

# The Teacher

## Improving Teaching and Its Evaluation: A Survey of Political Science Departments<sup>1</sup>

J. Clifford Fox and Scott Keeter, *Virginia Commonwealth University*

The Department of Political Science at Virginia Commonwealth University has long viewed undergraduate instruction as a critical element of its mission, and teaching effectiveness has been a key criterion in its hiring decisions over the past several years. However, as with many departments in the United States, VCU's political science department has used few systematic methods of evaluating and improving the quality of instruction. It was fortuitous that a 1994 grant from the University's Committee on Roles and Rewards provided both the incentive and the means to address this problem.

While a variety of activities were undertaken under the grant to improve instruction, the principal activity was a mail survey of 140 randomly selected political science departments in U.S. colleges and universities. There were two broad goals for this survey: first, to gather information that would help us (a) improve our own department's methods for evaluating teaching, (b) discover effective techniques for faculty development related to instruction, and (c) reexamine how assessments of teaching effectiveness are used in administrative decisions about faculty; second, to document the "state of current practice" in the discipline related to teaching evaluation and improvement.

### Survey Methods

The two-page survey included four major sections. The first was a list of nine commonly used methods for evaluating instruction in

political science departments with a yes or no checkbox beside each. A tenth item allowed other techniques of evaluation to be identified by write-in. A second section consisted of a list of items to ascertain methods of instructional development. Follow-up questions determined how available these methods were and whether each of the methods was required of faculty or merely voluntary.

A three-column table enabled respondents to estimate the weight given to teaching, research, and other factors in each of these administrative decisions. Finally, open-ended items obtained detailed information on how evaluation was used in faculty development and what innovative techniques were used in evaluation and development. Departments were also asked to provide copies of any documents that would be useful to our study—most commonly, forms used for student evaluation of instruction.

A systematic sample of 140 political science departments was drawn from the 1993–94 *APSA Directory of Political Science Department Chairpersons* (a population of 1,282 departments). The sample was stratified to include an equal number of graduate and undergraduate departments. Two follow-up mailings were sent to nonresponding departments. The response rate for the total sample was 71% (100/140), and the response rates for each of the strata were 69% for graduate departments (48/70) and 74% for undergraduate departments (52/70). We also received 55 supporting documents, including 40 examples of student evaluation forms.

### Findings

Responses to the survey reveal a general concern for the quality of teaching in administrative decision making. For the three categories of decision making that we included in the questionnaire—salary, tenure, and promotion to full professor—significant weight was reported to be given to teaching in each case. While teaching was weighted most heavily in the tenure decision (a mean weighting of 45%, compared to 41% for research, and 14% for other factors), it was also reported to be important in both promotion (39% for teaching, 47% for research, 14% for other) and salary (42% for teaching, 38% for research, 19% for other). Not surprisingly, undergraduate departments weighted teaching more heavily than graduate departments. While the magnitude of the actual emphasis on teaching in decision making may be questioned, these results indicate that, at the very least, departments engage in the rhetoric of emphasizing teaching.

If teaching plays an important role in tenure and promotion in political science departments, then systems for evaluating teaching and developing instruction should be important for the advancement of faculty. The survey results support this idea, indicating widespread commitment to some type of systematic teaching evaluation by political science departments, with all but one department reporting that they use one or more methods of formally evaluating teaching. However, the depth and complexity of evaluation varies widely, and there is little evidence of a systematic

**TABLE 1**  
Number of Information Sources for Teaching Evaluation

Number of Evaluation Methods Used	All Departments	Type of Department	
		Undergraduate	Graduate
None	1%	0%	2%
One	5%	2%	8%
Two	22%	19%	25%
Three	31%	29%	33%
Four	25%	33%	17%
Five	12%	15%	8%
Six	3%	2%	4%
Seven	1%	0%	2%

commitment to developing instruction in response to the results of these evaluations.

