DOCUMENT RESUME

ED 137 321 TM 006 111

AUTHOR Sirois, Herman A.; Iwanick, Edward F.
TITLE Delphi-Discrepancy Evaluation: A Model for the Quality Control of Federal, State, and Locally Mandated Programs.
PUB DATE [Apr 77]
NOTE 25p.; Paper presented at the Annual Meeting of the American Educational Research Association (61st, New York, New York, April 4-8, 1977)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Educational Programs; *Evaluation Methods; Evaluation Needs; Formative Evaluation; *Models; Needs Assessment; *Program Evaluation; Summative Evaluation
IDENTIFIERS *Delphi Technique

ABSTRACT

Legally mandated educational programs often lack specificity and guidelines for such programs are often vague and subject to considerable variability in interpretation. This situation presents perennial problems for evaluators. Few evaluation models have the flexibility for dealing with this ambiguity and variability while at the same time achieving program evaluations which are both formative and summative. The Delphi-Discrepancy Evaluation model is designed to assess and prescribe. The five phases of the model include: (1) establishing a standard by using the procedures of the Delphi Technique, (2) gathering data about programs based on the standard, (3) determining the discrepancy status of programs, (4) analyzing these discrepancies in relation to criteria and other programs in the population, and (5) prescribing and remediating based on the analyses of discrepancies. Bringing together successful techniques such as the Delphi Method, Discrepancy Evaluation, and the notions of Euclidean Distance to bear on a particular evaluation problem should facilitate analyses and render decision-makers with more valid information for formative and summative purposes. (Author)

DELPHI-DISCREPANCY EVALUATION: A MODEL FOR THE QUALITY CONTROL OF FEDERAL, STATE, and LOCALLY MANDATED PROGRAMS

Herman A. Sirois

Edward F. Iwanick

University of Connecticut

Paper presented at the annual convention of the American Educational Research Association: New York, April, 1977.

Delphi-Discrepancy Evaluation:
A Model for the Quality Control of Federal, State and Locally Mandated Programs (EVA)

HERMAN A. SIROIS EDWARD F. IWANICKI
University of Connecticut

-ABSTRACT-

Problem. Legally mandated educational programs often lack specificity and guidelines for such programs are often vague and subject to considerable variability in interpretation. This situation presents perrenial problems for evaluators. Few evaluation models have the flexibility for dealing with this ambiguity and variability while at the same time achieving program evaluations which are both formative and summative.

Methodology. The Delphi-Discrepancy Evaluation model is designed to assess and prescribe. The five phases of the model include: (1) Establishing a standard by using the procedures of the Delphi Technique. (2) Gathering data about programs based on the standard. (3) Determining the discrepancy status of programs. (4) Analyzing these discrepancies in relation to criteria and other programs in the population, and (5) Prescribing and remediating based on the analyses of discrepancies.

Conclusions. Bringing together successful techniques such as the Delphi Method, Discrepancy Evaluation, and the notions of Euclidean Distance to bear on a particular evaluation problem should facilitate analyses and render decision-makers with more valid information for formative and summative purposes.

TABLE OF CONTENTS

LIST OF FIGURES

The evaluation of federal, state, and local educational programs has become a large volume and high finance business. The evaluation industry has evolved around a growing cult of "experts," each passing the cult's initiation rites of developing a new or modified evaluation model which promises to resolve the perennial evaluation problems of methodology, application and implementation, relevance, impact, value, policy, and politics.

Of the numerous evaluation models in existence, such as Scriven's Formative-Summative model, his Goal-Free model, his Modus Operandi model, Stake's Countenance model, Stufflebeam and Guba's CIPP (content, input, process, product) model, Provus' Discrepancy model, Rippey's Transactional model, Owens' Adversary model, Wergin's Political model, and Cook's PERT (program evaluation and review technique), etc., almost all were designed to meet specific needs of the various programs to which they were initially applied. Much of the development (and progress) in educational evaluation appears to be a result of the failure of one model or another

of evaluation to meet the evaluation needs of a program when applied inappropriately to that program, and the subsequent refinement and/or redevelopment of an evaluation model to meet the unique needs of the situation or program in question.

