

A device for interpreting the results of student evaluation forms

George L. Findlen

Writing program administrators often have the task of evaluating the teaching of writing instructors. As the debate between Robert Powell and Kenneth Eble indicates ("Comment and Response," *College English*, 40 (January 1979), 559-567), interpreting the data from student evaluations of teaching is no simple matter. I agree with Eble and Larson, who believe that student evaluation data, despite their limitations, are an important source of information—although only one—about faculty teaching performance. In most cases, student evaluation involves an evaluation form of some fairly common type. Designing such a form is not my concern here. What concerns me are two problems raised by the use of any such form.

The first problem is logistical. Take a faculty of twenty, for example, each member teaching four sections, twenty students in each section. Giving every student a twenty-item evaluation form produces 32,000 pieces of information each semester. Arranging and interpreting this mass of data requires considerable skill. Few WPAs, however, have studied methods of descriptive statistics for reducing large quantities of data to a meaningful and manageable form. Even fewer WPAs have studied data processing.

A second problem involves interpreting results, once the data has been analyzed. When students record their estimate of how much a faculty member displays a trait or does a task, a higher rating is presumably better than a lower one. But a rating of 3 alone, on a scale of 6, really tells us little, since it does not tell us how high or low students usually rate instructors. A rating of 3 is low, not average, for example, if students never rate instructors 1 or 2. Thus, rating scale results become meaningful only when we compare results for a given instructor in a given course in a given semester either to that instructor's results for that course in other semesters, or to the results obtained by other instructors who teach the same course in the same semester.

The fairest way to rate teachers with student evaluation forms, therefore, is comparative. And the quickest way to make a fair comparison is visually. We must acknowledge, however, that visual comparison is approximate. It can give us only one rough indication of an instructor's performance. To attempt more precision denies that teaching and learning are very complex behaviors.

The most efficient means for representing figures visually is a line graph, because it can reduce a mass of numbers into an intelligible, useful form to make meaningful comparisons. Two department chairs at Tarkio College developed the graph described here: John E. Atkinson of the Science and Mathematics Department, and John R. Dubinski of the Language and Literature Department. (Tarkio College is an independent, four-year liberal arts college of 350 students.)

The graph is designed so that departments large or small can use it, and so that the data can be hand-tabulated.

How to make and use a line graph

For each faculty member, tally the ratings from all composition students and average each item's ratings. Then plot each item average on a blank evaluation form and connect each point with a line. The following example should clarify this procedure.

First, prepare a tally sheet of each instructor; write the evaluation form's item numbers in a vertical column on the left and the rating choices across the top. Second, record each student's choices on the tally sheet. Since students sometimes do not mark every item, and since only the number who mark the item is used to determine the item average, record how many students mark each item in the far right column.

Figure 1. Sample Tally Sheet

Item	Rating						Number of student answers
	6	5	4	3	2	1	
9	IIII	IIII	IIII	I			20
		IIII	II				
		I					

If you are curious, make a separate tally for each section an instructor has taught. If time is short, make only one tally for all courses an instructor has taught that semester. At Tarkio, the secretary makes only one tally for each faculty member; the tally includes evaluations from all courses taught by the faculty member in that term.

After finishing each tally, calculate the item averages. To do this, multiply the number of students who circle each rating by that rating; add the products; and divide the resulting sum by the total number of students who rated the instructor on that item. For example, averaging the data from Item 9 in Figure 1 produces the following: $[(4 \times 6) + (9 \times 5) + (6 \times 4) + (1 \times 3)] \div 20 = 4.8$. A secretary can do this calculation quickly and easily using a pocket calculator that has a memory key. Put each product into memory plus (M+); when all multiplications are done, touch the memory recall key (MR) and divide by the total number of students who rated the instructor on that item. Repeat for each item. Record each item average in the margin of a blank evaluation form. This is the individual instructor's summary-of-results data sheet. See Figure 2.