**Variety of Data Sources for Instructional Evaluation**

Despite arguments in the literature for a system of teaching evaluation that includes “several sources of data,” (Miller 1987, 27), our results indicate that political science departments generally rely upon limited sources of data for this purpose (see Table 1). Of the 10 possible sources of evaluation data that were listed on the questionnaire, 58% of responding departments relied upon three or fewer sources of data for teaching evaluation. Moreover, this figure is slightly inflated by the fact that some departments use multiple forms for student assessment of instruction (e.g., a college-mandated evaluation instrument and one specifically created for department use). Using multiple forms for students to assess teaching does not increase the number of evaluating sources (i.e., students), although it does increase the richness of the data and provide a possible check on its reliability. Moreover, the use of students as the *only* source of data for the evaluation of teaching raises many troubling questions, not only concerning reliability and validity, but also about the impact such “consumer satisfaction” evaluations may have upon the faculty member’s perception of her or his role. It would be more accurate, therefore, to count student assessments as only one

source of information no matter how many different forms are used.

When the duplication of student assessment forms is eliminated, the sources of data for assessing instruction are even less varied. Fifty-five percent of the responding departments depend upon two or fewer sources of data. This includes those departments that rely upon student assessments alone (21%) or student assessments combined with only one other source of evaluative data (34%)—most often an infrequent or nonuniform classroom visitation. Of the other responding departments, 28% used three sources of data, 9% used four sources of data, and only 3% used five or more sources of data. Graduate departments had a tendency to use fewer sources of data for teaching evaluation than undergraduate departments. The number of graduate departments relying upon two or fewer sources of data was 65%, while 45% of undergraduate departments were in this category.

**Student Assessment of Instruction**

The most common source of information for evaluating teaching is a written student assessment of instruction, completed near the end of the term. Made up of some combination of Likert-scale items (e.g., how much do you agree or disagree with this statement), other multiple choice items, and open-ended type questions—sometimes accompanied by questions concerning student demographic data—this type of evaluation is used by every responding political science department that employed a formal evaluation process (99 of 100). Three separate types of student evaluation forms were included in our list

of evaluation techniques (see Table 2). Of the three, the most commonly used form was one “specifically created for your college or university” (66%). This was in contrast to a form “specifically constructed for your department” (44%) and “a nationally standardized evaluation form completed by students” (3%). Fourteen of the 100 responding departments (14%) used a combination of more than one of these forms for student assessment of instruction.

Our results also indicate a significant reliance upon “open-ended, narrative-style questions” as a part of these systems of evaluation. Seventy-two percent of the responding departments reported including open-ended questions as a part of their evaluation form. Student evaluations are typically administered at the end of each term. Only seven of the responding departments reported the use of mid-term evaluations (7%), although the number of individual faculty members who use such evaluations for their own benefit outside of the regular departmental evaluation process was not indicated by our data.

Another method of gaining student feedback about instruction is the exit interview with graduating seniors. Thirty-two percent of the responding departments reported using exit interviews for this purpose. While there is no significant difference in the responses of graduate and undergraduate departments in the use of contemporaneous student assessments of instruction as discussed above, the use of exit interviews as a means of evaluating instruction is more common in undergraduate political science departments (39%) than in graduate departments (22%).

**TABLE 2**  
Percentage of Departments Using Selected Student Assessment Methods

Type of Student Assessment	All Departments	Type of Department	
		Undergraduate	Graduate
Nationally Standardized Form	3%	4%	2%
College or University Form	66%	63%	69%
Specific Departmental Form	42%	44%	40%
Narrative-Style Questions	72%	73%	71%
Mid-term Student Evaluations	7%	4%	10%
Exit Interviews	32%	40%	23%

Given the heavy reliance by political science departments upon student assessments of instruction, a more detailed consideration of them is needed. Many of the critiques of this technique are beyond the scope of this paper (for a survey, see Aleamoni 1981). However, as teaching becomes a greater factor in salary, tenure, and promotion decisions, the use of student assessments for summative evaluation purposes<sup>2</sup> presents special problems of reliability and validity (Scriven 1981). Our results provide some information in this regard.