A lesson we can derive from the history of educational program evaluation, to date, is that no one evaluation model can serve all evaluation needs. Rather, the evaluation needs of each particular program or situation must be analyzed and a program-specific evaluation plan must be developed, drawing from any and all existing sources of evaluation models, to meet the information, feedback, judgemental, and decision-making needs of the program under consideration. A Popham suggests, educational evaluators should disregard Polonius' dictum: "Neither a borrower nor a lender be (Popham, 1975, p.42)."

The evaluation design presented in this paper is a hybrid model, begged and borrowed from proven models in the areas of educational evaluation, educational planning and futuristics, and special education needs assessment. This model was developed to meet both the formative and summative evaluation demands of mandated educational programs where at least loosely defined guidelines or criteria exist.

The concepts of Delph-Discrepancy Evaluation is presented as a model for district-wide, state-wide, and/or nation-wide monitoring and needs assessment of mandated educational programs. The model, originally presented in May, 1976 to the Connecticut State Department of Education by Sirois & Levin (1976), was designed to meet the state department's desires to (1) monitor mandated LEA teacher evaluation programs in relation to state guidelines, and

(2) diagnose specific LEA needs (inservice, assistance, etc.) in relation to developing teacher evaluation program.

The model incorporates the Delphi Technique for establishing critical or criterion needs, and the notions of discrepancy evaluation for arriving at judgements regarding the remedial needs of local school districts. The model is essentially a blend of the Delphi Planning Technique (Weaver, 1973), the Discrepancy Evaluation model (Provus, 1971), and the notions of Euclidean Distance (from Gable & Gillung, 1976), and is appropriately named the "Delphi-Discrepancy" model of program evaluation.

The Delphi Planning Technique

The Delphi Planning Technique was originally developed as the Delphi Forecasting Method by Olaf Helmer at the Rand Corporation "think tank" during the mid-sixties (Helmer, 1967). Cyphert & Gant (1971) characterize the Delphi Technique as a model or set of procedures for achieveing consensus. Weaver defines the process as "...an intuitive methodology for organizing and sharing 'expert'" opinion (1973, p. 44).

In its ideal application, the Delphi Technique has been used to achieve a consensus of opinion amongst experts regarding the probability of the occurrance of certain events at some future time. The methodology has also been used to estimate the desirability of the occurance of such events. Redefining the Delphi Technique in terms of its applicability to the monitoring of state-mandated teacher evaluation, it might read: A technique

to achieve consensus amongst selected local school districts, previously judged to either have or have the potential of having "good" programs, "good", that is, in regards to the elements and activities that might comprise the "ideal" program.

Weaver's concise summary of the Delphi Technique (1973, p. 45) is paraphrased for the present context: Typically, the procedure includes a questionnaire, mailed to respondents who remain anonymous to one another. Respondents first generate several rather concise statements of criteria and/or activities that they consider important to the mandated educational program under consideration. In a second round, respondents are provided with a coalated list of all respondents' criteria and are asked to prioritize the items on the list in terms of their importance to the mandated educational program. A third round provided each respondent with another coalation of items, this time in rank order according to the averaging of the prioritization responses in round two. At this time, respondents are asked to prioritize the items once again either agreeing or disagreeing with the averaged list provided them. The final coalation of the averaged prioritization of the items becomes the "standard" against which comparisons are made.

Discrepancy Evaluation

Discrepancy Evaluation, as developed by Provus (1971), begins at the point of establishing a "standard". In program evaluation, this standard is usually considered to be the design of the program.

In the present application, the Delphi Technique accomplishes this task. This design or standard is next compared to the actual initial installation of the program. If a discrepancy exists, it is corrected before the program is allowed to procede. The third step in the discrepancy model is to compare the standard to the "process" which develops and exists after the program has been installed. Again, discrepancies are noted and remediated. A fourth step in the model compares the products or outputs of the program to those which were hypothesized in the design or standard. A final stage involves cost-design discrepancies.

