
Title	Towards the development of valid (additional) selection instruments for pre-service teacher education
Author(s)	Eng Soo Peck
Source	<i>Singapore Journal of Education</i> , 5(2), 13-19
Published by	Institute of Education (Singapore)

This document may be used for private study or research purpose only. This document or any part of it may not be duplicated and/or distributed without permission of the copyright owner.

The Singapore Copyright Act applies to the use of this document.

Towards the Development of Valid (Additional) Selection Instruments for Pre-Service Teacher Education

Eng Soo Peck
Institute of Education

Introduction

Selection: A Non-issue in a Teacher-shortage Situation

The search for valid selection instruments to act as a screen for potential teachers is not a new concern. In the past, the concern was not matched by resources diverted to their development or by an adequate fund of knowledge and skill necessary for developing such instruments. Also conventional wisdom has it that the traditional criteria of teacher selection based on levels of educational attainment would be adequate to ensure that candidates selected for teacher education would be able to measure up to the demands of the teaching task. In UK, Singapore, Malaysia and other ex-British colonies, the GCE "A" level Certificate or a degree from a recognised university provides the first screen in the selection process for entry into teacher education institutions. Other criteria may also be specified such as relevance of university courses for school teaching purposes and a pass in "O" level Mathematics in addition to the GCE "A" Certificate. Whilst academic criteria for admission are based on sound and valid educational and professional principles, they are never sacrosanct. Academic and professional standards notwithstanding, admission criteria fluctuate with the forces of the supply of and demand for teachers. Even today, admission criteria vary across societies according to the developmental level of that society, and within society across time.

The harsh fact is that forces of supply and demand often determine admission criteria. In a seller's market, colleges of education and the school system cannot afford to be selective in whom they choose to be the teachers of the young. In a buyer's market for teachers, as it is today in the USA, UK, New Zealand and Australia, the issue of whom to select as teachers assumes an importance and urgency it cannot assume in a teacher shortage situation. Hence the issue of teacher selection re-surfaces with renewed vigour and a meaningful concern reflecting the over-supply of teachers and genuine concern that the future citizens cannot be short-changed through sloppy selection procedures and shoddy teacher performance.

Selection Criteria in the Singapore Context

In the Singapore context, the question addressed is whether past selection procedures based on academic attainment, relatively rigorous vis-a-vis many other countries, coupled with the interview are adequate to ensure a high quality teaching force. Whilst there is no detailed and systematic documentation of the level of competence of the teaching force, there is prima facie evidence that both the academic and professional competence of some teachers are inadequate in terms of the demands of the teaching task today. Also it is common knowledge that an adequate academic preparation is no guarantee of adequate teacher performance. Indeed, teachers who are less bright academically but whose attitudes towards work are thoroughly professional are likely to be more effective teachers than those who impress selection boards with their academic brilliance. In brief, academic attainment alone as a selection criterion is a necessary but not a sufficient condition for effective teaching. Indeed, the recognition that successful teaching requires more than academic brilliance provides the impetus to the search for valid instrument(s) to screen out misfits from the teaching profession.

Whilst there is a consensus on the need to select teachers based on valid criteria the task of coming up with a valid procedure for such selection is a daunting one. There is a plethora of selection instruments on the market for screening purposes for entry into different occupational groups but they have never been validated. (Mosher 1971, Cannedy, 1969). Indeed the reliance on test scores is more an article of faith than a proven fact. Among others, the one criterion these instruments must satisfy is the predictive power for on-the-job performance and there is a paucity of research evidence that the screening instruments currently in use in the job market are valid.

Some Problems and Issues in Selection Instrument Validation

It is perhaps a relatively easy task to come up with instruments for screening prospective teachers into teacher colleges if the objective is

