DOCUMENT RESUME

ED 131 111 TM 005 836

AUTHOR Erickson, Glenn R.; Sheehan, Daniel S.
TITLE An Evaluation of a Teaching Improvement Process for University Faculty.
PUB DATE [Apr 76]
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (60th, San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-$0.83 HC-$1.67 Plus Postage.
DESCRIPTORS Classroom Observation Techniques; *College Instruction; *College Teachers; Effective Teaching; Feedback; Graduate Students; Higher Education; Inservice Teacher Education; *Instructional Improvement; Program Effectiveness; *Program Evaluation; Questionnaires; Student Attitudes; Teacher Attitudes; Teacher Evaluation; *Teacher Improvement; Teaching Skills; Video Tape Recordings

ABSTRACT

Forty faculty volunteers from 30 departments were randomly assigned to one of three experimental conditions: "full process," with teaching performance data collection, feedback, diagnosis, applied instructional improvement strategies and data re-collection; "diagnostic," which excluded the applied improvement strategies; and "data collection only." Full process and diagnostic condition faculty were assisted throughout by graduate student "teaching improvement specialists." Late semester findings were that full process instructors considered the process to be effective, worth their time and effort, and appropriate for their peers. There were no across treatment differences in faculty attitudes or self-ratings of teaching skill improvement, need for improvement, or overall teaching competence. Nor were there consistent across treatment differences in their students' attitudes or ratings of instructor teaching skill and improvement. (Author)

An Evaluation of a Teaching Improvement Process for University Faculty

Glenn R. Erickson

Instructional Development Program

University of Rhode Island

Daniel S. Sheehan

Research, Evaluation and Information Systems

Dallas Public Schools

Paper presented at the annual meeting of the

American Educational Research Association

San Francisco, April 1976

Abstract

Forty faculty volunteers from 30 departments were randomly assigned to one of three experimental conditions: "full process," with teaching performance data collection, feedback, diagnosis, applied instructional improvement strategies and data re-collection; "diagnostic," which excluded the applied improvement strategies; and "data collection only." Full process and diagnostic condition faculty were assisted throughout by graduate student "teaching improvement specialists." Late semester findings were that full process instructors considered the process to be effective, worth their time and effort, and appropriate for their peers. There were no across treatment differences in faculty attitudes or self-ratings of teaching skill improvement, need for improvement, or overall teaching competence. Nor were there consistent across treatment differences in their students' attitudes or ratings of instructor teaching skill and improvement.

AN EVALUATION OF A TEACHING IMPROVEMENT PROCESS FOR UNIVERSITY FACULTY

There are more than 250 college and university instructional improvement centers and programs in this country (Gaff, 1975). Their intellectual traditions, assessments of the most critical teaching and learning problems, and prescriptions vary, but all of those identified by Gaff are attempting to improve instruction through the in-service education of college teachers. Almost all offer at least some one-on-one consultation service to faculty members with course design or classroom teaching problems. Many offer consultation which is intended to help instructors improve their performance of underdeveloped skills or to broaden often limited teaching skill repertoires.

The Clinic to Improve University Teaching (Clinic)* has developed a systematic and individualized teaching improvement process which it believes has broad applicability to other instructional improvement programs. That process has several features which make it unique. First, its diagnostic and improvement procedures center upon some twenty teaching skills which are related to students' and instructors' perceptions of what constitutes effective teaching across a broad range of instructional situations. Second, the Clinic's diagnostic procedures include data collection techniques designed to obtain and integrate information from a variety of sources. These include student and instructor questionnaires, classroom observation and videotaping,

---

*The Clinic is supported by a grant from the W.K. Kellogg Foundation. Both authors were on the Clinic staff when the evaluation study described herein was conducted, and were ably assisted by B. Erickson, B. Howard, M. Jackson, L. Klemer-Hruska, D. McCarthy, and D. Wright.

interviews, and the examination of course materials. Third, improvement goals are negotiated and worked toward by selecting or designing and using any of a number of teaching improvement strategy alternatives. And, finally, faculty members are provided with continuing individualized assistance throughout the process by people thoroughly trained as "teaching improvement specialists." The process is designed to help an instructor identify significant strengths and problems within some course and to acquire or develop teaching skills which will increase his or her effectiveness within that course.

