DOCUMENT RESUME

ED 270 408 SP 027 621

AUTHOR Menges, Robert J.; Brinko, Kathleen T.
TITLE Effects of Student Evaluation Feedback: A Meta-Analysis of Higher Education Research.
PUB DATE Apr 86
NOTE 19p.; Paper presented at the Annual Meeting of the American Educational Research Association (70th, San Francisco, CA, April 16-20, 1986).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Research; *Feedback; Higher Education; Meta Analysis; *Student Evaluation of Teacher Performance; Teacher Effectiveness; *Teacher Evaluation; *Teacher Improvement

ABSTRACT

On many campuses, student ratings of courses and teachers are routinely given to instructors as feedback about their teaching. This review applies meta-analytic techniques to 30 studies of the effectiveness of student ratings feedback. The studies met the following criteria: (1) they used student ratings as the primary source of feedback; (2) they investigated post-secondary instruction; (3) they were conducted in the classroom rather than laboratory settings; (4) they employed a control group for comparison purposes; and (5) they stood apart from larger training programs in which the effects of feedback are inseparable from the effects of training. Results indicated that feedback from student ratings alone produces a positive but small effect on subsequent ratings. When ratings were accompanied by consultation and/or other types of feedback, considerably larger positive effects were likely. In the few studies of feedback's effects on student achievement and affect, results were less clear. Future research should: (1) investigate additional dependent variables; (2) more carefully document and investigate feedback implementations; and (3) explore additional characteristics of the recipients of feedback. The studies analyzed in this review are attatched, as well as a listing of those studies which did not meet criteria for meta-analysis (JD)

Effects of Student Evaluation Feedback: A Meta-Analysis of Higher Education Research

Robert J. Menges and Kathleen T. Brinko
Northwestern University

The Center for the Teaching Professions
Northwestern University
2003 Sheridan Road
Evanston, IL 60201

Paper presented at the meeting of the
American Educational Research Association,
San Francisco, April, 1986.

Effects of Student Evaluation Feedback: A Meta-Analysis of Higher Education Research

The institutionalization of course and teacher evaluations on many campuses in recent years has stimulated controlled research on the effects of feedback to teachers. In this paper, we examine how feeding back information about teaching affects the subsequent teaching of instructors in postsecondary education. Such information may come from many sources, but this review is limited to studies in which information for feedback comes from student evaluations. We report a meta-analysis of 30 studies, most of which take the following form: Results of student evaluations, usually collected at midterm, are fed back to some of the participating teachers, and end-of-term student evaluations are examined to identify differences between teachers who did and did not receive feedback.

Method

Searching the Literature

Computer-assisted searches were conducted on the ERIC (Education Resources Information Center), DAI (Dissertation Abstracts International), PsycInfo (Psychological Abstracts), and MEDLINE (Index Medicus) databases. The Business Publications Index and Abstracts and the Business Periodicals Index were searched manually. Bibliographies of the items identified in these searches were scanned in order to locate additional pertinent references.

Over 300 books, journal articles, dissertations, and unpublished reports and papers were reviewed for consideration and examination. Seventy-one empirical studies were identified which evaluated the effectiveness of some form of feedback to postsecondary instructors for the purpose of improving their teaching (Menges, Brinko, & L'Hommedieu, 1986). In 52 of the 71 studies, feedback was in the form of student ratings.

Criteria for Inclusion in the Meta-Analysis

The 30 studies finally included in the meta-analysis are studies 1) which investigated post-secondary instruction, 2) which used student ratings as the primary source of feedback, 3) which were conducted in classroom rather than laboratory settings, 4) which employed a control group for comparison purposes, and 5) which stood apart from larger training programs in which the effects of feedback are inseparable from the effects of training. These studies are listed in Attachment A. Studies of student ratings feedback which do not meet these criteria are listed in Attachment B.

Procedures for Coding and Analyzing Data

Information about these studies was coded into 46 categories which describe each study along five dimensions: document characteristics, participant characteristics, treatment characteristics, design and analysis characteristics, and variable characteristics.