limited to the successful completion of the teacher education programme. Such instruments would be useful if success in programme completion is highly correlated with effective teacher performance after graduation. There is no such evidence that this is so. Therefore any instrument or instruments used to screen pre-service teachers must also have high predictive validity in terms of teacher effectiveness whilst on the job. Another issue which complicates research into the development of effective selection instruments is the time frame in which the teacher is expected to be effective after graduation. Can instruments developed at this point in time with the present level of the state of the arts predict teacher effectiveness within the first five years of service or beyond? The longer the time-frame expected, the lower is the predictive validity of any instrument currently available. The next issue in the development of selection instruments pertains to the concept and measurement of teacher effectiveness. Effectiveness is an elusive concept. First, conceptions of effectiveness shift with prevailing notions of what constitutes teacher effectiveness. People occupying different status positions (such as MOE inspectors, IE lecturers, principals, teachers, pupils and parents) have different conceptions of teacher effectiveness. The question remains as to whose conceptions of effectiveness should the instruments developed be validated against. Also if teacher effectiveness shifts with the lapse of time, would instruments validated at one point in time be valid instruments at another point in time? There are more complex questions which need resolution than there are answers for them. One response to the apparently hopeless situation is to abandon research in this area. The other is to venture forth and contribute towards the resolution of some of the vexing problems. A very apt paragraph from Medley (1977) should spur us on:

"Very few decisions worth making can be put off until there is adequate information to base them on. In medicine and poker, most actions must be taken, most decisions made, on insufficient data. Patients die, and money is lost, because action is taken when data are inadequate but more patients and more money would be lost if no action were taken at all . . ." and "If Columbus had waited until he had a complete and accurate map of the world before setting sail his little fleet would still be sitting in Genoa". (p. 3)

It is with the above spirit that we set sail in the little charted sea with contradictory pointers of

research into valid selection instruments for pre-service teachers.

Highlights on the Teacher Effectiveness Literature

If selection instruments to be developed are to be validated against effective teaching, then the first task is to define and identify effective teaching (patterns of teaching behaviour considered effective) and/or the effective teacher (the person with such behaviour patterns). There is a voluminous body of literature of varying qualities on teacher effectiveness. Fortunately there are some excellent reviews of the literature available (Dunkin & Biddle, 1974; Smith 1971; Travers, 1973; Medley, 1977; Peterson & Walberg, 1979; Ottobre, 1980, Avalos & Haddad, 1981). It would then appear superfluous to review the reviews. What would be done, however, would be to focus on some of the highlights of the research on teacher effectiveness germane to the objectives of the present project.

The Teacher Characteristics Stage

Generally speaking, research into teacher effectiveness falls into several broad categories, each with its own focus, concern and assumptions. The first category of research focuses on teacher characteristics as indicators of teacher effectiveness. The assumption made is that more effective teachers can be distinguished by certain characteristics not possessed or possessed in a lesser degree by less effective teachers. Researchers with such a focus do not observe teachers actually teaching or attempt to measure differential learning outcome attributable to certain teacher characteristics. Researchers usually rely on "experts" in the field (usually principals, administrators and sometimes students) to identify the effective teacher. Working backwards from the sample of teachers identified as effective/non-effective, the researchers then try to identify such characteristics as teacher's age, sex, attitude towards teaching, appearance, personality variables, I.Q. or grade point average obtained in college (Medley, 1979; Rosenshine, 1979). The only piece of research with the same focus on teacher characteristics as indicators of "best" and "worst" teachers in Singapore reaches the not unexpected findings that "good" teachers are warm, systematic, stimulating and fair whilst "bad" teachers have the opposite characteristics (Derby, 1966). Research belonging to this category is flawed by both the conceptual frame used and the research designs employed to get at answers to the question of effective teaching through the iden-

tification of teacher characteristics. The characteristics identified indeed have no universal applicability.

The Effectiveness Equals Right Methods Stage

The second discernible focus in the teacher effectiveness research is more on what the teacher does than on what he is. The concern shifts to the teaching methods used by the teacher. The assumption is that if the "right" method is used then it is more effective than, say, method "B" vis-a-vis pupil learning (Medley, 1979; Rosenshine, 1979). The research design calls for the use of different methods of teaching in different classes and the criterion variable of effectiveness pertains to the average gain in knowledge attributable to the method used. Traditional versus "progressive" methods is an example of the sorts of teaching which interest researchers (Thaxton, Rothstein & Thaxton, 1977). As in research on teacher characteristics, teachers are not observed whilst they are teaching. Most of the studies come up with negative findings, that is, there is no significant difference between methods used and average knowledge gains (Anderson & Scott, 1978; Locke, 1979).