Unfortunately, evidence about the impact and the success or failure of the consultation strategy used by the Clinic has not been available. Nor have other teaching improvement programs (excepting, perhaps, instructional design programs) produced substantial evidence about their effectiveness. This is particularly unfortunate in a time when interest in "teacher effectiveness" and its improvement is high even while resources are shrinking. Such evidence is critical now for the many institutions considering the development, elimination or modification of their own instructional improvement programs.

We believe that a major reason for the scarcity of evaluative information has been a paralyzing preoccupation with what many consider to be the "ultimate" questions: Are teaching skills, behaviors and attitudes amenable to significant, quantifiable and lasting change? And, even if so, will students either learn more or more efficiently as a result of such changes? Important questions, yes, but not the only significant ones and probably not the best ones with which to begin. For programs directed toward improving university teaching to have any chance for success with reference to those "ultimate" questions, it would seem that they must be perceived as practical, effective and

attractive by consumers. Patricia Cross (1976, p. 3) is right: ". . . the reality is that instructional improvement is dependent upon faculty cooperation." This evaluation study, then, deals with faculty satisfaction with the Clinic's teaching improvement process, with student and faculty perceptions of qualitative changes in instructional performance, and with student and faculty attitudes toward themselves, their courses, and teaching.

Method

Subjects. The subjects were 40 volunteer faculty (28 males; 12 females) from 30 academic departments of a large New England university. Six were staff associates or instructors, 17 were assistant professors, 11 were associate professors, and six were full professors. Class sizes ranged from eight to over 200 students, with most falling into the 15 to 60 range. The subjects were probably not representative of the faculty of the university, but were similar in age, sex, rank, discipline and class size to faculty members who had volunteered to work with the Clinic in the past.

The teaching improvement process. The Clinic's improvement process is undertaken by instructors within the context of one of their courses over a semester. In collaboration with a graduate student trained as a "teaching improvement specialist," they work through four stages: 1) data collection; 2) feedback and diagnosis; 3) applied teaching improvement strategies; and 4) data re-collection and review. During the first stage, data regarding the instructor's teaching goals and objectives, procedures, and performance of certain teaching skills are collected through an interview, classroom observation and videotape, and a student and instructor teaching questionnaire. During the second stage, the instructor and improvement specialist review the data, jointly identify significant

strengths and problems, and establish reasonable improvement goals. Various teaching improvement strategies are devised and used during the third and longest stage. And, finally, additional data about perceived changes in the quality of teaching performance are collected.

Treatment conditions. Subjects were randomly assigned to one of three treatments: a "full process" condition which entailed going through the entire Clinic process as just described; a "diagnostic" condition which excluded the development and application of teaching improvement strategies (although possible improvement strategies were offered orally after diagnosis); and a "data collection only" condition which included only the first and the last stages of the teaching improvement process, with no feedback until the study was concluded. Subjects were then randomly assigned to five experienced teaching improvement specialists, all doctoral level students in Education.

Diagnostic instruments. The Teaching Analysis by Students (TABS) questionnaire was developed by the Clinic in 1974 for use as a diagnostic instrument within their teaching improvement process. It includes 38 teaching skill statements which are responded to on a four-point "amount-of-improvement-needed" scale. An additional item asks for a general rating of the course (TABS is reproduced in Bergquist, Phillips and Quehl, 1975). The TABS was administered to each instructor's class during the fourth week of the semester. At that time, each instructor also rated his or her own teaching performance on the TABS. These results were used to check on the early semester comparability of the three groups of subjects and as covariates for post measures.

Late semester measures. Near the end of the semester each instructor's students completed the three-section Student Questionnaire. Section I, different for each instructor, asked students to indicate on a five

point scale (much better to much worse) the extent of qualitative change they had observed over the preceding six weeks in their instructor's performance of the ten teaching skills which they had rated on the earlier TABS as most needing improvement. Section II included 35 teaching skill items with the scale from the TABS, and an item for an overall rating of the instructor. Section III was a three concept ("As a student in this course I feel. . .," "As a teacher, this instructor seems. . .," and "This course seems. . .") semantic-differential with seven-point bipolar adjective scales (see Figure 1).