When possible, effect sizes were calculated using Glass's (Glass, McGaw, & Smith, 1981, p. 102) formula (nine studies):

$$\hat{\Delta} = (\overline{X}_E - \overline{X}_C) / s_x.$$

For studies where that formula could not be used, i. e., when F or t statistics rather than means and standard deviations were reported, this formula (Glass, McGaw, Smith, 1981, p.107) was used (nine studies):

$$\Delta_p = t(1/n_E + 1/n_C)^{1/2}.$$

Where it was not possible to use either formula, e. g., when only chi squares were reported, we made conservative estimates of effect size based on level of significance and sample size (one study). In instances where an effect size could neither be calculated nor estimated but where reported results clearly indicated only small and random differences between groups, an effect size of zero was assigned (nine studies). Where effect sizes could neither be calculated nor estimated but where reported results indicated a positive significant difference between groups (two studies), we substituted an effect size equal to the mean of the effect sizes of those studies which

reported positive significant differences that were calculable or estimable.

Many of the studies contained multiple comparisons; for example, several studies compared treatment groups on each item of a questionnaire. In these cases we chose to average effect sizes of comparisons within studies because averaging yielded a more straightforward and interpretable result. Averaging also avoids problems of unfairly weighting the results of multiple comparison studies and of capitalizing on sampling error. The effects of theoretically relevant factors and subcomparisons are discussed elsewhere (L'Hommedieu, Brinko, & Menges, 1986).

Results

Of the 30 studies in the meta-analysis, 10 found significant differences between groups, and all comparisons favored the feedback group. One study found mixed results. The remaining 19 studies found no significant differences between groups. Our findings are summarized in Table 1.

Effects on Subsequent Student Ratings

All implementations versus no feedback

Twenty-seven studies with 31 comparisons compared instructors who received some form of feedback from student ratings with instructors who received no feedback. The average effect size was .44 with a standard deviation of .64 ($p < .001$). Thus, on the average, student ratings feedback raised subsequent ratings by almost one-half of a standard deviation. This indicates that at the end of the experimental treatment, ratings of the average teacher in the experimental groups were higher than 67 percent of the teachers in the control groups.

Summarizing across all studies, however, obscures important distinctions. These 27 studies include three distinct implementations of feedback: 1) student ratings feedback, e. g., statistical summaries sometimes accompanied by interpretative texts and/or written suggestions for improvement; 2) student ratings feedback with consultation, i. e., student ratings feedback discussed with a consultant for one to two hours,

sometimes accompanied by verbal suggestions for improvement; and 3) augmented student ratings feedback, i. e., student ratings feedback with consultation accompanied by feedback from other sources, such as self-evaluation, peer evaluation, peer group discussions, videotape analysis, and so on. Below we report effect sizes separately for these three types of student ratings feedback implementations.

Student ratings alone versus no feedback. (23 studies reporting 23 comparisons.) Sixteen of these studies, or 70 percent, found no significant differences between feedback and no feedback groups. Student ratings feedback alone produced a very small effect (M_Δ = .22; SD = .32; $p < .01$). Feedback of this type raised subsequent ratings by one-fifth of a standard deviation unit; in other words, after receiving student ratings feedback, the average instructor was rated higher than 59 percent of control group teachers. Thus, it appears that systematic feedback from student ratings alone has a positive but small effect.

Student ratings with consultation versus no feedback. (Five studies reporting five comparisons.) Student ratings feedback with consultation produced greatly varied results. In these studies, effect sizes ranged from 0 to 2.50 (M_Δ = 1.10; SD = 1.14; n.s.). However, four of the five studies reported significant differences, favoring the group which received student ratings feedback with consultation. On the average, feedback of this type raised subsequent ratings by more than one standard deviation unit. Thus, the average instructor who received student ratings feedback with consultation was subsequently rated higher than 86 percent of teachers in the control groups. We conclude that systematic student ratings feedback accompanied by interaction with a consultant can have large positive effects on subsequent performance; however, variables critical to effective consultation are yet to be identified.

