



12-1-1978

Evaluative Procedures for Staff Development Programs

Thomas P. Fitzgerald

Ellen F. Fitzgerald

[How does access to this work benefit you? Let us know!](#)

Follow this and additional works at: <https://commons.und.edu/tl-journal>



Part of the [Scholarship of Teaching and Learning Commons](#)

Recommended Citation

Fitzgerald, Thomas P. and Fitzgerald, Ellen F. (1978) "Evaluative Procedures for Staff Development Programs," *Journal of Teaching and Learning*: Vol. 4: Iss. 1, Article 3.

Available at: <https://commons.und.edu/tl-journal/vol4/iss1/3>

This Article is brought to you for free and open access by UND Scholarly Commons. It has been accepted for inclusion in *Journal of Teaching and Learning* by an authorized editor of UND Scholarly Commons. For more information, please contact und.common@library.und.edu.

Evaluative Procedures for Staff Development Programs

Thomas P. Fitzgerald
Bureau of Reading Education
New York State Education Department
and

Ellen F. Fitzgerald
State University of New York at Albany

While staff development programs have increased in number and importance over the past ten years, efforts to evaluate their content and impact have not improved proportionately. Most efforts to review the effectiveness of training programs are limited to the collection of basic information such as numbers attending, type of material presented and reactions of individuals to the worth of the sessions (Moffitt, 1963). In short, basic program features have been tabulated without an accompanying analysis of interactions, tone and quality of individual presentations. Consequently, the process and product involved in in-service programs, as well as the implementation of principles presented, have seldom been effectively analyzed.

Although the need for improved evaluations has been recognized often, several problems are inherent in designing appropriate evaluation instruments. The first real obstacle stems from the lack of measurable objectives basic to constructing process and product evaluations. Frequently, program leaders either fail to recognize the importance of measurable objectives to see them as impractical due to time constraints, lack of pre-planning sessions with consultants, or the use of several consultants during the program. Yet the fact remains that without measurable objectives, sessions tend to reflect the presenters' interests, interpretations and even biases rather than the focus designed for the inservice program.

Other impediments to improved evaluation can be attributed to a lack of expertise or necessary resources. In spite of the fact that analysis and support of staff efforts to apply information provided

in the classroom may be recognized as a key issue in planning inservice programs, the application stage does not receive proper emphasis due to the scarcity of human and material resources as well as the cost involved in effective follow-through.

Finally, the very dearth of instruments and methods for process evaluation restrains progress in inservice evaluation efforts. Current practices of gathering basic information and applying statistical treatments contribute little to an effective analysis of tone and quality of sessions or process evaluation. However, devising more appropriate instruments again requires time, resources and expertise.

In this study, selected instruments were employed during a series of training sessions in order to establish (1) whether they would produce the information necessary to determine the quality, tone and interactions occurring and (2) whether the ethnographic research paradigm would produce the information necessary to improve and re-direct inservice programs. This research approach emphasizes questionnaires, interviews and observations in gathering data for a balanced study (Strang, 1962; Wold and Tymitz, 1976). Through such analysis, participants in the inservice program can evaluate themselves with regard to workshop effects of perceptions, knowledge and application of principles.

The objective of the present study was to involve both presenters and participants in evaluating the quality, tone and interactions that occurred during each session. In order to accomplish this end, self-analysis, both introspective and retrospective, was required through the use of selected instruments. Three areas were to be evaluated: pre-session estimates, during-session interaction and post-session judgment. The following instruments were thought to have potential application to these areas: "ideal presentation" represented pre-session judgments by presenters and post-session judgments by the evaluation team; the Flanders Interaction Scale (Flanders, 1965) provided a coding of both verbal and behavior activities; and questionnaires completed by the

audience (at the end of the session) evaluated the tone and quality of the presentation.

These instruments were used in New York state during two Right to Read institutes designed to improve reading instruction as a necessary step to eliminating illiteracy. In New York, as in other states, leadership training for district personnel was established with the objective that participating districts would, as a result of the training, construct a comprehensive reading improvement plan. The training institutes were organized around measurable objectives for each topic and consultant. Within this framework, this paper addresses the questions of the effectiveness of each presentation and the appropriateness of redefining the continuing needs of participants (Fitzgerald and Marino, 1976). More precisely, the following questions have been addressed:

1. Does an analysis of "ideal presentation" reveal the differences between presenters and evaluators and between the ideal and actual presentations?
2. Does an interaction scale differentiate session characteristics?
3. Does prior knowledge of the form of an evaluation instrument influence the performance of the presenter?