As with the Delphi Technique, the application of Discrepancy analysis to the monitoring of mandated educational programs involves analogous procedures and techniques. Once a "standard" has been developed (via the Delphi Technique), local school district educational programs are compared to the appropriate standard(s) at appropriate stages of development.

Delphi-Discrepancy Evaluation

Delphi-Discrepancy Evaluation has a dual purpose: to assess and to prescribe. The five phases of the Delphi-Discrepancy Evaluation model described below are designed to accomplish these two purposes.

The phases of the Delphi-Discrepancy Evaluation model are described in context below as they would apply to the evaluation of a specific mandated educational program, i.e. state-mandated teacher evaluation in Connecticut.

In 1974, the Connecticut State Legislature mandated teacher evaluation programs of all LEAS. The Connecticut State Department

of Education (CSDE) developed a set of eleven guidelines to which all such programs would necessarily conform. The Educational Resources and Development Center (ERDC) at the University of Connecticut was contracted to conduct annual evaluations of state-wide teacher evaluation program progress.

The data from the ERDC evaluations represent the point of departure for the following description of the Delphi-Discrepancy Evaluation Model as applied to state-mandated teacher evaluation in Connecticut.

Phase I: Establishing a Standard

Phase I involves the use of the Delphi Technique for the purpose of generating a set of indicators or "standards" which will form the basis for the discrepancy analysis and prescriptive-remedial phases of the Delphi-Discrepancy model. The sequence of tasks in phase I are itemized below:

(1) Identification of an "Expert" Sample: Using the prior CSDE/ERDC ratings of Connecticut local school district teacher evaluation programs (Sheathelm, et. al., 1975; Sirois, et. al., 1976), determine which fifteen or twenty districts have been rated as having the "best" teacher evaluation programs. This group will be considered the "expert" respondents in the Delphi phase of the Delphi-Discrepancy Evaluation process.

(2) For each of the eleven state teacher evaluation guidelines, request of the respondents a detailed list (akin to brainstorming) of criteria, indicators, activities, etc. which they consider to be indicative of "good" teacher evaluation programs.

(3) Coalate all responses (i.e., within each of the eleven categories as defined by the eleven categories as defined by the

guidelines).

(4) Provide each respondent with this coalated list, asking each respondent to prioritize all items within each of the eleven guidelines.

(5) Secure an average priority rating for each item and coalate all items again, rank-ordering them under each of the eleven guidelines.

(6) Provide each respondent with this coalated, prioritized list. Ask each respondent to re-prioritize the lists either agreeing with the averaged priorities as they appear on the lists or changing them as opinion may dictate.

(7) Perform a final coalation of the eleven lists of items, again ranked by average priority rating. (For convenience, only the top ten, fifteen, or twenty items might be retained for the final lists). This final list should take the form of the sample displayed in Figure 1.

(8) The final form of the list(s) as presented in Figure 1 will serve as a working checklist for the "standard". This form should include three rating categories with values: (0)=indicator not in operation at all, (1)=indicator in partial operation, and (2)=indicator is in full operation (or has occurred fully) in the district's teacher evaluation program.

(9) The indicators should be weighted according to their priority ranking. One method for such a weighting has been used in the sample in Figure 1. The indicator ranked as top priority is given the greatest weight, in this case a value of 10. The second ranked indicator is assigned a value of 9, and so forth. At a later stage in analysis, these values are multiplied by the ratings

FIGURE 1: A SAMPLE PRIORITIZED LIST OF INDICATORS FOR GUIDELINE I.