The Process-Product Stage

Another broad class of research interest in teacher effectiveness comes under the general rubric of 'process-product research' (Dunkin & Biddle, 1974; Medley 1977) in which, unlike the research interest mentioned above, teacher classroom performance is systematically observed *in situ* or via the use of video-tapes and a serious attempt is made to relate teacher performance to pupil outcomes usually measured by standardized achievement tests and more rarely affective outcomes such as self-concept or attitude towards learning. The process-product paradigm, fashionable since the early seventies, sustains its robustness for over a decade through a number of evolutionary changes which seek to accommodate perceived complexities of the teacher effectiveness issue.

Rosenshine and Furst (1971) who review and synthesise 50 of the process-product research studies claim that they have identified ten generic behaviour patterns which consistently distinguish the more effective from the less effective teacher: clarity, variability, enthusiasm, task-oriented or business-like behaviour, student opportunity to learn criterion material, indirectness, criticism, structuring, questioning and probing.

Process-Product and Teacher Effectiveness: Beyond the Stage of Innocence

Before long, however, the complexities of classroom life compel researchers to rethink the usefulness of the concept of generic behaviours in teacher effectiveness studies. Further observation and research show that certain teacher behaviours considered effective for teaching are far from generic. Indeed, many teacher behaviours are situation-specific and relate to subject-matter, grade-level, and the socio-economic status of pupils (Berliner, 1976; Berliner & Tikunoff, 1976; Brophy, 1976; Stallings, 1976). Two generic behaviours, however, originally identified by Rosenshine & Furst (1971) as discriminating indicators of effective and non-effective teachers remain robust variables: student opportunity to learn criterion material and time spent on materials, refined to allocated time and engaged time. (Berliner, 1979; Bloom 1980; Rosenshine, 1979).

A further extension of the process-product research emerges from the consistent finding that the amount of time pupils spent on the criterion materials correlates positively with pupil learning. The research proceeds to identify teacher behaviour associated with higher engaged time or allocated time. The result is a synthesis of some of the findings of the process-product research subsuming effective teacher behaviours under the concept of "Direct Instruction". Ten dimensions of "Direct Instruction" are especially promising in identifying the effective teacher. One is a management function (maintaining-on-task behaviour), six refer to instructional functions (feedback, grouping, praise, questioning, structuring and student choices) while three refer to climate dimensions of warmth, task orientation and teacher expectancy (Graham & Siedentop, 1978; Berliner, 1979; Good & Grouws, 1977; Arehart, 1979; Good & Beckerman 1978; Gage, 1976; Medley, 1979; Kounin & Doyle, 1975).

Towards a More Comprehensive Model

Some of the highlights of the research effort to get at teacher effectiveness demonstrate in no uncertain terms that progress has been achieved. Shortcomings and gaps in conceptualisation and design remain and await resolution. For example, in the process-product research, teacher's intention and pupils' learning experiences are not taken into account. A more serious problem is the use of pupils' learning outcomes as the criterion variable for teacher effectiveness, attributing to a particular teacher's performance what may be attributable to other

teachers' efforts, the students' characteristics, instructional materials and setting and the context of teaching (Millman, 1981; McKenna 1981). The problematic issue is really one of disentangling the multiple impact of various factors on pupil learning outcome to arrive at a "pure" or "unique impact of the target teacher under investigation". On the methodological issue of teacher performance Michael Scriven (1981) takes the extreme position that classroom observation to evaluate teaching either by experts or others (except for formative evaluation) is not just incorrect but a disgrace. Amongst other reasons he notes that "nothing that could be observed in the classroom (apart from the most bizarre special case) can be used as a basis for an inference to any conclusion about the merit of the teaching" and that "the number of visits is too