Augmented student ratings versus no feedback. (Three studies reporting

three comparisons.) Student ratings feedback with consultation augmented by feedback from other sources produced a mean effect size of .996 (SD = 55; n.s.). Although there are only three such studies, each showed positive results and two reached statistical significance. Augmented feedback raised instructor performance by one standard deviation, and ratings of these instructors exceeded the ratings of 84 percent of control group instructors. When student ratings feedback is accompanied by other types and sources of feedback, the process can result in large positive effects on instructor performance.

Augmented student ratings versus student ratings alone

Two studies reporting three comparisons (not shown in Table 1) investigated the effects of augmentation on subsequent student ratings. Both studies reported no significant differences between those receiving student ratings feedback alone and those receiving student ratings feedback with consultation augmented with other types of feedback. Because we were unable to calculate or estimate actual effect sizes in both studies, all three comparisons were assigned effect sizes of 0. It is interesting to note that videotape recording was the other source of feedback in each comparison. It may be that videotape feedback and student ratings feedback are interactive and the effects of one adds nothing to the effects of the other.

Effects on Student Achievement

All implementations versus no feedback

Three studies reporting four comparisons compared the effects of student ratings feedback on student achievement. These studies yielded greatly varied results (M_Δ = .25, SD = .61; n.s.). This diversity could be due to the nature of the instruments used (one was standardized and two were locally constructed); however, comparison of results as measured by standardized and nonstandardized instruments yields no conclusive trends. It is also possible that the variation in results is due to differences in teacher

expectations, e. g., teaching to the test, or to the implementation of the feedback process. Clearly, more research is needed to determine the effect of student ratings feedback on student achievement. Nevertheless, we report comparisons separately for each type of feedback implemented.

Student ratings alone versus no feedback. (Two studies reporting two comparisons.) The results of the two studies which investigated the effects of student ratings feedback alone on student achievement found greatly discrepant results. One study mean was -.53 and the other was .94. Thus the effect size found (M_Δ = .20, SD = 1.0; n.s.) is inconclusive. The study which reported the mean of -.53 used a standardized measure of achievement, and the other study which reported the mean of .94 used a nonstandardized measure.

Student ratings with consultation versus no feedback. (Two studies reporting two comparisons.) Less discrepancy was found in the two studies which compared the achievement of students of instructors who received student ratings feedback with consultation and the achievement of students of instructors who received no feedback (M_Δ = .30, SD = .15; n.s.). One of the studies used a standardized measure and the other used a nonstandardized measure. Although the statistic did not reach significance, differences in both studies favored the feedback groups.

Augmented student ratings versus student ratings alone

The only study which compared augmented student ratings feedback with student ratings feedback alone (not shown in Table 1) reported no significant differences between groups on student achievement. The effect size calculated for this comparison is .18, a very small but positive effect.

Effects on Student Affect

All implementations versus no feedback

Five comparisons from three different studies were located which compared the effects of some form of student ratings feedback with a no

feedback control group. Two of the three studies used locally constructed instruments which measured attitude toward the subject and attitude toward self. Both of these studies reported significant differences between groups. The third study used a measure similar in content but standardized. This study reported no significant differences between groups. The mean effect size was .40 (SD = .25; $p < .05$). Thus the average class whose instructor received some type of student ratings feedback scored higher on measures of affect than 65 percent of the no feedback control group.

However, when effects are analyzed for the three types of student ratings feedback, each fails to reach significance. This is not surprizing given the small number of comparisons for each type of student ratings feedback. More research is needed in this area in order to determine the effects of student ratings feedback on student affect.

Comparison with Earlier Reviews

Three previous reviews on this topic have appeared. In their qualitative review, Rotem and Glasman (1979) examined 13 empirical studies, six of which investigated feedback to postsecondary instructors. They concluded that "feedback from student ratings (as was elicited and presented to teachers in the studies reviewed) does not seem to be effective for the purpose of improving performance of university teachers" and suggested that "educational consulting services may be required as an integral part of evaluation aimed at improving teaching" (p. 507).