Figure 2. Data for Professor X, All Courses, Fall, 1977

Please rate your professor on each of these items by drawing a circle around the number that best indicates his/her position in comparison with other teachers you have had. Rate each item as thoughtfully as possible. Do not omit items.

	6	5	4	3	2	1	Average
1. Interests me in the subject	6	5	4	3	2	1	4.6
2. Has helped broaden my interests	6	5	4	3	2	1	4.4
3. Has given me new viewpoints	6	5	4	3	2	1	4.4
4. Has increased my skills in thinking	6	5	4	3	2	1	4.4
- 5. Has motivated me to do my best work	6	5	4	3	2	1	4.3
6. Is clear and understandable in his/her explanations	6	5	4	3	2	1	4.8
7. Makes good use of examples and illustrations	6	5	4	3	2	1	5.1
8. Interprets abstract ideas and theories clearly	6	5	4	3	2	1	4.7
9. Stresses important material	6	5	4	3	2	1	4.8
10. Inspires confidence in his/her knowledge of the subject	6	5	4	3	2	1	5.1
+ 11. Has a sense of humor	6	5	4	3	2	1	5.5
12. Considering everything, I would rate him/her	6	5	4	3	2	1	5.1

Your professor would like to know if there is something you believe he/she has done especially well in the teaching of this course:

Your professor would also like to know what specific things you believe might be done to improve his/her teaching in this course:

Thus far your judgments have been restricted to characteristics of the teacher. For the item below indicate your feeling for the *subject matter* of the course by checking the appropriate entry:

The subject matter or content of the course is:

highly interesting	<u>17</u>
moderately interesting	<u>24</u>
not very interesting	<u>6</u>

○ = rating circled most often; + = item rated highest; - = item rated lowest

Tallying answers and recording averages of a twelve-item instrument takes only twenty minutes for a class of twenty students. For a twenty-member department, three secretaries can produce one line graph for every instructor in one full day of work. At Tarkio, where five faculty members teach a total of fifteen sections (a full load is three four-hour courses per semester), the one department secretary does all tallying, calculating, and recording in four hours.

Once item averages are calculated, plot a graph connecting the averages on each instructor's summary sheet. The writing program administrator may do this personally to maintain faculty privacy. Then copy each line graph onto a departmental summary. If the faculty is small, say ten, use different colored pencils or other linear codes to trace each person's information onto the department summary. Keep a legend on a separate sheet of paper to identify the chart's confidentiality. *Figure 3* illustrates this kind of summary.

If the faculty is large, anything over ten, trace every faculty member's information onto the department summary in straight lines in black ink. Then clip an overhead projector transparency sheet over each instructor's summary. Trace each instructor's line in red ink, one transparency per instructor. Be sure to identify each transparency. Also, trace the number at the left of the first and last items in order to know where to place the transparency sheet when laying it on the department summary. To make the comparison, hold the transparency over the department summary. An individual's line which is further than the others on the left indicates a strength. An individual protrusion to the right indicates a weakness. When all lines tend to the left or right, the group as a whole has a strength or weakness in that area.

At Tarkio College, department chairs make up both the individual summary sheet and the divisional summary sheet. In addition to examining the latter, they read, for each instructor, two letters of peer evaluation (one person chosen by the instructor, one chosen by the chair) and a letter of self-evaluation. The chairs then meet with each instructor to discuss the information received and individual plans for continued professional development. The chairs give instructors their individual summary in advance of this meeting. The instructors retain the summary after the meeting. Individual faculty see the department summary, which the chair then keeps.

Benefits of using a line graph

A line graph reduces a mass of data to a manageable form. The results are compact: one sheet per instructor plus a department summary sheet for large departments, and only one summary sheet for small departments. The graph ends the confusion that results from looking at a ten-foot-long, green and white computer printout with numbers of respondents, means, standard deviations, decile ranks, and split-half reliabilities for all forty-seven items and every section of every instructor. In addition, putting the graph together takes relatively little effort. Teaching is too important not to give an hour per instructor each semester to prepare information about performance.