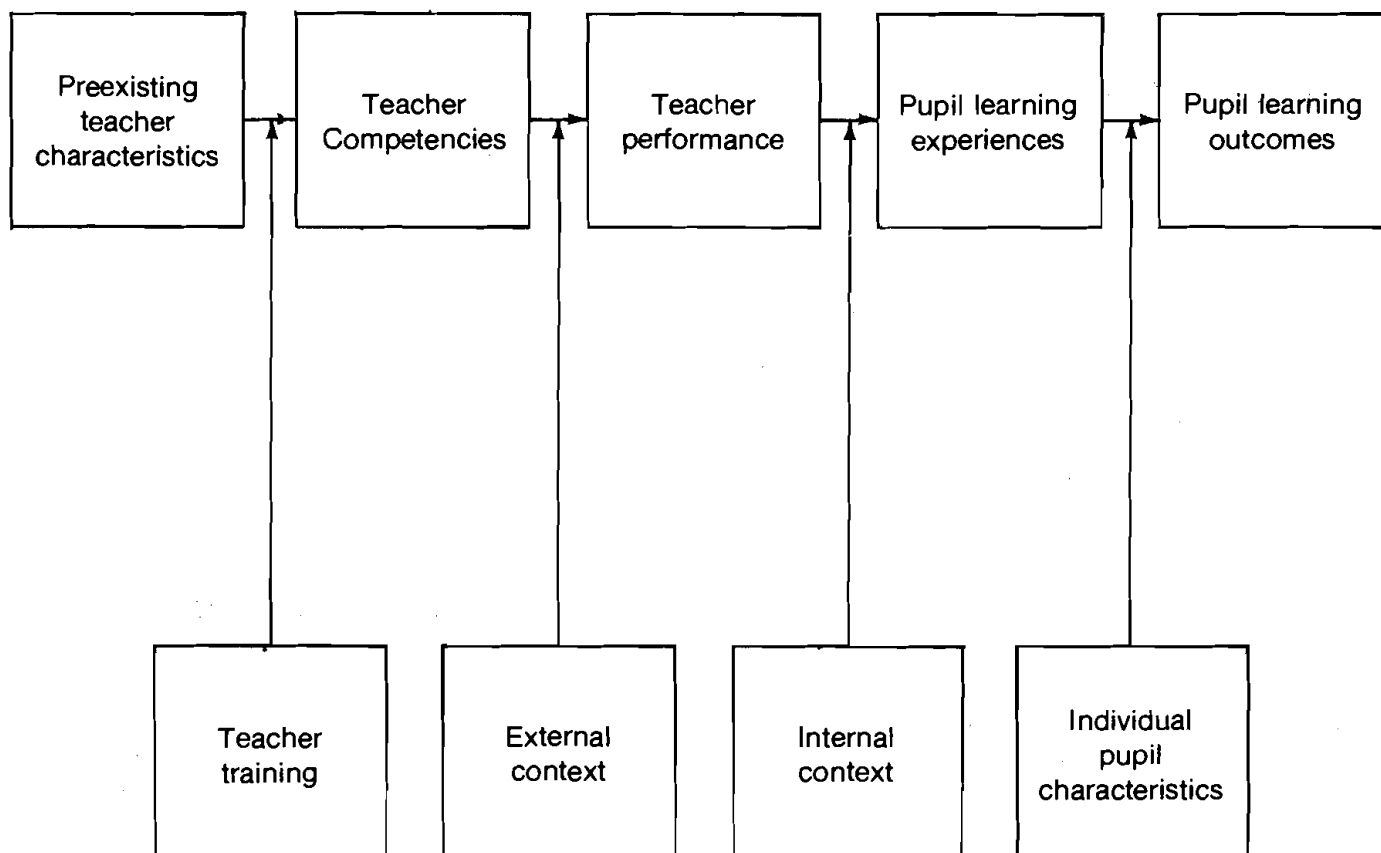
small to be an accurate sample from which to generalise, even if it were a random sample." (p 251)

In short, research into teacher effectiveness is fraught with difficulties both conceptual and methodological. Over-specification or under-specification in the model used could be misleading. Also statistical limitations could limit fine-grained analysis. Difficulties notwithstanding, research on teacher effectiveness must press on if policy decision-makers and teacher educators are to be able to formulate policies and actions that are firmly grounded in theory and empirical evidence.

To correct the underspecification of the process-product model, Medley (1982) proposes a more adequately specified model as given below.