The present review expands and updates Cohen's (1980) meta-analysis. Cohen analyzed 17 studies, all but one of which are also included in the present review. He found a mean effect size of .38 across all studies, a mean effect size of .20 for studies of student ratings alone, and a mean effect size of .64 for studies of student ratings feedback with consultation and/or other types of feedback. Cohen concluded that "comparatively large effect sizes emerged from studies using augmented feedback or consultation in conjunction with student ratings feedback. Studies using only student rating

feedback produced much smaller effects. These results clearly suggest that instructors need more than just student ratings feedback to markedly improve their instruction" (p. 338).

In a discursive review of the literature on improving postsecondary instruction, Levinson-Rose and Menges (1981) noted 24 studies investigating the effects of student ratings feedback. They concluded that "feedback from students can positively affect subsequent teaching, particularly if ratings are accompanied by consultation" (p.419). They also rated each study according to its design features as warranting low, moderate, or high confidence and observed that "the greater our confidence in the study, the less likely it supports the intervention" (p.417).

Compared with these previous reviews, the larger pool of studies now available permits differentiation of three rather than two treatment implementations. Unfortunately the number of studies was quite uneven across implementations. Nevertheless, effects from student ratings feedback alone are modest. When accompanied by face-to-face consultation, however, the effects are more than quadrupled. And when accompanied by face-to-face consultation plus other types of feedback, the effects are similarly strong.

These results are counterintuitive in that we expected student ratings, consultation, and feedback from other sources to have additive effects. Instead, student ratings feedback with consultation and augmented student ratings feedback yield approximately the same effect. We offer three explanations for this finding. First, it may be that student ratings feedback and feedback from other sources are interactive, that is, the effects of one may cancel out the effects of the other. Second, the true gains made by instructors in their efforts to improve their instruction may be masked by a ceiling effect. Since most student ratings instruments utilize a 5-point Likert scale, there is little room to report a wide range of improvement for the already average or above average instructor. Third, the similarity in

results may be due to differences in types of studies. Whereas all studies of student ratings feedback with consultation were empirical studies of the feedback process per se, all studies of augmented feedback were evaluation studies of faculty development programs. This difference in focus may have produced different expectations in experimenters, in instructors, or in the students, which in turn produced differential results.

Since the present review is the most recent and the largest quantitative synthesis, it enables us to ask whether more recent studies differ from earlier studies in their findings and in their design. We divided the studies into two groups: those reported prior to 1980, the year Cohen's meta-analysis appeared, and those reported during or since 1980. With regard to results, we found a marked difference: The mean effect size of studies reported since 1980 (seven studies reporting nine comparisons; M_{Δ} = .83) is three times as large as the effect size of the pre-1980 group (20 studies reporting 22 comparisons; M_{Δ} = .28). However, the variance of the later group (SD = .91) is twice as great as the pre-1980 group (SD =.41).

With regard to design, studies appearing in 1980 or later are no more likely than earlier studies to employ feedback from sources other than student ratings (what we have termed augmented feedback) or to define impact beyond student ratings, e. g., using measures of student affect or achievement.

What Major Questions Remain?

We have found that across studies, the effects of feedback from student evaluations are evident in subsequent student ratings, especially when feedback from students is augmented with consultation or with consultation plus feedback from other sources. These findings are quite clear, and we contend that additional studies like those of recent years would be redundant. Instead, studies should be refined to deal with three problem areas apparent in this literature.

1. The first area concerns **dependent variables**. The dependent variables of student achievement and student affect should have additional research attention, given the inconclusive findings so far. Other indicators of the impact of feedback should also be studied, including comments by students on open-ended evaluations or in interviews, changes in faculty and student attitudes and values, modifications of course materials, and so on.

A more fundamental problem stems from the nature of student evaluation instruments. Most are based on unarticulated theories of teaching effectiveness, a problem which Abrami (1985) among others has called to our attention. While such theory is being developed, researchers can take the interim step of closely matching dependent measures to the content of feedback. For example, if feedback includes information on classroom interactions, then student evaluation items should elicit data on that topic. Few investigators appear to have designed dependent measures in this way, thus reducing their sensitivity.