These results indicate that very few political science departments collect student assessment information in a way that facilitates checks of reliability. The reliability of a student assessment instrument is its ability to give relatively consistent results, both within a class and over time. The use of multiple instruments for the student assessment of instruction would be one way to allow cross-checks of reliability. Yet, only 14 of the 100 responding departments (14%) use more than one instrument for student assessment of instruction. However, an examination of the sample forms returned with the questionnaires shows that several of the forms include repetitive items that might be used to check the internal reliability of student responses. It is not possible from our data to determine whether such checks actually take place.

Reliability may also be checked through the multiple administration of the same form at different times in the semester, e.g., in consecutive weeks, with the results of each administration being correlated for each responding student (Aleamoni 1981, 131–2). By comparing results between students within a class, an intraclass correlation coefficient (Winer 1962) may also be calculated as a check of the reliability of the student assessment form. Obviously, such an effort is difficult and, thus, unlikely within individual departments—as is the tracking of reliability over time. However, for departments that use university-level instruments, checks of reliability are more feasible—although

not necessarily more likely to occur. While the reliability of nationally standardized student assessment instruments is the most likely to be regularly examined, our results demonstrate that such forms are, by far, the least likely to be used.

Two final notes on the reliability of student assessments are important. Research has demonstrated that as the number of students participating in the evaluation process and the number of classes evaluated increases, so does the reliability of the results for any given instructor (Marsh 1987). Reliability is also increased by the consistency with which the evaluations are administered and completed. Therefore, those who administer the evaluations (frequently the faculty themselves) and the students who complete them should always be made aware of the importance of such consistency if student assessments are to be relied upon for summative evaluation purposes. An institutional commitment to the enforcement of consistent administration of evaluations should be pursued if reliability is to be promoted (Ory 1990). Moreover, administrators who use student assessments of instruction for summative purposes must be cognizant of all the factors potentially affecting the reliability of such data (Franklin and Theall 1990).

The validity of forms used for student assessment of instruction is related to the form's capacity to measure what it is designed to measure; in this case, the quality of instruction. A threshold issue, therefore, is the definition of the quality of instruction. In a 1987 article, Thomas M. Sherman, with the assistance of four of his graduate students, surveyed the literature to determine the "characteristics of excellence" in college teaching. This survey identified five general characteristics that were consistently identified in the literature as contributing to good teaching (Sherman, et. al. 1987). We used these characteristics—enthusiasm, clarity, preparation and organization, ability to stimulate interest and thinking, and knowledge of the subject matter—to analyze the

**TABLE 3**  
Student Evaluation Forms  
Classified by Characteristics of  
Excellence

Number of Characteristics	Percentage of Forms (Number)
Five	10% (n = 4)
Four	28% (n = 11)
Three	35% (n = 14)
Two	10% (n = 4)
One	8% (n = 3)
Zero	3% (n = 1)

Based upon Sherman's Characteristics of Excellence (Sherman, et. al. 1987).

40 student assessment forms received from responding political science departments. Each form was examined to determine how many of Sherman's characteristics of excellence were assessed by it. The results are shown on Table 3. As the table shows, only four of the 40 forms examined (10%) assessed all five characteristics. Three forms used open-ended questions to elicit general responses that cannot be reliably classified. These results would indicate that based upon our data and Sherman's characteristics of excellence, the forms used by the large majority of political science departments are open to criticism on the grounds that they fail to measure the generally identified characteristics of excellent teaching. Moreover, a lack of clarity in question construction is likely to exacerbate the problem.

Despite these shortcomings, however, most of the forms that were examined did assess *some* characteristics of teaching effectiveness. They were more likely to examine basic classroom issues than Sherman's more elevated characteristics of excellence. For example, 14 of the forms sought information concerning the effectiveness of texts and other teaching resources. They also assessed details of classroom management, such as punctuality and attendance, classroom time management, meeting office hours, and availability to students generally.

