

DOCUMENT RESUME

ED 417 664

HE 031 120

AUTHOR Murray, Harry G.; Jelley, R. Blake; Renaud, Robert D.
 TITLE Longitudinal Trends in Student Instructional Ratings: Does Evaluation of Teaching Lead to Improvement of Teaching?
 PUB DATE 1996-04-00
 NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (New York, April 1996).
 PUB TYPE Reports - Research (143) -- Speeches/Meeting Papers (150)
 EDRS PRICE MF01/PC01 Plus Postage.
 DESCRIPTORS *College Faculty; Educational Trends; Faculty Development; *Faculty Evaluation; Foreign Countries; Higher Education; *Instructional Improvement; Longitudinal Studies; Social Sciences; *Student Evaluation of Teacher Performance; Trend Analysis
 IDENTIFIERS *University of Western Ontario (Canada)

ABSTRACT

This study examined effects of student evaluation of faculty teaching for 7 departments in the Faculty of Social Science at the University of Western Ontario over a 21-year period. The sample of teachers included 1322 faculty members who had taught undergraduate courses in one or more year between 1973-74 through 1993-94. The same 10-item teaching evaluation form was used continuously throughout this period. The evaluation form focused on classroom teaching skills such as explaining clearly, showing enthusiasm, and encouraging student participation. Significant improvement across years was found for 5 of the 7 departments, for the faculty as a whole, and for a fixed group of 72 faculty members who had taught continuously throughout the 21-year observation period. These results, in combination with similar evidence from faculty opinion surveys and field experiments on student feedback, support the view that student evaluation of teaching contributes significantly to improvement of teaching quality. (DB)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED 417 664

Longitudinal trends in student instructional ratings: Does evaluation of teaching lead to improvement of teaching?

Harry G. Murray, R. Blake Jelley, and Robert D. Renaud
Department of Psychology
University of Western Ontario
London, Canada N6A 5C2

Presented at annual meeting of the
American Educational Research Association
New York, April, 1996

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

- This document has been reproduced as received from the person or organization originating it.
- Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

Abstract

Mean student ratings of teaching for 7 departments in the Faculty of Social Science, University of Western Ontario, were compared longitudinally over the 21-year period since the advent of student evaluation in 1973. Significant improvement across years was found for 5 of 7 departments, for the the faculty as a whole, and for a fixed group of 72 faculty members who had taught continuously in the faculty throughout the 21-year observation period. These results, in combination with similar evidence from faculty opinion surveys and field experiments on student feedback, support the view that student evaluation of teaching contributes significantly to improvement of teaching quality.

1/28
1/28
1/28

PERMISSION TO REPRODUCE AND
DISSEMINATE THIS MATERIAL
HAS BEEN GRANTED BY

Harry G. Murray

BEST COPY AVAILABLE

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)



Longitudinal trends in student instructional ratings: Does evaluation of teaching lead to improvement of teaching?

Introduction

Student instructional ratings have gained widespread acceptance over the past 30 years as a measure of teaching effectiveness in colleges and universities. Nearly 100% of postsecondary institutions now have some sort of plan for student evaluation of teaching, with results used both as feedback to faculty members and as input to personnel decisions.

Given that student evaluation of teaching is so widely implemented, and given that one of the main justifications for introducing student evaluation was to improve teaching, it would be interesting to know whether student evaluation has in fact contributed to improvement of teaching. Despite the large volume on the reliability and validity of student evaluation of teaching, it has yet to be established that student evaluation has a positive impact on quality of teaching.

One way of assessing the formative impact of student evaluation of teaching is to survey the opinion of faculty members who have undergone the evaluation process. Across 8 faculty surveys reviewed by Murray (1996), 73% of respondents said that student evaluation provided useful feedback and 69% said that it had led to improved teaching. Although this type of data is potentially affected by limited return rate, self-report bias, and uncontrolled variables, it is interesting to note that a clear majority of faculty members seems to believe that student evaluation has indeed contributed to improvement of teaching.