TEACHER EFFECTIVENESS

Figure 1. Structure of teacher effectiveness



The full-blown model is derived from Mitzel's. Medley explains that the top row cells of five types of variables have been used at one time or another as criteria for evaluating teachers. The second row cells are four additional types of variables that affect the outcomes of teaching and are not controlled by the teacher. The arrows indicate the flow of impact from one variable to the next. He stresses the importance of research in distinguishing the separate influence of each on the criterion variable. Two points are worthy of note: one is the non-specification of interaction effects and the other the omission of a suggested methodology to get at indirect effects. Presumably, path analysis could be used to get at the direct and indirect impacts of each of the independent variables. Indeed, his proposal of a triangular design for investigating three types of variables in turn using each of the following as the dependent variable (Teacher competencies, Teacher performance, Pupil learning experiences and Pupil learning outcomes) seems to point to path analytic techniques as a suitable analytical tool. If interaction among the variables is more isomorphic with the social reality of the classroom, then a non-recursive model would have to be specified. Such a specification can only take place after justification can be provided by a reading of the literature to provide both empirical and theoretical support for specifying interacting phenomena. Pioneering work in interaction design and analysis has already been carried out by Cronback and his studies. For example, Winne and Marx (1977) and Winne (1977) have carried out preliminary studies using interaction analysis to throw some light on the proverbial black box of classroom life. Their works would be invaluable in variable linkage specifications in carrying out research using Medley's structure of teacher effectiveness.

Relating Screening Devices to Future Teacher Effectiveness

It is clear that the identification of the effective teacher or effective teaching is difficult enough. To come up with selection instruments with a high predictive validity (for successful teaching after graduation) for pre-service teachers (prior to training) may prove an even more daunting task. The search of the literature has not been very helpful or encouraging. Much of the literature on teacher selection refers to selection post training for entry into the teaching profession and even so selection procedures are based on presumed effectiveness rather than proven effectiveness. Pre-service selection procedures are based by and large on conventional criteria of

education attainment (eg GPAs) and the usual interview. Additionally, tests on reading ability or basic Mathematics may be administered. Variations of the interview technique have also been used but again there is no empirical evidence to support that pre-service teachers screened into training institutions would be effective teachers whilst in service.

In spite of the importance of screening techniques to ensure quality teachers for the school system, it seems surprising that very few studies seem to exist that evaluate the effectiveness of screening criteria. Turner (1975), in an overview of research in teacher education could only find one study "in which factors on which teacher education entrants might be selected were correlated with teacher success". Of the eight predictor variables, only scores from the pre-teacher training interview were related significantly to administrator rating of teaching success. Other studies (Simon and Asher, 1964; Cornett, 1969; Draba and Steinkellner, 1977; Barnes, Blaisdell and Hill, 1977; Frankel and Milgram, 1975; Van Patten, 1977) using a variety of screening techniques involving as many as twenty-five predictor variables and some non-conventional techniques such as observation of the student in a functioning, unstructured nursery or primary class, observation of the student in a structured situation and applicants reaction to a videotape of a school classroom experience, have not met with any more significant success. Ellsworth, Krepelka and Kear (1979) failed to locate any significant number of studies which set out to demonstrate the validity of various criteria so far used in predicting future effectiveness of teachers. What is evident from the few studies they reviewed is that grades are not particularly good predictors of administrator ratings of teaching effectiveness.

Conclusion

The body of literature pertaining to screening techniques/procedures and teacher effectiveness makes dismal reading not unlike the school effects literature. It does not imply, however, that they are non-researchable problems. The state of the art must necessarily limit the kind of research that can be done. What has been done so far provides the knowledge and research base for reconceptualisation of teacher effectiveness and teacher presage characteristics. A more comprehensive model similar to the one postulated by Medley coupled with some attention addressed to the criterion problem, the problem of the characteristics of the effective teacher and the problem of the evaluation of such characteristics (Schulsmans, 1980) prior to research on screen-

ing procedures would resolve the crucial problem of teacher effectiveness research and enable meaningful research on the screening of pre-service teachers problem. If the objectives of research in this area are realistic, based on the understanding that no system is foolproof and that what society ought to seek is a selection system that increases the odds above chance that the applicant chosen is in fact the best teacher available, then disenchantment with such research will not arise.