At the same time, instructors completed the Instructor Questionnaire which, excepting the semantic-differential, was identical for each instructor to the Student Questionnaire completed by his or her students. The final section was also a semantic-differential, with similar adjective scales (see Figure 1) but different concepts ("In general, as a teacher I am. . .," "As a class, these students seem. . .," and "For me, this course is. . .").

Insert Figure 1 about here

Finally, after finishing the diagnostic stage of the process, full process and diagnostic condition subjects completed the 25 item Evaluation of Teaching Clinic: Part I. Full process condition subjects completed an additional 14 item Part II at the conclusion of the process. Both parts included a mix of forced-choice (mostly strongly agree to strongly disagree) and free-response items designed to solicit reactions to their improvement specialist and the elements of the improvement process.

Results

Student data. A comparison of class means across the three

treatment conditions revealed statistically significant group differences on the early semester TABS results. Consequently, subsequent analyses were restricted to data from students who completed both early and late semester measures so that analysis of covariance techniques could be used to adjust final class means for initial differences. (Comparisons of the scores of students who completed both instruments with those of students who completed only one or the other revealed no significant differences.)

Using class means as the unit of analysis, one-way analyses of covariance were carried out on the following late semester Student Questionnaire scores: 1) ratings of qualitative changes in the performance of the ten teaching skills which, according to TABS results, were most in need of improvement for each instructor; 2) ratings of the degree of improvement still needed on the 35 teaching skill items carried over from the TABS; and 3) overall instructor ratings. The covariates were, respectively: 1) the average amount-of-improvement-needed scores on the corresponding ten TABS items; 2) the average amount-of-improvement-needed scores on the corresponding 35 TABS items; and 3) the initial overall course rating scores from the TABS. The latter were also used as covariates for the multivariate analysis of covariance on the semantic-differential concept average scores.

The Student Questionnaire results are summarized in Table 1. There were no differences across treatment conditions in student perceptions of improvement on the ten weakest teaching skills for each instructor ($F=2.60$, $df=2/36$, $p<.10$) or for overall instructor ratings ($F=2.68$, $df=2/36$, $p<.10$). There were, however, significant differences ($F=3.30$, $df=2/36$, $p<.05$) in final amount-of-improvement-needed ratings for the 35 teaching skill items. (Post hoc comparisons with Scheffe's test

showed that full process condition adjusted mean scores were significantly higher than those of the diagnostic, but not data collection only, condition classes. There were no differences between diagnostic and data collection only condition scores.) With regard to attitude scores as measured by the semantic-differential, there were no differences

Insert Table 1 about here

across treatments in attitudes toward "self as student" or the course. But there were significant differences in attitudes toward the "instructor as teacher." (Scheffe's post hoc test indicated that full process condition adjusted mean scores were significantly greater than those of data collection only, but not diagnostic, condition classes. There were no other differences.)

Instructor data. As there were no significant across treatment condition differences on early semester instructor self-assessments, Instructor Questionnaire results were compared through analysis of variance. The results of these analyses are summarized in Table 2. They indicate that there were no differences across treatment conditions on late semester instructor self-assessments or in attitudes toward themselves as teachers, their courses, or their students.

Insert Table 2 about here

Finally, instructors' responses to the Evaluation of Teaching Clinic questionnaire are summarized in Table 3. For all Likert scale items, an average of 93.6% of the responses were positive. Of particular interest, perhaps, are the strongly positive responses to the items having

to do with the utility of improvement strategies for helping to improve their teaching (#31), whether the entire process was useful and worth their time and effort (#36), and whether they would recommend the process to their peers (#37).

Insert Table 3 about here

Discussion

This evaluation study did yield one unambiguous finding: faculty members who undertook the Clinic's entire teaching improvement process were pleased with the experience. They agreed that it helped them to improve their teaching, that it was worth their time and effort, that they would recommend it to their peers, and that their graduate student teaching improvement specialists were competent and helpful. It is true that all participants were volunteers and that Gaff (1975, p. 168) has noted that it has been the experience of several programs that ". . . participants generally find what they come for." Nonetheless, client satisfaction is an important goal for the Clinic and for other teaching improvement programs. Furthermore, some may be surprised that graduate students can be a credible medium for providing faculty members with what they obviously consider to be a very useful service.