Finally, impact of feedback over time with repeated interventions should be investigated more carefully.

2. The second problem area is the nature of feedback **implementations**. In most of these reports, feedback as a treatment is described only sketchily. Few details are given about how information is communicated. Investigators seldom verify even that feedback has been received, leading us to wonder how much of the effects of consultation occur merely because consultation insures that feedback is actively attended to and processed.

What goes on during consultation is also inadequately reported. We wonder which content is emphasized and how much emphasis is given to interpretation, to diagnosis, and to suggested correctives. We also wonder whether it makes a difference if faculty members themselves control decisions, i. e., if the teacher decides whether or not to receive feedback, what the content of that feedback should be, and what assistance is needed to interpret information and to plan for change.

Finally, other treatments should be investigated. There might be assessments of such sources of feedback as videotapes, peer observers, interaction analysis, and so on, relative to student ratings, with and without consultation. Such treatments might also be tried in combination, given adequate theoretical justification and appropriately sensitive measures.

3. Characteristics of **the recipient** of feedback comprise the third problem area. Some faculty are surely more ready to use feedback than others. Several studies suggest that one indicator of readiness is discrepancy between self-evaluaton and student evaluation, at least if that discrepancy is moderate. Other variables deserving further study include objective characteristics such as sex, years of experience, tenure status (as a possible indicator of professional vulnerability).

Individual differences more closely related to pedagogy include professors' own cognitive styles and learning styles; their definitions of the teaching role, e. g., the relative importance given to content coverage versus student outcomes; and the priorities they assign to various professional responsibilities, e. g., to teaching, scholarship, and service.

Psychological variables which may yield interesting results include efficacy (Bandura, 1977), self-monitoring (Synder, 1985), and causality orientation (Deci and Ryan, 1985). Researchers informed by the "theory of reasoned action" (Ajzen, 1985) might investigate 1) the teachers' own attitudes toward the changed behavior and 2) teachers' beliefs that others who are important to them think they should change their behavior.

Conclusion

It is apparent that feedback from students, when augmented with consultation or other types of feedback, can powerfully influence subsequent teaching. But much remains to be learned about the extent or the impact of feedback, about the details and dynamics of the feedback process, and about the characteristics of those most receptive to feedback.

References

Abrami, P. C. (1985). Dimensions of effective college instruction. Review of Higher Education, 8, 211-238.

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beekmann (Eds.), Action control: From cognition to behavior (pp. 11-39). New York: Springer-Verlag.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. Psychological Review, 84, 191-215.

Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. Research in Higher Education, 13, 321-341.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. Journal of Research in Personality, 19, 109-134.

Glass, G. V., McGaw, B., & Smith, M. L. (1981). Meta-analysis of social research. Beverly Hills, CA: Sage Publications.

Levinson-Rose, J., & Menges, R. J. (1981). Improving college teaching: A critical review of research. Review of Educational Research, 51, 403-434.

L'Hommedieu, R., Brinko, K. T., & Menges, R. J. (1986). Effects of student ratings feedback to college teachers: A meta-analysis. Evanston, IL: Center for the Teaching Professions, Northwestern University.

Menges, R. J., Brinko, K. T., & L'Hommedieu, R. (1986). Improving college teaching through feedback to teachers. Evanston, IL: Center for the Teaching Professions, Northwestern University.

Rotem, A., & Glasman, N. S. (1979). On the effectiveness of students' evaluative feedback to university instructors. Review of Educational Research, 49, 497-511.

Snyder, M. (1979). Self-monitoring processes. In L. Berkowitz (Ed.), Advances in experimental social psychology (pp. 85-128). New York: Academic Press.