Another important use of the forms was to assess instructor attitudes toward students. An item concerning the instructor's interest

in students' opinions and in whether they improved their knowledge of the subject was the most frequently included item (33/40; 83%). The only item asked as frequently was one that asked for an overall assessment of the instructor—either in absolute terms or compared to a given population (e.g., instructors you have had in this university) (33/40; 83%). Thus, the forms give substantial emphasis to fundamental issues of teaching, with less attention for identified measures of "excellence."

Most political science departments in the survey appeared to use student assessments of instruction to head off potential problems, rather than to improve instruction in all of its dimensions. This point is reinforced by the responses to the open-ended question concerning how teaching evaluations are used to improve the quality of instruction. The most common response was that the chair reviewed the evaluations and discussed them with faculty members (24/100; 24%). Sixteen of the responding department chairs identified some form of problem intervention as their only use of evaluations for development purposes. Two examples are typical: "We use evaluations to weed out bad instructors (particularly part-time instructors)"; and, "Teaching evaluations sometimes call attention to truly bad performances." With such a limited use of student assessments, validity is more likely.

It is also important to note that two forms included an item concerning the instructor's sensitivity to issues of diversity and gender. If problem intervention is a primary goal of student assessments of instruction, then this kind of question on an anonymous evaluation form could be quite useful. However, using student assessments in this manner leads to the most common worry of faculty concerning the validity of student evaluations of instruction—bias.

Much has been written on the subject of bias in student assessments of instruction. In this context, bias is the impact upon evaluation results of noninstructional

factors—such as the characteristics of the student evaluators or of the class and its setting. John Centra has recently discussed the subject of bias (Centra 1993, 60–78) presenting an excellent summary of the current state of the literature. His conclusions are enlightening:

Although the research is largely supportive of student evaluations, they have been shown to have limitations. Student evaluations do differ somewhat for some course and instructor characteristics, including class size, method of teaching, academic field, prior student interest in the subject, grading leniency, and teacher expressiveness or "seductiveness" . . . [However, e]ven several biasing factors considered together account for only about 15 percent of the variance in student ratings (Marsh 1987). Although such variance is not excessive, it could alter a close personnel decision. In tenure and promotion decisions, consideration of the ratings of several different courses over several years minimizes the possibility of bias (Centra 1993, 78).

---

*Most political science departments in the survey appeared to use student assessments of instruction to head off potential problems, rather than to improve instruction in all of its dimensions.*

---

The potential for bias, while a small problem, must always be considered when using student assessments of instruction. To this end, 21 of the 40 forms that were reviewed (53%) collected data on some or all of the "course and instructor characteristics" that Centra identifies.

#### Classroom Visitation

The second most common source of data for the evaluation of in-

struction is the classroom visitation. Our results indicate that 54% of the responding departments use this technique, though the frequency and character of visitations varies widely. Of departments that use classroom visitation, only 10 (17%) report doing so biannually (or every semester). Seventeen departments hold annual visitations (29%), with the remainder conducting visitations even less frequently. Often, visitations *only* occur as a part of the process of review for tenure or promotion. *After* tenure has been granted, visitations become infrequent or largely nonexistent. For example, several departments conducted visitations of tenured faculty every five years.

Classroom visitations also varied widely in who conducted them. Visitations were most often conducted by colleagues (excluding the department chairperson or head; 24/54; 44%). Department chairpersons conducted the visitations in 24% (13/54) of departments, while a college dean did so in 6% (3/54) of the departments. In 12 (22%) departments, a committee made up of a combination of people—such as other faculty, administrative officials, and/or students—conducted the visitation.

Classroom visitations as described in our survey (insofar as our results reflect any consistent procedure) raise several questions. Given the fact that 23% of the departments that use classroom visitations disclose a direct use of them for summative purposes,<sup>3</sup> the same problems of reliability and validity occur as with student assessments of instruction. In the most comprehensive study of this subject, Centra (1979, 75) reported that such ratings are statistically unreliable. His results indicated a correlation among the ratings of different colleagues of only .26 for each item used in his research. He also found colleague ratings to be consistently more favorable than student ratings, which had fewer reliability problems when properly administered. While some of the unreliability of classroom visitation can be rectified by team visitation (Miller 1987, 77; Hennessy 1975,

## The Teacher

12), our results indicate that a small number of departments follow such a procedure (12/54; 22%).