A second way of investigating whether student evaluation improves teaching is to carry out a field experiment in which randomly assigned experimental teachers receive feedback concerning mid-course student evaluation of teaching, whereas control teachers are evaluated at midterm but given no feedback. The two groups are then compared on end-of-course student ratings, with the expectation that experimental teachers will show higher ratings as a result of the beneficial effects of feedback. Cohen (1980) conducted a meta-analysis of 22 field experiments of this type, and concluded that feedback from student ratings alone leads to modest improvement in faculty teaching performance, whereas student feedback supplemented either by expert consultation leads to more substantial gains in quality of teaching. Field experiments provide further support for the view that student evaluation leads to improved teaching, even with extraneous variables controlled and self-report biases eliminated, but field experiments have their own methodological limitations, including (1) artificiality, and (2) a very short time frame, usually 2 to 3 months.

A third way of assessing the contribution of student evaluation to improvement of teaching is compare mean student ratings of teaching longitudinally over over a period of several years in a particular academic unit (department or faculty) following the introduction of student evaluation of teaching in that unit. If student evaluation contributes to improvement of teaching, this improvement should be reflected in a gradual increase in the average teacher rating for the unit as a whole. This approach, which was followed in the present study, has the advantage of assessing improvement under real-world conditions and from a long-term perspective.

Ideally, a valid test of the longitudinal improvement hypothesis requires the following conditions: (1) mean ratings are compared across a minimum of 10 years, or 10 semesters for a fixed group of teachers; (2) tracking of mean ratings across years begins in the same year where student evaluation was first introduced; (3) the same student rating form is used throughout the study; and (4) all faculty and all courses undergo student evaluation in all years.

Published research on longitudinal trends in student ratings of teaching has yielded mixed results. Of 14 studies located by the present authors, 8 reported significant longitudinal improvement and 6 reported no significant change in student ratings over time. However, as outlined below, most studies conducted to date have failed to fulfill the four methodological conditions identified above. For example, Gray and Brandenberg (1985) found significant longitudinal improvement in mean student ratings of teaching in a sample of 304 faculty members from various academic disciplines at the University of Illinois, but ratings were tracked over only four consecutive semesters, and the study did not begin in the semester where student evaluation was introduced. Vogt and Lasher (1973), on the other hand, found no significant improvement in mean student ratings for a group of 50 business professors at Bowling Green State. Longitudinal tracking of mean ratings began concurrently with the advent of student evaluation in the Vogt and Lasher study, but ratings were compared across only eight academic quarters between 1969 and 1972.

Marsh and Hocevar (1991) conducted a large-scale longitudinal study of student ratings of teaching that fulfilled all of the four methodological conditions listed above. The sample of teachers consisted of 195 faculty members from 31 departments at the University of Southern California, each of whom had been evaluated in each of at least 10 different years over a 13-year period from 1976 to 1988. All instructors were evaluated by the same evaluation form, namely the Students' Evaluations of Educational Quality (SEEQ) instrument. Ratings of a given instructor on each of the 11 SEEQ dimensions were averaged across all courses taught in the same year, and trends across years were assessed by multiple regression procedures. It was found that there was virtually no change in mean

student ratings across the 13-year observation period. The correlation between year and rating was significant (but in a negative direction) for only 2 of 11 SEEQ dimensions, and year accounted for less than 1% of variance in student ratings. Thus, despite the use of a large sample and powerful design, the Marsh and Hocevar study provided no evidence that mean student ratings improve longitudinally following the introduction of student evaluation of teaching.

Method

The present study also fulfilled the four methodological conditions identified above, and was conducted with a larger sample and over a longer time frame than any previous study. The sample of teachers included 1322 faculty members who had taught undergraduate courses in the Faculty of Social Science, University of Western Ontario, in one or more of 21 consecutive academic years extending from 1973-74 to 1993-94. Each of the seven constituent departments of the Faculty of Social Science (Anthropology, Economics, Geography, History, Political Science, Psychology, and Sociology) has used the same 10-item teaching evaluation form continuously since 1973, the point at which student evaluation was introduced in the Faculty. The evaluation form focuses on classroom teaching skills such as explaining clearly, showing enthusiasm, and encouraging student participation, each of which is rated on a 5-point scale. The evaluation form is administered annually in all courses under standard conditions, with results used on a compulsory basis in promotion and tenure decisions.