Table 1: Meta-Analysis of Feedback Versus No Feedback Studies

FEEDBACK IMPLEMENTATION

DEPENDENT MEASURE	All implementations	Ratings alone	Ratings with consultation	Ratings with consultation augmented by other feedback
Subsequent Ratings	31 comparisons	23 comparisons	5 comparisons	3 comparisons
	$\underline{M}_\Delta$ = .44 $\underline{SD}$ = .64 $\underline{t}$ = 3.78***	$\underline{M}_\Delta$ = .22 $\underline{SD}$ = .32 $\underline{t}$ = 3.33**	$\underline{M}_\Delta$ = 1.10 $\underline{SD}$ = 1.14 $\underline{t}$ = 1.927	$\underline{M}_\Delta$ = .996 $\underline{SD}$ = .55 $\underline{t}$ = 2.56
	67th %ile	59th %ile	86th %ile	84th %ile
Achievement	4 comparisons	2 comparisons	2 comparisons	No comparisons located
	$\underline{M}_\Delta$ = .25 $\underline{SD}$ = .61 $\underline{t}$ = .71	$\underline{M}_\Delta$ = .20 $\underline{SD}$ = 1.04 $\underline{t}$ = .20	$\underline{M}_\Delta$ = .30 $\underline{SD}$ = .15 $\underline{t}$ = 2.00	
	60th %ile	58th %ile	62nd %ile	
Affect	5 comparisons	2 comparisons	2 comparisons	1 comparison
	$\underline{M}_\Delta$ = .40 $\underline{SD}$ = .25 $\underline{t}$ = 3.19*	$\underline{M}_\Delta$ = .24 $\underline{SD}$ = .12 $\underline{t}$ = 2.00	$\underline{M}_\Delta$ = .42 $\underline{SD}$ = .34 $\underline{t}$ = 1.25	$\underline{M}$ = .672
	66th %ile	60th %ile	66th %ile	75th %ile

*** $p < .001$
** $p < .01$
* $p < .05$

Attachment A: Studies Included in the Meta-Analysis

Aleamoni, L. M. (1978). The usefulness of student evaluations in improving college teaching. (Research Report No. 356). Urbana: University of Illinois, Office of Instructional Resources. (ERIC Document Reproduction Service No. ED 088 317).

Braunstein, D. N., Klein, G. A., & Pachla, M. (1973). Feedback expectancy and shifts in student ratings of college faculty. Journal of Applied Psychology, 58, 254-258.

Carter, K. R. (1974). The effect of student feedback in modifying teaching performance. Dissertation Abstracts International, 35, 5110A. (University Microfilms No. 75-2572)

Centra, J. A. (1973). Effectiveness of student feedback in modifying college instruction. Journal of Educational Psychology, 65, 395-401.

Cohen, P. A., & Herr, G. (1979). A procedure for diagnostic instruction feedback: The Formative Assessment of College Teaching (FACT) model. Educational Technology, 19(12), 18-23.

Emmer, E. T. (1974, April). Instructor perception, content of scale, and feedback effectiveness. Paper presented at the annual meeting of the American Educational Research Association, Chicago. (ERIC Document Reproduction Service No. 103 399)

Erickson, G. R., & Erickson, B. L. (1979). Improving college teaching: An evaluation of a teaching consultation procedure. Journal of Higher Education, 50, 670-683.

Erickson, G. R., & Sheehan, D. S. (1976, April). An evaluation of a teaching improvement process for university faculty. Paper presented at the annual meeting of the American Educational Research Association, San Francisco. (ERIC Document Reproduction Service No. 131 111)

Froman, R. D., & Owen, S. V. (1980, April). Influence of different types of student ratings feedback upon later instructional behavior. Paper presented at the annual meeting of the American Educational Research Association, Boston. (ERIC Document Reproduction Service No. 187 724)

Hoyt, D. P., & Howard, G. S. (1978). The evaluation of faculty development programs. Research in Higher Education, 8, 25-38.

Kellams, S. E., & Bohm, B. H. (1985, April). Are student ratings useful for improving the instruction of part-time faculty members? Paper presented at the annual meeting of the American Educational Research Association, Chicago.