GUIDELINE I: Each professional shall cooperatively determine with the evaluator(s) the objectives upon which his or her evaluation shall be based.	RATING		
	COMPLETE 2	PARTIAL 1	NOTHING 0
10. Written objectives for the evaluatee are developed.			
9. Objectives are developed cooperatively.			
8. Procedures for selecting and agreeing on objectives are developed in writing.			
2. Objectives are stated in measureable terms.			
1. A master list (or pool) of objectives is developed for the school system.			

(0, 1, or 2) which results from the data gathering phase (II).

Phase II: Determining the "Status" of Local Educational Programs

The checklists developed in phase I may be used to gather data in a number of ways. Each of these methods utilizes the notions of discrepancy evaluation (Provus, 1971). In each case, the indicators on the checklists represent the optimum standards, which the ratings secured from observations and/or self-reports represent degrees of actual status of local teacher evaluation programs.

"Type A" Discrepancy: In phase II of Delphi-Discrepancy Evaluation we are seeking to eliminate any occurrance of a "type A" discrepancy, and arrive at a collaboration as to the "status" of the teacher evaluation program in question. Each of the conditions of data gathering requires collaboration from at least two sources. A type A discrepancy is said to occur when any of the conditions of data gathering, described below, result in significantly different ratings on any of the checklists.

At least three options or conditions of data gathering exist for the Delphi-Discrepancy evaluator:

(1) On-Site Visitation/Superintendent Report: In this case, the superintendent completes the checklist(s) while a representative of the mandating agency (CSDE) conducts an on-site visitation-interview to complete a second such checklist. A "type A" discrepancy exists if the two assessments are significantly different.

(2) Local Reference Groups Method: The superintendent, a principals' group, and one or more teachers' groups complete separate checklists. A "type A" discrepancy exists if the superintendent's report differs significantly from either of the other two reports.

(3) The Adversary Method: The superintendent appoints two parties or teams to complete separate checklists. One party is designated as a critical adversary, while the other party represents the superintendent. A type A discrepancy exists if the two reports are significantly different.

A "significant" difference between groups cannot be determined statistically when each group consists of only one subject as is called for in the present design. A somewhat arbitrary level of significance must, therefore, be established. This fiat is easily accomplished: (1) Determine what is the highest possible "score" on the particular guideline checklist by summing the products of two (rating category 2=completion of indicator) times the weightings of each indicator (this total might be approximately 110 for any guideline in the present application). (2) Determine the significant discrepancy value by taking 15% of this total (in the present example the significant discrepancy value would be 17). (3) Consider any type A discrepancy which has a value greater than the significant discrepancy value to be a significant type A discrepancy.

Resolving a Type A Discrepancy: If a type A discrepancy exists, it must be resolved prior to testing for "type B" discrepancy in phase III of the Delphi-Discrepancy model. Three methods of resolution of type A discrepancies are suggested here.

(1) First, for insignificant type A discrepancies, the various forms of the checklists may simply be averaged for a consensus rating.

(2) A second method attempts to resolve a significant type A discrepancy by returning the forms to the respondents and asking them to get together to complete a consensus form of the checklist. Failing consensus at this point, the third method of discrepancy resolution is instituted.

(3) The third method also attempts to resolve a significant type A discrepancy. This discrepancy is referred to the mandating agency (CSDE) for further verification, on-site visitation, etc. in an attempt to gather an accurate assessment of the status of the teacher evaluation program in question.

Whatever method is used to resolve a type A discrepancy, the resolution becomes the "status" of the particular teacher evaluation program under consideration.

Phase III: "Type B" Discrepancy: Standard - Status

The "status" of any particular teacher evaluation program is represented by the ratings (on the eleven checklists) assigned to that program after any type A discrepancy has been resolved during phase II of the Delphi-Discrepancy Evaluation process.

"Type B" Discrepancy: Once the question of a type A discrepancy has been resolved and the "status"of a particular teacher evaluation program has been determined, we may then analyze the nature of any "type B" discrepancy which may exist. This is done by comparing the "status" of teacher evaluation programs to the "standard". As defined in Delphi-Discrepancy Evaluation, a "type B" discrepancy is not considered to be an absolute measure. Rather, it is considered to be the relative discrepancy of a district's teacher evaluation program when compared to the "standard" in relation to all other

district teacher evaluation programs.