Results

To obtain an annual measure of overall teaching effectiveness for each faculty member, student rating data were averaged across all items of the evaluation form and across all courses taught in a given academic year. Trends across years in department or faculty mean ratings were assessed by fitting a regression line to the data points and testing the deviation of its slope from zero. The major results of the present study were as follows:

1. Mean student ratings of teaching increased significantly across the 21-year observation period for the Faculty of Social Science as a whole (see Figure 1). It may be noted that the average teacher rating increased from approximately 3.70 in the mid-1970's to approximately 3.90 in the mid-1990's, which corresponds to a gain of approximately .67 standard deviation units. A regression line fitted to the faculty-wide data was found to deviate significantly from zero, and the correlation between year and faculty mean rating was .85. This result differs from what was reported by

Marsh and Hocevar (1991) and shows that is possible to get longitudinal improvement in student ratings under some conditions.

2. Significant longitudinal improvement in mean student ratings was found in some individual departments but not in others (see Figure 2). Departments A, B, C, E, and F showed significant improvement, as indicated by correlations between mean rating and year ranging from .56 to .85, whereas Departments D and G, with correlations of -.10 and .11 respectively, did not show significant improvement. These results suggest that it is possible to get conflicting longitudinal results even among similar academic units in the same institution using the same teaching evaluation form. Thus the conflicting results of previous studies, and in particular the negative results of the Marsh and Hocevar (1991) study, are perhaps not so surprising.
3. An important limitation of the data in Figures 1 and 2 is that annual mean ratings are based on a sample of teachers that varies somewhat from year to year due to faculty turnover. Thus, the possibility exists that year-to-year gains are due, not to longitudinal improvement in a fixed group of teachers (improvement by development), but rather to a tendency for newly appointed faculty members to be better teachers, on average, than the individuals they replace (improvement by selection). To check on this possibility, a subsample of 72 faculty members was identified who had held positions in various departments of the Faculty of Social Science for 21 consecutive years and had taught in undergraduate courses in at least 17 of those 21 years. Data for missed years (of which there were never more than two in succession) were estimated by interpolation. Figure 3 shows annual mean student rating scores for the fixed group of 72 faculty members and for the Faculty as a whole. Statistical analysis indicated that the fixed group of teachers showed significant longitudinal improvement over the 26-year observational period, but the amount of improvement shown by this group was significantly less than that for the department as a whole. The correlation coefficient between year and mean student rating was .49 for the fixed group of teachers, as compared to .85 for the Faculty as a whole. These results indicate that the longitudinal gains in teacher ratings depicted in Figures 1 and 2 are due in part to true longitudinal development in individual teachers and in part to the tendency of new faculty members to be more effective teachers than the individuals they replace.