Koffman, M. L. (1974). A study of the effects of instructional analysis and feedback on the classroom behavior and student achievement of university teaching assistants. Dissertation Abstracts International, 35, 302A.

Larsgaard, J. O. (1971). An experiment using a student rating scale as a form of feedback for the modification of teaching behavior at the college level. Dissertation Abstracts International, 32, 2485A. (University Microfilms No. 71-28438)

Lo Guidice, J. R. (1973). Effects of form of feedback from student ratings on faculty teaching practices and faculty attitudes. Dissertation Abstracts International, 34, 1728B. (University Microfilms No. 73-23,810)

McKeachie, W. J., Lin, Y-G., Daugherty, M., Moffet, M. M., Neigler, C., Nork, J., Walz, M., & Baldwin, R. (1980). Using student ratings and consultation to improve instruction. British Journal of Educational Psychology, 50, 168-174.

McLean, D. F. (1979). The effects of midsemester feedback upon weekly evaluations of university instructors. Unpublished master's thesis, Department of Psychology, University of Western Ontario, London, Ontario.

Miller, M. T. (1971). Instructor attitudes toward, and their use of, student ratings of teachers. Journal of Educational Psychology, 62, 235-239.

Murphy, J. B., & Appel, V. H. (1978, April). The effect of mid-semester student feedback on instructional change and improvement. Paper presented at the annual meeting of the American Educational Research Association, Toronto.

Overall, J. V., & Marsh, H. W. (1976, June). The relationship between students' evaluation of faculty and instructional improvement. Paper presented at the International Conference on Improving University Teaching, Newcastle-upon-Tyne, England. (ERIC Document Reproduction Service No. 138 165)

Pambookian, H. S. (1972). The effect of feedback from students to college instructors on their teaching behavior. Dissertation Abstracts International, 33, 4950A. (University Microfilms No. 73-6893)

Payne, D. A., & Hobbs, A. M. (1979). The effect of college course evaluation feedback on instructor and student perceptions of instructional climate. Higher Education, 8, 525-533.

Polahar, K. L. (1981). The effect of student feedback to college business instructors on the evaluation of instruction. Dissertation Abstracts International, 43, 2208A. (University Microfilms No. DA8210635)

Ramsey, R. S. (1980). The effect of evaluative feedback from students on faculty performance at the University of Mississippi. Dissertation Abstracts International, 41, 4622A. (University Microfilms No. 8108768)

Rotem, A. (1978). The effects of feedback from students to university professors: An experimental study. Research in Higher Education, 9, 303-318.

Scott, N. H. (1976). A pilot study testing the effects of feedback on teaching behavior as measured by student ratings at New River Community College, Dublin, Virginia. Dissertation Abstracts International, 37, 1950A. (University Microfilms No. 76-23,235)

Skeff, K. M. (1983). Evaluation of a method for improving the teaching performance of attending physicians. The American Journal of Medicine, 75, 465-470.

Smith, D. L. (1977). The relationship of feedback to professors of the results of student ratings of their teaching effectiveness at mid-semester to their end-of-semester ratings (Doctoral dissertation, University of Southern California). Dissertation Abstracts International, 38, 187A.

Thomas, H. B., Jr. (1969). Changing teacher behavior: An experiment in feedback from students to teachers. Dissertation Abstracts International, 30, 1903A. (University Microfilms No. 69-17,281)

Toney, J. W. (1973). The effects of feedback to teachers from student evaluations of the instructional process. Dissertation Abstracts International, 34, 7652A. (University Microfilms No. 74-12,216)

Weerts, R. R. (1978). The use of feedback from student ratings for improving college teaching (Research Report No. 6, Student Perceptions of Teaching [SPOT]: VIII). Iowa City, IO: University of Iowa, Evaluations Examination Service.

Attachment B: Student Ratings Studies Which Did Not Meet Criteria for the Meta-Analysis

Aleamoni, L. M., & Stevens, J. J. (1983, April). The effectiveness of consultation in support of student evaluation feedback: A ten year follow-up. Paper presented at the annual meeting of the Rocky Mountain Psychological Association, Albuquerque, NM. (ERIC Document Reproduction Service No. 231 313).

