored to the individual pupil.

This does not seem feasible in a normal classroom with 25 or 30 pupils; at one extreme the teacher must run 25 or 30 schedules simultaneously, and at the other each pupil is actively learning only about two minutes per hour. It also seems undesirable because of the complete dependence of the pupil on the teacher.

Process-product research began with the assumption that teachers' behavior has in it a sizable component that is stable across most or all of the different kinds of activity that go on in the classroom. These stable patterns of behavior are symptomatic of (if not the source of) what has been called the socioemotional climate or the learning environment in the classroom; and this was originally hypothesized to be an important element in teacher effectiveness. Under such an assumption it was logical to observe classroom processes at random times, at times when a variety of different activities were going on, so that the unstable parts of the behavior would tend to cancel out and only what was stable would remain.

Looking back, this seems to be a strange way of attacking the teacher effectiveness problem. It is analogous to trying to find out what makes physicians effective by observing them at work at random times, and counting how often each physician looks at a patient's throat, prescribes penicillin, or (perhaps) asks a higher-order question--all without paying any heed to what was wrong with the patient or whether the patient was in for examination, treatment, shots, or whatever.

The reason this seems to be such a footless way of studying medical practice is that what physicians do or how often is much less important

than when or why they do it. About all one could learn from such a study is something about their "bedside" manner, about how they relate to patients. This, one feels, is not likely to be a very important factor in their effectiveness in curing patients' ills.

It is here that the analogy between doctor and teacher is weakest. It is somewhat more sensible to assume that the learning environment a teacher creates and maintains is related to pupil learning than to assume that a physician's "bedside" manner is an important factor in determining how well patients recover.

The 600 correlations reported in the monograph seem to me to suggest rather strongly that the concept of the learning environment may be a useful one, and that we might set as one goal of teacher education to help each student learn how to create and maintain an environment favorable to learning. If so, we might be quite prescriptive. If we follow my reading of the findings reported in the monograph, we would have something to say about maintaining an orderly classroom without punitive behavior, maintaining pupil involvement in learning tasks, maintaining a low cognitive level in class discussions, etc. This is a third, and more prescriptive, way of using interim research findings in teacher education.

SUMMARY

How teacher educators decide to relate these research findings--and others as well--to the practice of teacher education is much less important than whether they do so. No profession, least of all ours, can afford to operate in this day and age without the soundest possible research base for what it does. It has been proven again and again that practice based on tradition, on common sense, or on folklore is likely to do more harm than good to society. Up to now, very little of what we tell teachers to do in order to be effective has any empirical basis at all, and even if all we know empirically were implemented tomorrow, there would still be a vast gap between research and practice. Which makes it all the more important that we use what we know.

READER RESPONSE

The Educational Resources Information Center (ERIC) is a nationwide information system of the National Institute of Education, whose basic objective is to provide ideas and information on significant current documents in education, and to publicize the availability of such documents. Through a network of specialized clearinghouses, ERIC gathers, evaluates, abstracts, and indexes these materials, and processes them into a central computerized data system. The scope of the ERIC Clearinghouse on Teacher Education is the preparation and continuing development of education personnel, as well as selected aspects of health education, physical education, and recreation education.

We are convinced that the knowledge base on teacher competence and teacher effectiveness is in need of expansion and that practitioners possess considerable expertise to contribute. We encourage you, therefore, to submit to us any manuscript you have developed on this topic and to encourage your colleagues to do the same.

We need a reproducible copy (two copies, if available) of any materials and, if possible, a brief abstract. Documents submitted are selected on the basis of their relevance to the current needs of the field. Those accepted are abstracted and indexed in the monthly journal, Resources in Education (RIE), and are made available in microfiche at over 600 locations and reproduced in xerographic form through the ERIC Document Reproduction Service. Copyrighted materials will receive only an announcement in RIE if permission to reproduce is not given.

Documents announced in RIE typically are unpublished or of limited distribution, and include research reports, program descriptions, speeches, annotated bibliographies, and curriculum guides. Dissertations available elsewhere are not announced in RIE.

We believe there are benefits in submitting documents to ERIC. Your work will be widely publicized since over 5,300 organizations subscribe to Resources in Education. Publications that have limited distribution or are out of print can continuously be made available to readers through the microfiche collections and reproduction service. And you will be performing a professional service for your colleagues.

Please send relevant documents to: Information Analyst, ERIC Clearinghouse on Teacher Education, Suite 616, One Dupont Circle, N.W., Washington, D.C. 20036.