Discussion

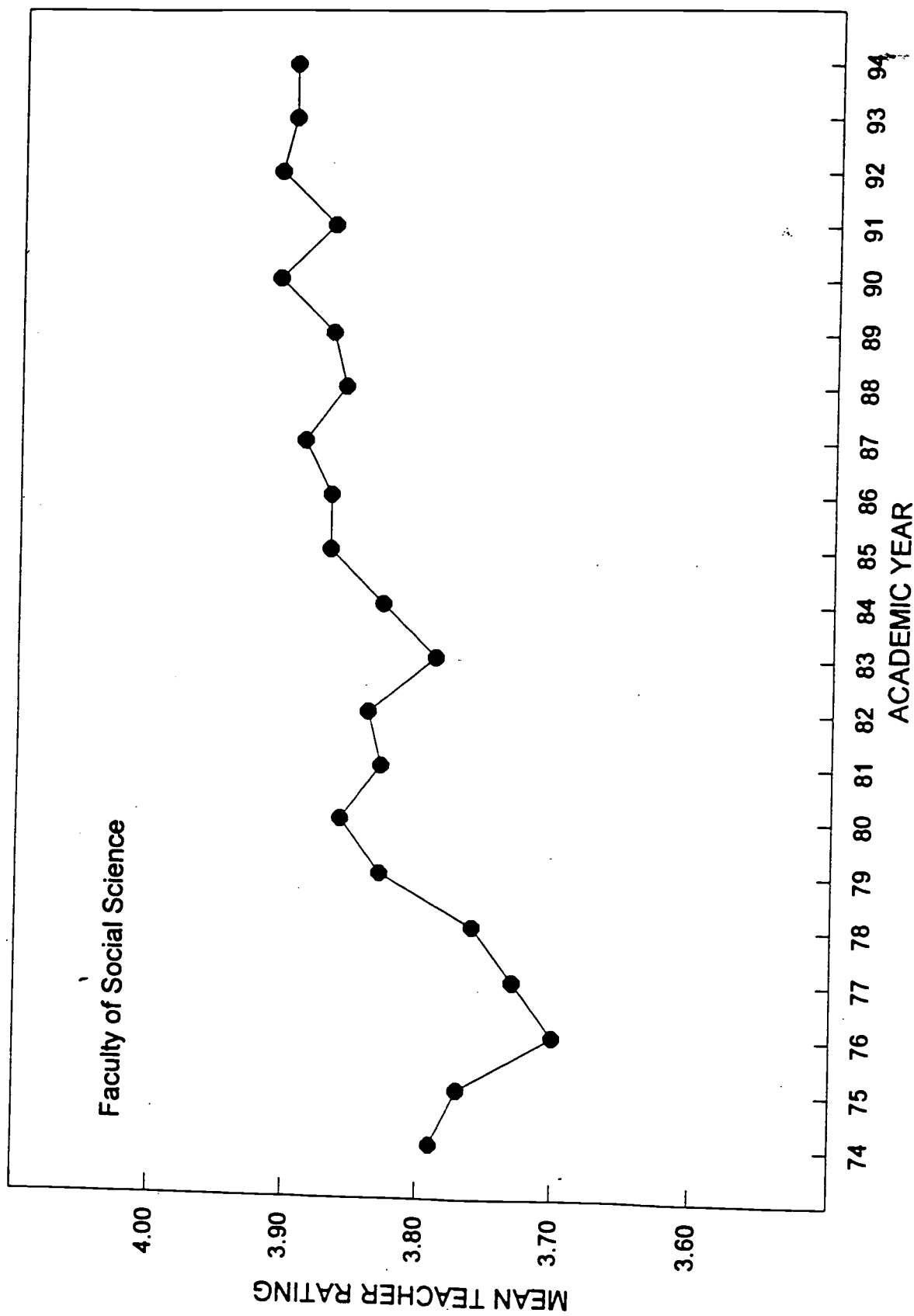
The results of this study suggest that, at least under certain conditions, the introduction of student evaluation of teaching in an academic unit can lead to long-term improvement in teaching in that unit. This finding is consistent with positive evidence from faculty opinion surveys and field experiments reviewed above. This convergence of evidence across three methodologically distinct areas of research (faculty surveys, field experiments, and longitudinal comparisons) gives credibility to the view that student evaluation does indeed contribute significantly to improvement of teaching. Despite this positive conclusion, there are some important questions that arise in relation to the present data:

1. It appears that longitudinal improvement in teaching sometimes occurs and sometimes does not occur following the introduction of student evaluation in an academic unit. But the reasons for this inconsistency are not clear. What factors are responsible for finding long-term improvement in rated teaching effectiveness in some academic units but not in others? Could faculty participation in instructional development programs, such as workshops, courses, and peer consultation, be one of the factors that makes a difference? Could mandatory use of student evaluation of teaching in faculty personnel decisions be a factor that contributes to longitudinal improvement in an academic unit? These are interesting questions that invite further research.
2. The finding that student ratings of teaching increase significantly across years for a fixed group of faculty members is difficult to reconcile with the conclusion of several previous researchers (including two of the present researchers, Renaud & Murray, 1996) that faculty age correlates negatively with student instructional ratings? Is this anomaly related to the that a longitudinal design was used in the present study (at least for the fixed group of teachers), whereas a cross-sectional design was used in most studies finding a negative correlation between age and ratings?
3. One possible interpretation of the present results is that student evaluation of teaching leads to improvement of certain aspects of teaching only, namely those aspects that are measured by the typical student evaluation form (eg., clarity of explanation, promptness of feedback, encouragement of participation). These improvements notwithstanding, is it possible that other aspects of teaching, such as grading standards, academic requirements, and willingness to innovate, have not benefited from student