A variation in the application of the Euclidean Distance formula as suggested by Gable & Gillung (1976) will provide an index or number, the magnitude of which indicates the similarity between the status" and the "standard" (a low value indicating a high similarity). By generating such an index for each school district teacher evaluation program, comparisons can be made across districts on the basis of such variables as per pupil expense, size, unban-suburban-rural, mean income, etc.. More important for the immediate purpose of the CSDE (the mandating agency), such an index will enable a ranking of school districts according to the Euclidean Distance values (by guideline) thus facilitating the identification of the remedial needs of local school districts in relation to their teacher evaluation programs.

$$\Delta g = \sqrt{\frac{\sum A_{ig} - I_{ig})^2}{NG}}$$

Where: Δg = the Euclidean Distance value for guideline number g.

A_{ig} = the actual status rating for indicator number i of guideline number g.

I_{ig} = the ideal (standard) rating for indicator number i of guideline number g.

$\sum$: indicates that the value $(A_i - I_i)^2$ is computed for each indicator and these values are summed under each guideline.

Ng = the number of items or indicators under guideline number g.

A number of Euclidean Distance values may be generated. First, and most appropriate, eleven such values can be obtained for each

local teacher evaluation program (one for each of the eleven guidelines). Other such values may be obtained by considering groups of guidelines together for comparison to other districts or to other groups of guidelines, or simply totalling the eleven Euclidean Distance values for gross comparisons between districts. (It should be noted that in such variations, the values I_i and N_g would vary).

Any resulting Euclidean Distance value is not interpetable in any direct manner. Rather, the value and power of Euclidean Distance values lies in the fact that they may be compared directly to other Euclidean Distance values of the same population. It can be concluded, for example, that a teacher evaluation program in district X, with a distance value of 5 on guideline number I, is in need of greater remedial assistance in relation to such components as setting objectives and agreeing on objectives for evaluation cycles than in district Y, which may have received a distance value of 2 on guideline number I.

Phase IV: Analysis of Discrepancies

Given the standardized quality of the Euclidean Distance values and the reliability and utility of their comparison to other Euclidean Distance values from the same population, four different analytical procedures are suggested here to assist the CSDE in (1) assessing statewide status, and (2) developing remedial plans in relation to local teacher evaluation programs. The first three of these procedures "block on" or compare the various local teacher evaluation programs to each other in relative terms. The fourth procedure blocks on or compares the eleven state teacher evaluation guidelines to each other.

(1) Rank Order: Based on the Euclidean Distance values, a rank ordering of the local teacher evaluation programs should be performed. The higher the Euclidean Distance value, the more a particular teacher evaluation program departs from the "standard", and is in need of greater attention and/or assistance from the CSDE. Such a rank order, if performed for each of the eleven state teacher evaluation guidelines, would provide considerable diagnostic information for the CSDE in its efforts to assist and remediate. Also, such rank orderings would yield a direct measure of the "status" and "compliance" of local teacher evaluation programs in relation to the CSDE guidelines.

(2) Clustering: Based on the rank ordering of the teacher evaluation programs conducted in procedure number one above, clusters of school districts within each guideline may be identified (either visually or by a computer sub-program) which are similar in terms of their stages of development, level of "compliance" to guidelines, and level of attention and/or remediation required by their teacher evaluation programs.

(3) Item Analysis: As an optional procedure, for a more detailed diagnostic analysis, the CSDE may wish to cluster local teacher evaluation programs based on their Euclidean Distance values on each item or indicator within each of the eleven guidelines. Such analysis would not only identify the particular guidelines which are presenting problems for local districts, but would provide the CSDE with information regarding the particular indicators which are contributing to such problems. (It should be noted that changes in certain computational procedures should be considered when analyzing

data based on indicators: e.g. the weightings of the indicators would necessarily be ignored and not entered into any computations).