A line graph presents the data in a meaningful fashion. A writing program administrator can see at a glance how faculty members do in comparison to one

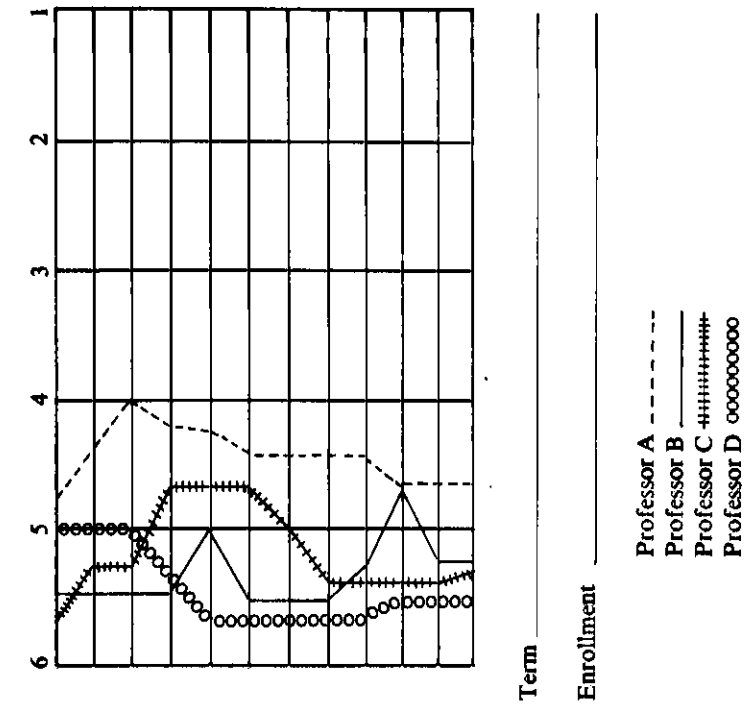


Figure 3. Student Evaluation Form, Department Summary, Fall, 1977

1. Interests me in the subject
2. Has helped broaden my interests
3. Has given me new viewpoints and/or appreciations
4. Has increased my skills in thinking
5. Has motivated me to do my best work
6. Is clear and understandable in his/her explanations
7. Makes good use of examples and illustrations
8. Interprets abstract ideas and theories clearly
9. Stresses important material
10. Inspires confidence in his/her knowledge of the subject
11. Has a sense of humor
12. Considering everything, I would rate him/her

Faculty Member

Course

Term

Enrollment

Professor A - - - - -
 Professor B _____
 Professor C +++++++
 Professor D ooooooooo

another. Thus, the graph permits the WPA to identify both individual and group strengths and weaknesses. With pertinent information, the whole point of evaluation, the WPA can make better decisions regarding future faculty development. If a majority of the faculty score low on "Has motivated me to do my best work," for example, and if instructional development personnel or money are available, then the WPA might arrange for one or more workshops on motivation. What holds true for the department holds true for individuals with particular weaknesses: the chair can give them specific directives and help make necessary improvements. At Tarkio, department chairs encourage faculty in need of improvement to make use of seminars and workshops organized by the Kansas City Regional Council for Higher Education. Merit increases reflect the faculty member's improvement from year to year.

Resources

Those wanting a practical guide to conducting faculty evaluation or to setting up an evaluation program should turn to Richard I. Miller's two books, *Evaluating Faculty Performance* (San Francisco: Jossey-Bass, 1972) and *Developing Programs for Faculty Evaluation* (San Francisco: Jossey-Bass, 1974). Kenneth Doyle reviews and analyzes the research on the student evaluation of teaching in his *Student Evaluation of Instruction* (Lexington, MA: Heath, 1975). Gilles Nadeau's article, "Student Evaluation of Instruction: the Rating Questionnaire," does the same in Christopher K. Knapper et al., *If Teaching Is Important...* (Toronto: Clarke, Irwin, 1977), pp. 73-128. Richard L. Larson reviews issues, problems, and methods of evaluation in *The Evaluation of Teaching College English* (New York: ERIC and MLA, 1970).