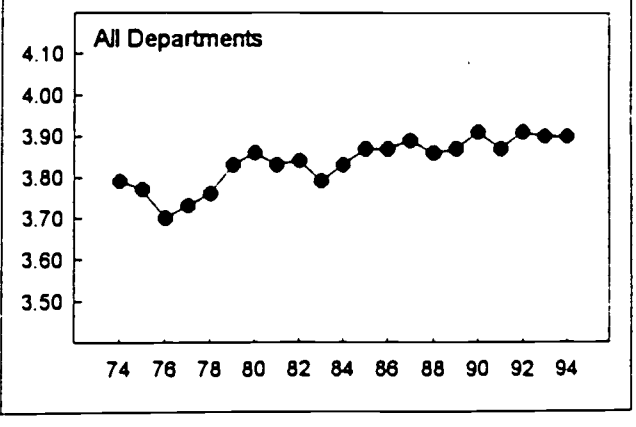
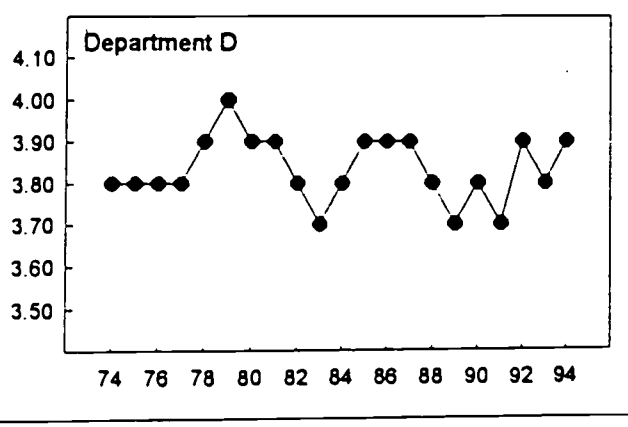
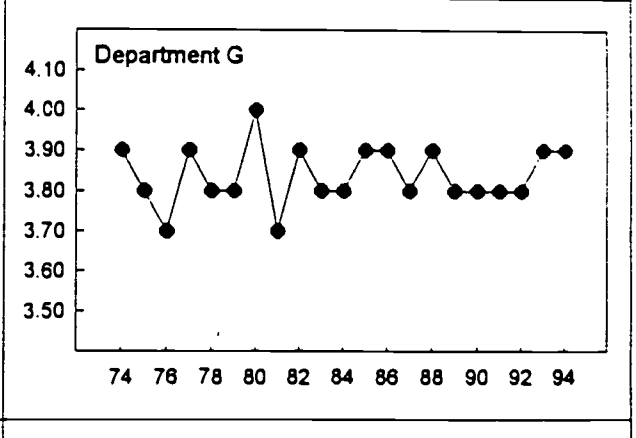
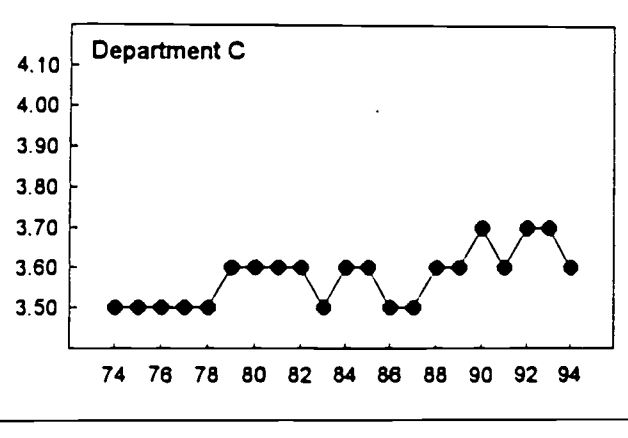
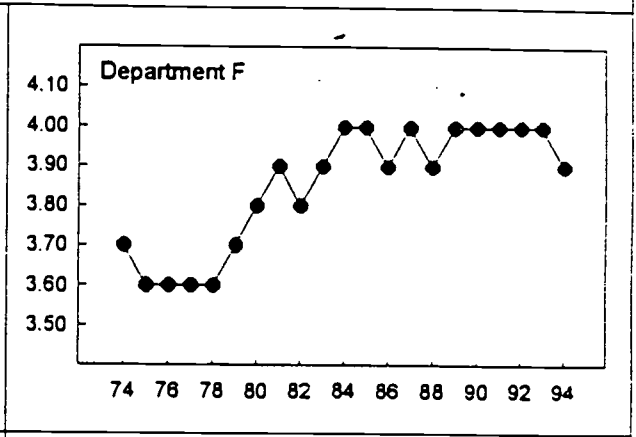
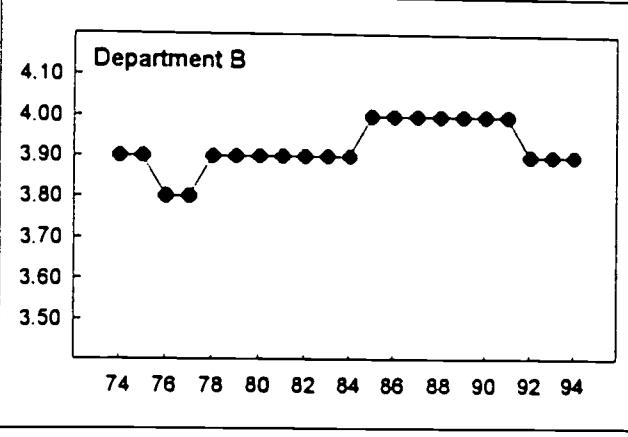