Data were collected using three evaluation instruments during eight presentations. The content of both institutes was parallel, permitting the matching of four sessions. An evaluation team consisting of State University of New York Albany staff and students (N=5) and a random sampling of participants (N=85) combined with the presenters (N=8) to generate the data at various stages. The participants (35 males, 50 females) ranged in age from 27 to 52 years old and had 3.7 median years experience as elementary teachers.

Ideal Presentation

On the form labeled "ideal presentation," presenters and observers from the evaluation team responded using a ten-point scale to describe a perfect session for a particular topic. Points considered were planning process, audience needs, material and information, and mode of presentation. This introspective analysis provided perceptions of how the sessions might be presented if free from constraints of time, personalities, abilities, etc. Such self-analysis from both leaders and observers are essential if process and product evaluations are viewed as interdependent (Halasa, 1977).

The "ideal presentation" format allowed quantification of elements which must be considered in the planning and execution of training sessions. Included in the planning segment were such features as the preparation of hand-outs, setting objectives, devising varied activities and the use of equipment. Those completing the form for audience segment considered identifying needs, generating feedback and interaction, as well as motivation.

The materials and information form included statements on encouraging commitment, addressing district needs, resources and experiences in addition to discussing strengths and weaknesses in the procedures presented. The final segment in the "ideal presentation" dealt with mode of delivery, emphasizing types and delivery of activities and tone of session.

The data from "ideal presentation" are summarized in Table 1 in the form of mean estimates (0 to 10 as possible scores) and differences between rater groups recorded before actual presentations. Included also are observations made by the authors during presentations using an identical form to reflect the actual, rather than ideal, perception. This data reveals that the ideal presentation was higher in the minds of the presenters than in the minds of the trained observers. Judgments of what actually transpired were equal to or higher than the observers' ideal estimates in three of

TABLE 1
 Mean Estimates (N=85) and Difference in Perceptions
 of Ideal Presentation vs. Actual Presentation

<u>Factors</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>Totals</u>
<u>Planning</u>							
Presenters (1)	9.0	8.4	8.4	6.5	6.3	7.8	7.7
Observers (2)	5.1	6.7	7.1	5.4	5.0	7.7	6.2
Actual (3)	5.5	7.3	7.5	3.7	3.8	7.8	6.2
Differences (1-2)	3.9	1.7	1.3	1.1	1.3	0.1	1.57
(1-3)	3.5	1.1	0.9	2.8	2.5	0.0	1.80
<u>Audience</u>							
Presenters (1)	9.0	7.7	8.1	7.7	7.7	6.3	7.7
Observers (2)	4.7	6.0	9.1	4.6	5.8	9.6	6.6
Actual (3)	5.4	7.5	7.5	5.6	6.7	5.8	6.4
Differences (1-2)	4.3	1.7	-1.0	2.1	1.8	-3.3	2.27
(1-3)	3.6	0.2	0.6	2.1	0.9	0.5	1.32
<u>Material</u>							
Presenters (1)	9.3	8.4	9.3	8.0	4.9	7.8	7.9
Observers (2)	4.2	5.8	6.1	5.7	5.8	7.8	5.9
Actual (3)	7.5	7.5	7.8	6.1	4.5	7.8	6.9
Differences (1-2)	5.1	2.6	3.2	2.3	-0.9	0.0	2.35
(1-3)	1.8	0.9	1.5	1.9	0.4	0.0	1.08
<u>Presentation Mode</u>							
Presenters (1)	6.5	5.9	9.4	6.5	5.0	5.7	6.5
Observers (2)	4.6	4.3	4.3	3.6	5.0	6.7	4.8
Actual (3)	6.7	5.5	7.5	5.1	6.7	5.0	6.1
Differences (1-2)	1.9	1.6	5.1	2.9	0.0	-1.7	2.20
(1-3)	-0.2	0.4	1.9	1.4	-1.7	0.7	1.05

the four areas. Also the differences between the observers and presenters were higher in three areas than the differences between the presenters and actual occurrences. Such data appear to indicate high presenter's ideal and different expectations between presenters and observers. Also, the presenters reported relatively equal and high ideal in three areas, with the lowest concern for the mode of presentation. Support for this observation is revealed in the judgments about what actually transpired since the mode again received the lowest rating while the material section received the highest.