The model as postulated by Medley has promise. At the planning stage every variable has to be clearly specified, and evaluation and measurement problems resolved. Lines of influence have to be decided based on both

theoretical and empirical grounds, taking care not to overlook backward looping influences. The factors to be considered in the first cell or presage characteristics deserve utmost care if the objective is to come up with instruments which can predict effective teaching after graduation. Assuming that the presage characteristics plugged into the model can predict effective teaching, the instruments developed to measure such characteristics cannot justifiably be used as screening devices until such instruments have been used on a fresh cohort of incoming students with follow-up studies over the short and medium term to ascertain if they do in fact predict teaching effectiveness in subsequent studies.

References

- Anderson, L.W. & Scott, C.C. (1978). The Relationship among Teaching Methods, Student Characteristics, and Student Involvement in Learning. *Journal of Teacher Education*, 29(3), 53-57.
- Arehart, J.E. (1979). Student Opportunity to Learn Related to Student Achievement of Objectives in a Probability Unit. *Journal of Educational Research*, 72, 253-258.
- Avalos, Beatrice and Haddad, Wadi. (1981). *A Review of Teacher Effectiveness Research in Africa, India, Latin America, Middle East, Malaysia, Philippines, and Thailand: Synthesis of Results*. Ottawa, Canada, International Development Research Centre.
- Barnes, C., Blaisdell, L. and Hill, Shirley, (1977). Darts or Criteria? How Do We Choose Teacher Education Students?" *California Journal of Teacher Education*, IV, 63-73.
- Berliner, D.C. Tempus Educare, (1979). In P.L. Peterson & H.J. Walberg (Eds.) *Research on Teaching: Concepts, Findings and Implications*. Berkeley, CA: McCutchan, 1979.
- Berliner, D.C. (1976). Impediments to the Study of Teacher Effectiveness. *Journal of Teacher Education*, 1, 5-13.
- Berliner, D.C., & Tikunoff, W. (1976). The California Beginning Teacher Evaluation Study: Overview of the Ethnographic Study. *Journal of Teacher Education*, 27(1), 24-30.
- Bloom, B.S. (1980). The New Direction in Education Research: Alterable Variables. *Phi Delta Kappan*, 61, 382-385.
- Brophy, J.E. (1976). Reflections on Research in Elementary Schools. *Journal of Teacher Education*, 27(1), 31-34.
- Cannedy, Rodney C. (1969). The Development and Validation of Psychological Tests to Predict Employee Tension. Dissertation Abstracts. International, 30(3), p 1391.
- Cornett, Joe D. (1969). Effectiveness of Three Selective Admissions Criteria in Predicting Performance of First-Year Teachers. *The Journal of Educational Research*, LX11, 247-250.
- Derby, Orlo L. (1966). Characteristics of 'Best' and 'Worst' Teachers. *Malaysian Journal of Education*, Vol. 3, No. 2.
- Draba, Robert E. and Steinkellner, L.L. (1977). "Screening Applicants for Teacher Training." *The Educational Forum*, XL11, 101-109.
- Dunkin, Michael J. & Biddle, Bruce, J. (1974). *The Study of Teaching*. N.Y. Holt, Rinehart and Winston, Inc.
- Ellsworth, Randy; Krepelka, Ernie; and Kear, Dennis, (1979) Teacher Education and Dilemmas of Selecting Students for Teacher Training. *Teacher Educator*, V. 15, No. 1, summer, 21-31.
- Frankel, J. and Milgram, Noel I. (1975). Developing an Educational Audition, *Childhood Education*, L1, 309-311.
- Gage, N.L. (1976). A Factorially Designed Experiment in Teacher Structuring, Soliciting and Reacting. *Journal of Teacher Education*, 27(1), 35-38.