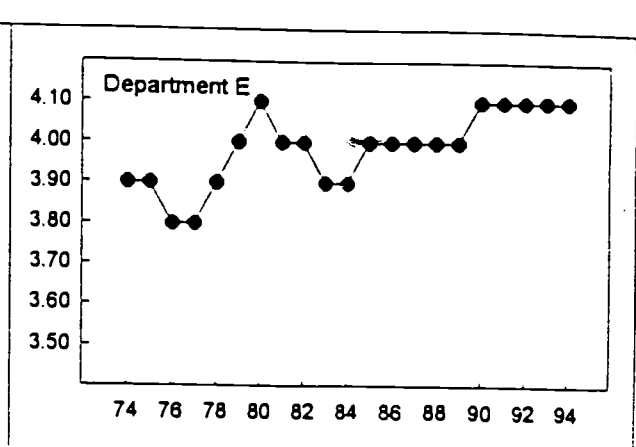
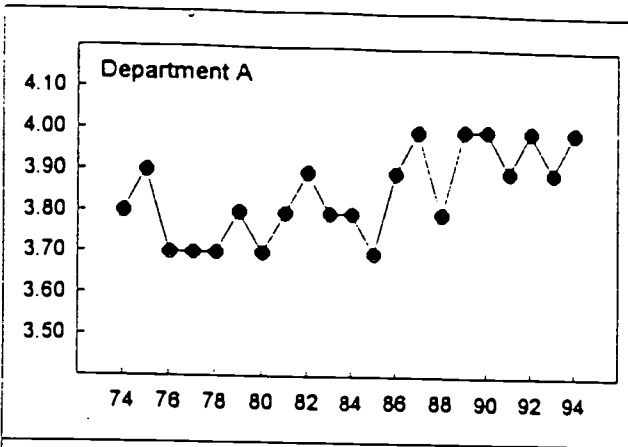
evaluation, and in fact, have actually gone in the opposite direction (i.e., gotten worse) as a result of student evaluation of teaching?