The second part of this form further developed self-analysis by asking the presenters to complete the following open-ended questions after the session:

1. What major factors did you consider in preparing the session?
2. What major factors do you use to determine the success of the session?
3. What would you do differently if you were to repeat the session?

The answers to these questions highlighted some central concerns for future planning. The major factors considered in preparing sessions were reported in order of importance as follows:

1. topics - materials, objectives, value processes useful in district/classroom
2. audience - needs, level of sophistication, size
3. program - time of presentation and place in overall program

When asked to indicate success factors, presenters responded in four categories:

1. audience questions

2. audience reactions - attention, willingness to interact
3. subjective judgments - were goals met?
subjective analysis
4. post-session activity - application
and request for information

Presenters seldom had major doubts about the effectiveness of their sessions, but the changes they made could be grouped into five areas:

1. group hand-outs differently; make transparencies available
2. have more time or different time slot
3. narrow the topic
4. include more or different activities
5. be more dynamic

In sum, the subjective analysis produced by the open-ended questions revealed that presenters tended to consider the objectives, materials, audience needs and reactions as their major concerns. These are the factors most frequently identified for session planning, evaluation and for making improvements. It is interesting to note that a number of presenters relied not on self-judgments about the program's success but rather on audience reactions.

Using this information for future planning, organizers of training sessions might assist presenters by carefully delineating topics to be covered and by describing in general the audience and the program.

Of equal importance is the assignment of a narrow, manageable topic and the availability of hand-outs with varied activities during the session.

Interaction Analysis

The second area of investigation deals with the tone set during the session which is often determined by verbal and non-verbal interaction. The question posed was whether an interaction scale would differentiate the characteristics of the session. A review of available scales indicated that most were designed for classroom use; none was constructed specifically for inservice sessions (Simon, 1970). Furthermore, it was doubtful that available scales would reflect the learning environment of inservice training, given the unique characteristics of staff training programs.

Nonetheless, a modified version of the Flanders Interaction Scale (1965) was selected for use in this study. For purposes of clarity, the category of presenting information was sub-divided into the following categories: orally, orally with audio-visual equipment, written, and pictures and graphs. A twelfth category of humor or storytelling was also added. This modified scale was applied to four presentations of approximately 1¼ hours each by the evaluation team (N=5), trained in the use of the scale. Each member was directed to record a code number for the activity occurring every ten seconds, for five minutes, four times during each session. Thus, 35 codings were recorded four times for each session by each member of the team in overlapping time blocks.

The data generated by the evaluation team depicting the coded activities during the four sessions was converted to a percentage of time devoted to each category for each presentation (Table 2). This data demonstrates that the majority of time observed in these sessions was devoted to presenting material orally. One of the directions given to the presenters was that the workshop was to emphasize a "hand-on" and task approach rather than lecturing. The data gathered indicates little interaction with the exception of the fourth session which shows a high percent of time involving participants asking questions. The first session used the major portion of time for orally describing information while using transparencies.

TABLE 2

Percent of Time Devoted to Activity Categories
as Observed by Evaluation Team (N=5)

<u>Modes</u>	<u>Presentations</u>			
	1	2	3	4
Oral	19	74.5	55	54
Oral + A.V.	74	14	7	2
Questioning	1	2	5	3
Interaction	0	0	3	17
Part. Ques.	0	1	1	9
Stories/humor	0	4	21	0
Silence	2	.5	7	2
Other	4	4	1	13

In sum, interaction scales, such as this modified Flanders Scale, are not sensitive to the meaningful characteristics of staff development programs. Therefore, the need exists for a scale to be devised which would permit the coding of presenters' behavior, participants' behavior, verbal interactions, types of activities and the tone created during training programs for professionals.

Prior Knowledge of Form

Previous research has tended to support the notion that presenters design their sessions to match the objectives set by themselves. The question then arises, would presenters be influenced by an evaluation scheme if shown the evaluation form before making their presentation? In order to explore this question, a two-page evaluation form was constructed. Four

general questions were asked on the first page, and three sets of different questions were listed on different second pages dealing with the amount of information dispensed, the tone of the session, or audience reaction.