(4) Guideline Block Analysis: Procedures one, two, and three are designed to compare the various local teacher evaluation programs to each other. Procedure number four compares, on a state-wide basis, each of the eleven state teacher evaluation guidelines with each other. The value of such a comparison is that those guidelines which are causing problems for many school districts are readily identified. This, of course, would have definite implications for the focus of state-wide teacher evaluation workshops, and other large scale CSDE teacher evaluation activity.

Note: The four procedures suggested above do not exhaust the analytical prossibilities nor do they resolve all the problems of assessment of status and identification of necessary remediation. The Delphi-Discrepancy evaluator should be prepared to conduct further analysis tailored to specific information needs of the mandating agency and of the local school districts. A combination of procedures numbers two and four, for example, would produce a cross sort table indicating the priority needs of both local districts and specific guidelines. Such an analysis is displayed in Figure 3 below and is discussed in phase V of this presentation of the Delphi- Discrepancy Evaluation model.

FIGURE 2: DELPHI-DISCREPANCY, EUCLIDEAN DISTANCE CROSS SORT::: DISTRICTS X GUIDELINES

DISTRICTS	GUIDELINES I	II	III	IV	XI
1	3	8	4	21	8
2	7	16	4	19	4
3	20	15	18	23	21
4	0	4	3	20	2
5	19	6	8	18	1
6	2	1	6	22	5
169	4	4	8	21	8

*The figures within this cross sort table are hypothetical Euclidean Distance values.

Phase V: Planning for Assistance, Prescription, and Remediation: The Use of Discrepancy Analysis

The CSDE (or mandating agency) will undoubtedly wish to provide assistance to local teacher evaluation programs, within the perennial constraints of available funds. The results of the Delphi-Discrepancy Evaluation procedures can assist the mandating agency in planning workshops and other more direct in-service activities which can be designed to meet both the particular and the common needs of the various school districts in an efficient manner. A discussion of Figure 3 should demonstrate how the analytical procedures suggested in phase IV of the Delphi-Discrepancy model can achieve this purpose. Please note that the discussion below is based on the purely hypothetical data from Figure 3.

Referring to Figure 3 and recalling that a larger Euclidean Distance value indicates a need for attention and/or assistance in relation to the teacher evaluation program in question, we notice that certain cells in the cross sort have obviously greater values than do other cells. Cell III-3 (guideline #III, district #3), for example has a noticably larger value than any other district under guideline III.

A look at guideline IV reveals that all districts are having problems with this guideline. This would suggest to the mandating agency that a one-topic, state-wide workshop in relation to guideline IV (job descriptions) might be in order as a remedial technique.

Considering the values under guideline I, however, such a state-wide workshop dealing with setting objectives (guideline I) would not be advised. A small group workshop, however, or some other form of assistance to school districts numbers 3 and 5 does seem warranted in relation to the high values of these two districts on guideline I.

Finally, it appears from the values in Figure 3, that district number 3 is in need of some special attention and assistance, since it has high values under every guideline. General workshops for district number 3 might prove useless at this point. Some intense consultation, financial assistance, etc. from the mandating agency might be in order for distric number 3.

The above discussion, hopefully, has made clear the remedial applications of the Delphi-Discrepancy Evaluation model. Without the actual data resulting from a complete application of the Delphi-Discrepancy method, specific examples and suggestions for in-service and remedial assistance, of course, cannot be made. The point is, simply, that the results of the Delphi-Discrepancy analysis can have definite implications for remediation; specific remedial needs of local districts can be identified via the Delphi-Discrepancy method. The effective and creative use of the model can render such assistance both more efficient and considerably more effective.