References

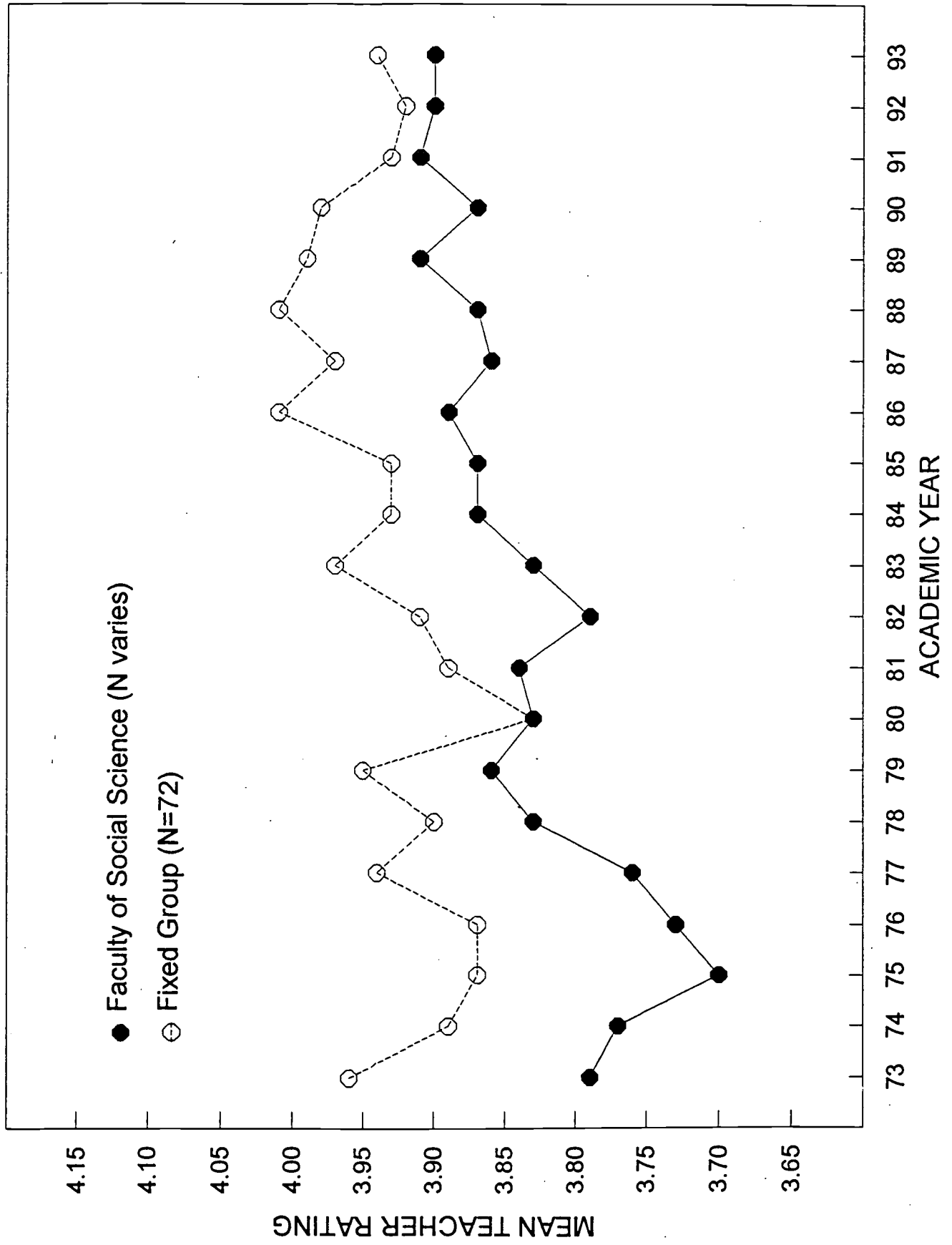
- Cohen, P.A. (1980). Effectiveness of student rating feedback for improving college instruction: A meta-analysis of findings. Research in Higher Education, *13*, 321-341.
- Gray, D.M. & Brandenburg, D.C. (1985). Following student ratings over time with a catalog-based system. Research in Higher Education, *22*, 155-168.
- Marsh, H.W. & Hocevar, D. (1991). Students' evaluations of teaching effectiveness: The stability of mean ratings of the same teachers over a 13-year period. Teaching and Teacher Education, *7*, 303-341.
- Murray, H.G. Does evaluation of teaching lead to improvement of teaching? (1996). International Journal of Academic Development, In Press.
- Renaud, R.D. & Murray, H.G. (1996). Aging, personality, and teaching effectiveness in academic psychologists. Research in Higher Education, In Press.
- Vogt, K.E. & Lasher, H. (1973). Does student evaluation stimulate improved teaching? College of Business, Bowling Green State University (ERIC Document No. ED 078 748).