Eight presenters were matched, based on similar topics addressed during one of two institutes. Four were shown an evaluation form (pages one and two) two weeks before their scheduled presentation, and the other four received only general evaluation forms (page one). For the purposes of this study, it was hypothesized that the evaluations area shown to the presenter before his/her session would produce the highest scores from participants, under the assumption that the presenters would bias deliveries toward that favored area.

Table 3 displays the data gathered from a random sampling of 15 participants following the eight sessions. The second page of the evaluation form (rating one to five) represented one of the areas of information, tone or reaction, generating five responses for each of the three areas. The means and differences displayed in Table 3 resulted from a five-point scale summed over the four questions on the four evaluation topics. Presenters were matched based on their topics with the even numbered presenter in Table 3 having been shown an evaluation form prior to their session. The particular form shown to a presenter is indicated by an asterisk.

The results indicate all four presenters received higher ratings in the evaluation area in which they had prior knowledge. However, the average evaluation across all four areas was also higher, raising the question whether the highlighted topic reflects a conscious effort or an overall superior performance. The differences between the paired presenters were higher in the highlighted area in three of the four comparisons. This data seem to indicate the possibility that presenters are influenced by the evaluation form. The situation is analogous to that of the classroom teacher influenced in developing curriculum by the measurement instruments to be used.

TABLE 3
 Mean Ratings by Participants (N=15)
 of Four Presentations

<u>Forms</u>	<u>1</u>	<u>2</u>	<u>Differences</u>	<u>3</u>	<u>4</u>	<u>Differences</u>
general	2.32	2.25	-.07	2.28	2.02	-.26
information	2.16	3.42*	1.26	2.04	1.58	-.46
tone	2.17	1.75	-.42	2.06	3.13*	1.07
reaction	1.71	1.57	-.14	1.64	2.29	.65
TOTALS	2.09	2.25	-.16	2.00	2.26	.26
	<u>5</u>	<u>6</u>	<u>Differences</u>	<u>7</u>	<u>8</u>	<u>Differences</u>
general	2.13	2.65	.52	2.00	2.50	.50
information	1.71	3.75	2.04	1.83	2.50	.67
tone	1.33	3.83	2.50	2.00	3.00	1.00
reaction	1.52	2.28*	.76	1.78	3.85*	2.07
TOTALS	1.67	3.13	1.46	1.90	2.96	1.06

*Indicates the area seen by the presenter before session.

Conclusions

This study focused on the problems inherent in evaluation of staff development programs. Three evaluation instruments were used in two training institutes to generate information prior to the session (ideals), during the session (interaction scale), and following the session (questionnaire). It was hypothesized that use of these instruments would produce self-analysis and expectations from presenters, participants and an evaluation team of university staff.

The data collected address three questions previously described with the following results: (1) differences did exist between presenters' and observers' ideal concepts and what actually occurred, (2) a modified Flanders Scale was not sensitive to the learning environment of these training sessions, (3) the data tended to reveal that presenters were influenced by having prior knowledge of the evaluation format.

Within the limitations of numbers of subjects, personalities and topic differences, certain general conclusions appear warranted by this study. Research in the area of process and product evaluation is both necessary and possible. A structured evaluation scheme, based on the pre-established objectives of the session, will have a positive effect on the presenters and the participants. The use of observation, interview and questionnaire devices such as the ideal form and interactive scale are useful but require further development.

The demands of staff development programs continue to grow. If such programs are to continue and increase in efficiency, more research should be devoted to evaluating the processes and products involved in inservice training.

REFERENCES

- Fitzgerald, Thomas and Marino, Jackie. Evaluation of inservice training in reading. Educational Technology, in press.
- Flanders, N. Interaction analysis in the classroom-- A manual for observers. Michigan: University of Michigan, 1965.
- Halasa, Ofelia. The interdependency of product assessment in educational evaluation. Educational Technology, March, 1977, pp. 55-57.
- Moffitt, John. Inservice Education for Teachers. Center for Applied Research in Education, Inc., Washington, D.C., 1963.
- Simon, Anita (ed.). Mirrors for Behavior II. Research for Better Schools, Philadelphia, 1970.
- Strang, Ruth. Reactions to research on reading. The Educational Forum, 26, 1962, pp. 187-192.
- Wolf, Robert and Tymitz, Barbara. Ethnography and reading: matching inquiry mode to process (guest editorial). Reading Research Quarterly, 12, 1976, pp. 5-11.