Conclusion

The methodology of the Delphi-Discrepancy Evaluation model as presented in this paper appears to be of significant value in achieving the dual purpose of (1) assessing the status of local educational programs in relation to guidelines established by mandating agencies, and (2) identifying the specific remedial needs of such educational programs in a diagnostic manner. Although the techniques that make up the Delphi-Discrepancy model (i.e., the Delphi Technique,

Discrepancy Evaluation, and the notions of Euclidean Distance) have not yet been applied or tested together, each has proven quite valuable and reliable in isolated applications of the various techniques. Bringing all three techniques together to bear on an evaluation problem such as that posed by the Connecticut State Teacher Evaluation Act (P. A. 74-278) should facilitate the analysis of the problem and render decision-makers sets of valid and reliable data on which to base judgements.

The use of the Delphi-Discrepancy Evaluation model is not limited to the applications as presented herein. Other programs, state, local, and federal would benefit from appropriately designed comprehensive evaluation models such as the Delphi-Discrepancy model. Again as noted earlier, an evaluation model such as the Delphi-Discrepancy technique should not be simply superimposed on a program. Rather, the evaluation needs of the program should be analyzed and the evaluation model should be modified and adapted to suit the program needs.

References

Cook, Desmond L. PROGRAM EVALUATION AND REVIEW TECHNIQUE. Washington: Bureau of Research, U. S. Govt., 1966.

Cyphert, Fredercik, & Walter Gant. The delphi technique. In Ernest R. House (Ed.), SCHOOL EVALUATION. Berkeley, Cal.: McCutchan, 1973.

Gable, R. K., & T. Gulling. A NEEDS ASSESSMENT MODEL FOR IDENTIFYING SPECIAL EDUCATION COMPETENCIES FOR REGULAR CLASS TEACHERS. Paper presented at the New England Educational Research Organization Annual Meeting, Provincetown, Mass., May, 1976.

Gage, N. L. MANDATED EVALUATION OF EDUCATORS. Stamford University: Stamford Center for Research and Development in Teaching, 1973.

Helmer, Olaf. ANALYSIS OF THE FUTURE: THE DELPHI METHOD. Santa Monica, Cal.: Rand Corp., 1967.

House, Ernest R. (Ed.). SCHOOL EVALUATION. Berkeley, Cal.: McCutchan, 1973.

Iwanicki, Edward F. ACTIVITY ANALYSIS: AN APPROACH TO IMPROVING THE EVALUATION DESIGN OF ONGOING EDUCATIONAL PROGRAMS. Paper presented at the Annual Meeting of the National Council on Measurement in Education, Chicago, April, 1974.

Popham, W. James. EDUCATIONAL EVALUATION. Englewood Cliffs, N. J.: Prentice-Hall, 1975.

Provus, Malcom M. DISCREPANCY EVALUATION. Berkeley, Cal.: McCutchan, 1971.

Rippey, Robert M. STUDIES IN TRANSACTIONAL EVALUATION. Berkeley, Cal.: McCutchan, 1973.

Sheathelm, Herbert H., J. Baker, W. Lopes, & T. Sergi. TEACHER EVALUATION IN CONNECTICUT, 1974-1975. Educational Resources and Development Center/University of Connecticut, 1975.

Sirois, Herman A., & Nehemia Levin. STATE MANDATED TEACHER EVALUATION: THE CASE OF CONNECTICUT. Connecticut State Department of Education/Bureau of Research, Planning, and Evaluation, May, 1976 (unpublished).

Sirois, Herman A., J. Senteio, T. Sergi, & H. Sheathelm. TEACHER EVALUATION IN CONNECTICUT, 1975-1976. Educational Resources and Development Center/University of Connecticut, 1976.

Weaver, W. Timothy. The delphi forecasting method. In Richard W. Hostrop (Ed.), FOUNDATIONS OF FUTUROLOGY IN EDUCATION. Homewood, Ill.: ETC, 1973.

Wergin, Jon F. The evaluation of organizational policy making: a political model. REVIEW OF EDUCATIONAL RESEARCH, 1976, 46, 1, 75-115.

Wolf, Robert L. Trial by jury: a new evaluation model. PHI DELTA KAPPAN, 1975, 57, 3, 185-188.