Less clear is the significance of the lack of across treatment differences in late semester instructor self-ratings of improvement demonstrated (all felt that there had been improvements) or needed and of overall teaching competence. Nor were there differences in late semester attitudes toward themselves, their courses, or their students. Perhaps those results suggest that each of the faculty members who volunteered was indicating a willingness to change, and all may have

been encouraged to improve by the knowledge that their teaching was being examined for evidence of change. That explanation suggests that the improvement strategies used may well have worked as reported on item #31 of the Evaluation of Teaching Clinic. But they may not have resulted in changes which were experienced as any more intense than were the adjustments made by the atypically motivated subjects in the other experimental groups. As for instructor attitudes, there is simply no evidence that the Clinic's process had any impact.

Student responses to the Student Questionnaire are not very illuminating either. Generally, there are no across treatment condition differences, except that full process adjusted mean scores are more positive than those of one of the two, but not both, other groups for two variables: average amount of improvement still needed on all of the skills, and attitudes toward their instructors as teachers. The significance of those results may have been clearer if the full process scores had been more positive than those of both other groups, or of the same group, for both measures. As it stands, there is no real evidence that student attitudes or perceptions of their instructors' teaching was affected by the treatment condition to which their instructors were assigned.

In sum, there is no compelling evidence from either instructor self-ratings or student ratings that their teaching changed any more or less than that of instructors in the other two treatment groups. In fact, instructors and students perceived positive changes in teaching performance in all three groups. What is clear about the Clinic's teaching improvement process from this evaluation study is that it produces consumer satisfaction with the process, the mediator, and the results.

Some of the other results may have been less ambiguous if the evaluation design had been more precise. For example, we did not ask teaching improvement specialists to restrict their instructional improvement prescriptions and efforts to any abbreviated list of skills. To do so may have meant that subjects/clients would have spent time working on less personally significant teaching problems. But to have limited their attention to five or ten teaching skills, and/or to have assigned subjects to groups after problem identification, may have allowed for more precise assessments of the consequences of working on or ignoring shared problems.

This evaluation study, then, put aside for the moment any concern with the questions of whether teaching skills and behaviors are amenable to significant change and whether any such change would significantly affect student learning. It focused instead upon consumer satisfaction, perceptions of change, and some related attitudes. That faculty members were satisfied with the Clinic's teaching improvement process is obvious and encouraging. But the study failed to demonstrate that instructors who complete the Clinic's process will appear to themselves or their students to have improved any more than instructors in diagnostic or data collection only groups. Nor did we find any clear impact upon student or instructor attitudes toward themselves, their classes, or each other. We believe, however, that there is just enough ambiguity in the data to suggest that these findings may be more of a reflection upon the imprecision of the study's design and measures than of the ineffectuality of the Clinic's teaching improvement process. Certainly we hope that the Clinic and other instructional improvement programs will continue and step up the pace of their efforts to assess the impacts of their programs--even if most of us still lack the skills or the resources for investigating program effects upon student learning.

References

Bergquist, W. H., Phillips, S. R., and Quehl, G. (General Editor). A handbook for faculty development. Washington, D.C.: Council for the Advancement of Small Colleges, 1975.

Cross, K. P. The instructional revolution. Pap[illegible] the National Conference on Higher Education, Chicago, March [illegible].

Gaff, J. G. Toward faculty renewal. San Francisco: Jossey-Bass Publishers, 1975.

Concepts	Adjective Pairs
	self-confident/not self-confident
As a student in this course I feel...	irresponsible/responsible
As a teacher, this instructor seems...	unsuccessful/successful
	organized/disorganized
	passive/active
In general, as a teacher I am...	not anxious/anxious
As a class, these students seem...	uninformed/informed
	competent/incompetent
	tense/not tense
	enthusiastic/not enthusiastic (about learning)
	rewarding/unrewarding
	stimulating/dull
This course seems...	relevant/irrelevant
For me, this course is...	unpleasant/pleasant
	impersonal/personal
	meaningful/meaningless

Figure 1: Semantic-differential concepts and bipolar adjective pairs used in the third sections of the Student Questionnaire and the Instructor Questionnaire.