MEAN TEACHER RATING



ACADEMIC YEAR





REPRODUCTION RELEASE

(Specific Document)

I. DOCUMENT IDENTIFICATION:

Title: Longitudinal Trends in Student Instructional Ratings: Does Evaluation of Teaching Lead to Improvement of Teaching?	
Author(s): Harry G. Murray, R. Blake Jelley and Robert D. Renaud	
Corporate Source:	Publication Date: Apr 96

II. REPRODUCTION RELEASE:

In order to disseminate as widely as possible timely and significant materials of interest to the educational community, documents announced in the monthly abstract journal of the ERIC system, *Resources in Education* (RIE), are usually made available to users in microfiche, reproduced paper copy, and electronic/optical media, and sold through the ERIC Document Reproduction Service (EDRS) or other ERIC vendors. Credit is given to the source of each document, and, if reproduction release is granted, one of the following notices is affixed to the document.

If permission is granted to reproduce and disseminate the identified document, please CHECK ONE of the following two options and sign at the bottom of the page.

The sample sticker shown below will be affixed to all Level 1 documents

The sample sticker shown below will be affixed to all Level 2 documents



Check here
For Level 1 Release:
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical) and paper copy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 1



Check here
For Level 2 Release:
Permitting reproduction in microfiche (4" x 6" film) or other ERIC archival media (e.g., electronic or optical), but *not* in paper copy.

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL IN OTHER THAN PAPER COPY HAS BEEN GRANTED BY

Sample

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

Level 2

Documents will be processed as indicated provided reproduction quality permits. If permission to reproduce is granted, but neither box is checked, documents will be processed at Level 1.

"I hereby grant to the Educational Resources Information Center (ERIC) nonexclusive permission to reproduce and disseminate this document as indicated above. Reproduction from the ERIC microfiche or electronic/optical media by persons other than ERIC employees and its system contractors requires permission from the copyright holder. Exception is made for non-profit reproduction by libraries and other service agencies to satisfy information needs of educators in response to discrete inquiries."

Sign here → please

Signature: <i>HGMurray</i>	Printed Name/Position/Title: <i>Harry G. Murray / Professor</i>	
Organization/Address: <i>Dept. of Psychology U. Western Ontario London, Canada N6A 5C2</i>	Telephone: <i>519-661-4109</i>	FAX: <i>519-661-3961</i>
	E-Mail Address: <i>murray@sscl.uwo.ca</i>	Date: <i>Dec 3 /97</i>

III. DOCUMENT AVAILABILITY INFORMATION (FROM NON-ERIC SOURCE):

If permission to reproduce is not granted to ERIC, or, if you wish ERIC to cite the availability of the document from another source, please provide the following information regarding the availability of the document. (ERIC will not announce a document unless it is publicly available, and a dependable source can be specified. Contributors should also be aware that ERIC selection criteria are significantly more stringent for documents that cannot be made available through EDRS.)

Publisher/Distributor:
Address:
Price:

IV. REFERRAL OF ERIC TO COPYRIGHT/REPRODUCTION RIGHTS HOLDER:

If the right to grant reproduction release is held by someone other than the addressee, please provide the appropriate name and address:

Name:
Address:

V. WHERE TO SEND THIS FORM:

Send this form to the following ERIC Clearinghouse:	ERIC CLEARINGHOUSE ON HIGHER EDUCATION THE GEORGE WASHINGTON UNIVERSITY ONE DUPONT CIRCLE, SUITE 630 WASHINGTON, D.C. 20036-1183
-----------------------------------------------------	-------------------------------------------------------------------------------------------------------------------------------------------

However, if solicited by the ERIC Facility, or if making an unsolicited contribution to ERIC, return this form (and the document being contributed) to:

ERIC Processing and Reference Facility
1100 West Street, 2d Floor
Laurel, Maryland 20707-3598

Telephone: 301-497-4080
Toll Free: 800-799-3742
FAX: 301-953-0263

e-mail: ericfac@inet.ed.gov
WWW: <http://ericfac.piccard.csc.com>