Bingman, R. M. (1983, May). Use of student evaluations to improve instruction. Paper presented at the meeting of the Association for Institutional Research Forum, Toronto.

Bledsoe, J. C. (1975). Insight into one's own teaching through feedback from students' evaluations. Psychological Reports, 37, 1189-1190.

Braskamp, L. A., Caulley, D., & Costin, F. (1979). Student ratings and instructor self-ratings and their relationship to student achievement. American Educational Research Journal, 16, 295-306.

Butler, J. R., & Tipton, R. M. (1976). Rating stability shown after feedback of prior ratings of instruction. Improving College and University Teaching, 24, 111-112+.

Friedlander, J. (1978). Student perceptions on the effectiveness of midterm feedback to modify college instruction. Journal of Educational Research, 71, 140-143.

Gage, N. L., Runkel, P. L., & Chatterjee, B. B. (1963). Changing teacher behavior through feedback from pupils: An application of equilibrium theory. In W. W. Charters and N. L. Gage (Eds.), Readings in the social psychology of education (pp. 173-181). Boston: Allyn and Bacon.

Green, M. J. (1979). Effect of feedback discrepancy upon university faculty Dissertation Abstracts International, 40, 1356A. (University Microfilms No. 79-19,724)

Gray, D. M., & Brandenburg, D. C. (1985). Following student ratings over time with a catalog-based system. Research in Higher Education, 22, 155-168.

Howard, G. S. (1977). A program to improve university instruction: A promising area for psychologists. Professional Psychology, 8, 316-327.

Marsh, H. W., Fleiner, H., & Thomas, C. S. (1975). Validity and usefulness of student evaluations of instructional quality. Journal of Educational Psychology, 67, 833-839.

Pambookian, H. S. (1974). Initial level of student evaluation of instruction as a source of influence on instructor change after feedback. Journal of Educational Psychology, 66, 52-56.

Pambookian, H. S. (1976). Discrepancy between instructor and student evaluations of instruction: Effect on instructor. Instructional Science, 5, 63-75.

Sherman, T. M. (1977-78). The effects of student formative evaluation of instruction on teacher behavior. Journal of Educational Technology Systems, 6, 209-217.

Siagian. E. (1983). The impact of student evaluative feedback on the perception of the teaching faculty at Andrews University. Dissertation Abstracts International, 44, 1254A. (University Microfilms No. DA8320339)

Smith, W. G. (1975). The influence of student feedback on changing teaching behaviors of college faculty members. Dissertation Abstracts International, 36, 4424A. (University Microfilms No. 76-1131).

Stevens, J. J., & Aleamoni, L. M. (1985). The use of evaluative feedback for instructional improvement: A longitudinal perspective. Instructional Science, 13, 285-304.

Stillman, P. L., Gillers, M. A., Heins, M., Nicholson, G., & Sabers, D. L. (1983). Effect of immediate student evaluations on a multi-instructor course. Journal of Medical Education, 58, 172-178.

Trank, D. M. (1978). The effect of student feedback on the improvement of instruction. Iowa City: University of Iowa. (ERIC Document Reproduction Service No. 177 642)

Tuckman, B. W., & Oliver, W. F. (1968). Effectiveness of feedback to teachers as a function of source. Journal of Educational Psychology, 59, 279-301.

Tuckman, B. W., & Yates, D. (1980). Evaluating the student feedback strategy of changing teacher style. Journal of Educational Research, 74(2), 74-82.

Vogt, K. E., & Lasher, H. (1973). Does student evaluation stimulate improved teaching? Ohio: Bowling Green State University, College of Business Adminstration. (ERIC Document Reproduction Service No. 078 748)

Waters, J. E. (1983). Evaluating and changing the social climate. An introductory sociology class. Teaching Sociology, 10, 219-223.

Wilson, R. C. (1986). Improving faculty teaching: Effective use of student evaluations and consultants. Journal of Higher Education, 57, 196-211.