Table 1

Late Semester Student Questionnaire Results

Measures		Treatment Condition: Full Process (N=13)[b]	Diagnostic (N=13)	Data Only (N=14)	p[a]
Perceived improvement on ten weakest teaching skills[c]	Adj. $\bar{X}$	2.56[d]	2.58	2.69	.10
	Obs. $\bar{X}$	[illegible]	2.61	2.70	
	SD	.22	.16	.15	
Average amount of improvement still needed on 35 teaching skill items[c]	Adj. $\bar{X}$	(1.89)	(2.07)[f]	2.01	.05
	Obs. $\bar{X}$	1.80	2.18	1.99	
	SD	.31	.19	.26	
Overall instructor ratings[g]	Adj. $\bar{X}$	1.95[d]	2.26	2.32	.10
	Obs. $\bar{X}$	1.85	2.51	2.18	
	SD	.61	.49	.65	
Semantic-differential concept scores[g]					
Toward self as student	Adj. $\bar{X}$	5.23[h]	5.00	5.09	.26
	Obs. $\bar{X}$	5.28	4.90	5.14	
	SD	.42	.39	.32	
Toward instructor as teacher	Adj. $\bar{X}$	(6.06)[h]	5.85	(5.64)	.005
	Obs. $\bar{X}$	6.11	5.74	5.70	
	SD	.33	.31	.39	
Toward course	Adj. $\bar{X}$	5.53[h]	5.31	5.14	.10
	Obs. $\bar{X}$	5.67	5.01	5.29	
	SD	.59	.63	.78	

[a]Probability of differences across treatments being by chance

[b]Number of classes

[c]Covariates were earlier scores on corresponding TABS items

[d]Scores could range from one (most positive) to five

[e]Scores could range from one (most positive) to four

[f]Adjusted means within parentheses and in the same row differ significantly according to Scheffe's post hoc test

[g]Covariates were overall course ratings on early semester TABS

[h]Scores could range from one (most negative) to seven

Table 2

Late Semester Instructor Questionnaire Results

Measures		Treatment Condition: Full Process	Diagnostic	Control	p[a]
Perceived improvement on ten weakest teaching skills	$\bar{X}$	2.32[b]	2.40	2.68	<.10
	SD	.42	.35	.27	
	N[c]	12	10	11	
Average amount of improvement still needed on 35 teaching skill items		97[d]	2.15	2.03	<.53
		.43	.31	.37	
	N	12	10	11	
Overall instructor rating	$\bar{X}$	2.18[b]	2.44	2.30	.60
	SD	.58	.50	.64	
	N	11	9	10	
Semantic-differential concept scores					
Toward self as teacher	$\bar{X}$	5.77[e]	5.72	5.65	.89
	SD	.60	.58	.57	
	N	12	13		
Toward course	$\bar{X}$	6.42[e]	5.86	5	.16
	SD	.64	.72	1	
	N	12	13		
Toward students	$\bar{X}$	5.19[e]	4.45	4.	.12
	SD	.79	.77	1.13	
	N	12	13	10	

[a]Probability of differences across treatments being by chance

[b]Scores could range from one (most positive) to five

[c]Number of instructors

[d]Scores could range from one (most positive) to four

[e]Scores could range from one (most negative) to seven

Table 3

Summary of Instructor Responses to
Evaluation of Teaching Clinic Questionnaire Items

Items[b]	Responses		
	Positive[a]	Negative	Undecided
Part I: completed by full process and diagnostic condition subjects after diagnosis			
1. By the end of the initial interview, the Clinic's goals were quite clear to me.	26	0	0
2. The procedures used within the Clinic's teaching improvement process were clearly explained to me.	25	0	1
3. Procedures and instruments for collecting data about my teaching were not made clear to me.	24	2	0
4. I do not believe that the preliminary information about my course and my teaching methods that was gathered through the Course Information Form, the initial interview, and the course materials which I made available gave the teaching improvement specialist (TIS) an adequate perspective from which to assess data to be collected later.	23	2	1
5. During the initial interview, I was able to clarify my personal objectives in working with the Clinic.	23	2	1
6. This initial stage did not give me a very clear idea of how much of my time and my class's time would be required by the Clinic's data collection procedures.	23	3	0
7. The TIS failed to answer some of my questions satisfactorily.	25	1	0
8. All in all, I am convinced that the initial interview was well worth my time and effort.	25	1	0

[a]Strongly agree or agree on positively stated items; strongly disagree or disagree for negatively stated items. Rule for negative responses is just the reverse.