- Good, T. & Beckerman, T. (1978). Time on Task: A Naturalistic Study in Sixth Grade Classrooms. *The elementary School Journal*, 78, 192-201.
- Good, T. & Grouws, D. (1977). Teaching Effects: A Process-Product Study in Fourth Grade Mathematics Classrooms. *Journal of Teacher Education*, 28(3), 49-54.
- Graham, G. & Siedentop, D. (1978). Developmental Stages of Teaching. In L. Gedvilas & M. Kneer (eds.) *Proceedings of the NAPECW/NCPEAM National Conference*. Chicago, 111: University of Chicago Circle Publication services.
- Kounin, J.S. & Doyle, P.H. (1975). Degree of Continuity of a Lesson's Signal System and Task Involvement of Children. *Journal of Educational Psychology*, 67, 159-164.
- Locke, L. (1979). Learning from Teaching. In J. Jackson (Ed.) *Theory into Practice*. Victoria, British Columbia: University of Victoria Press.
- Medley, D.M. (1979). The Effectiveness of Teachers. In P.L. Peterson & H.J. Walberg (eds.), *Research on Teaching: Concepts, Findings and Implications*. Berkeley, CA: McCutchan.
- Medley, Donald M. (1977). *Teacher Competence and Teacher Effectiveness: A Review of Process-Product Research*. Washington, D.C., American Association of Colleges for Teacher Education.
- Medley, D.M. (1982). Teacher Effectiveness. In H.E. Mitzel, (ed), *Encyclopedia of Educational Research* (5th Ed.) New York: The Free Press, Vol. 4.
- McKenna, Bernard H. (1981). Context/Environment Effects in Teacher Education. In Jason Millman (ed), *Handbook of Teacher Evaluation*. London, Sage Publications.
- Millman, Jason (1981). Student Achievement as a Measure of Teacher Competence. In Jason Millman (ed), *Handbook of Teacher Evaluation*. London, Sage Publications.
- Mosher, Lawrence, (1971). Court Rules Against Job Tests Screening Out Blacks. *National Observer*, March 15, p6.
- Ottobre, Frances M. (1980). *Assessing Teacher Effectiveness*. New Jersey, International Association for Educational Assessment.
- Peterson, P.L. & Walberg, H.J. (eds.) (1979). *Research on Teaching: Concepts, Findings and Implications*. Berkeley, CA., McCutchan.
- Rosenshine, B. (1979). Content, Time and Direct Instruction. In P.L. Peterson & H.J. Walberg (eds.), *Research on Teaching: Concepts, Findings and Implications*. Berkeley, CA: McCutchan.
- Rosenshine, B & Furst, N. (1971). Research in Teacher Performance Criteria." In B.O. Smith (ed.) *Research in Teacher Education*. Englewood Cliffs, N.J: Prentice-Hall.
- Smith, B.O. (Ed) (1971). *Research in Teacher Education: A Symposium*. Englewood Cliffs, N.J., Prentice-Hall.
- Schulsmans, Kathleen. (1980). What is an Effective Teacher. In Frances M. Ottobre (ed.), *Assessing Teacher Effectiveness*. New Jersey, International Association for Educational Assessment.
- Scriven, Michael (1981). Summative Teacher Evaluation. In Jason Millman (ed) *Handbook of Teacher Evaluation*, London, Sage Publications.
- Simun, P. Bates and Asher, J.W. (1964). The Relationship of Variables in Undergraduate School and School Administrators' Rating of First-Year Teachers. *The Journal of Teacher Education*, XV, 293-302.
- Stallings, J. (1976). How Instructional Processes Related to Child Outcomes in a National Study of Follow Through. *Journal of Teacher Education*, 27(1), 43-47.
- Thaxton, A.B., Rothstein, A.L. & Thaxton, N.A. (1977). Comparative Effectiveness of Two Methods of Teaching Physical Education to Elementary School Girls. *Research Quarterly*, 48, 420-427.
- Travers, M.W. (ed.) (1973). *Second Handbook of Research on Teaching*. Chicago. Rand McNally.
- Turner, Richard L. (1975). An Overview of Research in Teacher Education. In Kevin Ryan (ed.), *Teacher Education, The Seventy-Fourth Yearbook of the National Society for the Study of Education*. Part II, Chicago, University of Chicago Press.
- Van Patten, J.I. (1977). Selection and Education of Prospective Teachers: The Role of the First Education Course. *The Journal of Teacher Education*, XXCIII, 9-14.
- Winne, P.H. (1977). Aptitude Treatment Interactions in an Experiment on Teacher Effectiveness. *American Educational Research Journal*, 14, 389-409.
- Winne, P.H. & Marx, R.W. (1977). Reconceptualizing Research on Teaching. *Journal of Educational Psychology*, 69(9), 668-678.