[b]Items 24, 25, 32 and 39 have been omitted because of space limitations. Two of them (24 and 32) asked for ratings of the usefulness of the various data sources, and the other two asked for additional comments or elaborations.

Table 3 (continued)

Summary of Instructor Responses to
Evaluation of Teaching Clinic Questionnaire Items

Items	Responses		
	Positive	Negative	Undecided
9. The potential value of each type of data which was collected (observation, videotape, student TABS, self-assessment, and prediction of student responses to TABS) was persuasively explained to me.	23	1	2
10. The TIS's observation and videotaping of my class was disruptive or otherwise poorly handled.	24	1	0
11. I was not happy with the way the administration of the TABS questionnaire to my class was handled.	25	1	0
12. The procedures for completing the TABS' self-assessment and prediction of student responses were explained clearly.	24	2	0
13. Generally, I do not believe that the data which was collected was sufficiently comprehensive to provide the basis for making decisions about my significant teaching strengths and weaknesses.	21	4	1
14. I did not find the opportunity to preview the videotape and TABS results by myself useful.	25	1	0
15. The computer printout of the TABS results was confusing or otherwise not an adequate display of the data.	24	1	1
16. The TIS appeared to be well prepared for our localization session.	26	0	0
17. The TIS did a fine job of presenting and explaining the data which was collected about my teaching performance.	25	0	1
18. Even with the TIS's assistance, I was overwhelmed by the amount and/or the variety of data which was made available to me.	23	2	1
19. We were able to identify some significant teaching strengths which were indicated by the data.	25	0	1

Table 3 (continued)

Summary of Instructor Responses to
Evaluation of Teaching Clinic Questionnaire Items

Items	Responses		
	Positive	Negative	Undecided
20. The TIS and I did not agree on significant teaching problems which were substantiated by the available data.	26	0	0
21. The TIS was not able to suggest practical and potentially worthwhile teaching improvement strategies for me to undertake.	22	0	2
22. I was not convinced that I should go on to the teaching improvement strategy phase of the instructional improvement process.	23	1	1
23. So far, I am convinced that my TIS is competent to work with me.	26	0	0
Part II: completed only by full process condition subjects after entire process completed.			
26. I was impressed by the TIS's ability to suggest a wide variety of possible teaching improvement strategies.	11	1	0
27. The TIS was not adequately sensitive to my personal style, responsibilities, and specific instructional problems when suggesting possible teaching improvement strategies.	11	1	0
28. The teaching improvement strategies which I tried, or which were suggested to me, were not explained to my satisfaction.	12	0	0
29. The TIS was not of much help to me as I tried to implement teaching improvement strategies.	12	0	0
30. Adequate procedures for assessing the effectiveness of the teaching improvement strategies which I tried were not developed or used.	11	1	0
31. Overall, I believe that the teaching improvement strategies which I tried helped me improve my teaching.	12	0	0

Table 3 (continued)

Summary of Instructor Responses to
Evaluation of Teaching Clinic Questionnaire Items

Items	Responses		
	Positive	Negative	Undecided
33. I found it difficult to interpret the final questionnaire data as it was displayed and explained.	9	1	?
34. I thought that efforts to collect data about changes in my teaching performance were comprehensive and satisfactory.	10	2	0
35. After the final data re-collection process, the TIS helped me identify current teaching strengths and any significant teaching problems.	12	0	0
36. Overall, I found the Clinic's teaching improvement process useful and worth my time and effort.	11	0	1
37. I would not recommend the Clinic's teaching improvement process to my peers.	11	1	0
38. I found my TIS competent and helpful.	12	0	0