These results indicate serious problems with the use of classroom visitations for summative evaluation purposes. However, several formative evaluative uses of classroom visitations are also indicated by our data.<sup>4</sup> These formative evaluative uses of classroom visitation do not suffer the same limitations as using visitations for summative evaluation purposes. Colleague classroom visitation occurs as a part of a peer counseling or mentoring program in 20% (20/100) of the responding departments (or 37% of all departments that use classroom visitations). These programs range from a colleague visitation followed by an informal discussion of the visit to the formal assignment of a senior faculty member as mentor for an entering junior faculty member until tenure review, with regular classroom visitations as a part of the mentoring process.

Other departments report encouraging junior faculty to visit the classrooms of identified "master teachers" followed with discussions of instructional methods. One department reports using "visits by all members of the department to other members' classes. Through these visits we're exposed to a wide range of teaching styles and receive formal critiques of our in-class performance." When used in this kind of program, classroom visitation contributes to instructional development in a positive manner. However, a commitment to programs of this type are necessary from both junior and senior faculty if they are to be successful (Centra 1993, 121–3).

### Other Evaluation Techniques

Responses to the questionnaire show that departments are using several other techniques for gathering data to evaluate instruction. The most common of these is best described as "formally structured self-evaluations prepared by the instructor." This group of techniques includes teaching portfolios,

**TABLE 4**  
Percentage of Departments Using Selected Teaching Development Methods

Type of Development Activity	All Departments	Type of Department	
		Undergraduate	Graduate
Videotaping	26%	25%	27%
Presentations or Seminars	46%	54%	38%
Open Discussion Groups	48%	48%	48%
Other Methods	29%	23%	35%

written responses to student assessments, and end-of-the-year self-assessments to identify strengths and weaknesses. Often self-evaluations include the evaluation by the faculty member of areas other than instruction, as well (e.g., scholarship and service). Twenty-one percent (21/100) of the responding departments reported the use of some type of formally structured self-evaluation as a part of their evaluation system. The most common form of self-evaluation involved some kind of teaching portfolio including, among other items, copies of syllabi, goals for classes taught, and assessments of success. Formally structured self-evaluations can contribute to both summative and formative evaluations. In fact, research has shown that using self-evaluations as a tool for summative evaluation increases the care with which the self-evaluations are prepared, thus, increasing their benefit in instructional development (Centra 1993, 100–2).

Another method for gathering data for instructional evaluation is outcome-based evaluation. This form of evaluation is dependent upon the measurement of how many students achieve a predetermined set of standards or skills. Although used by a small number of departments (9/100; 9%), this technique was under study in several others. While a 1991 survey of outcomes assessment in political science noted that "fewer than one out of six responding institutions" formally collected "certain types of outcomes assessment data" (Julian, Chamberlain, and Seay 1991, 206–7), our results indicate that fewer still have programs in place in 1994.

### Mechanisms for Instructional Development

Evaluation is only one part of improving instruction in political science. For the feedback of formative evaluation to be effective, more than just incentives or threats of including instructional quality in the tenure and promotion process must be present. Mechanisms must be in place to assist those who seek to improve their teaching (Scriven 1981, 247). Moreover, these mechanisms are best included in a "flexible mix of improvement activities" (Weimer 1990, 82–110).

The survey responses to our items on instructional development reveal little diversity in the mix of methods used. A majority of the responding departments (54/100; 54%) reported providing either no methods of instructional development (19/100; 19%) or only a single method (35/100; 35%). Only 16 departments (16%) reported three or more mechanisms for improving instruction. Such results, rather than indicating a "flexible mix of improvement activities," indicate virtually nothing to be mixed.

The most common mechanism for the development of instruction was the "open discussion of instruction" (i.e., brown-bag-style discussion group). Open-discussion groups were used by 48 departments (48%). The next most common mechanism for instructional development was the "seminar on teaching" (46/100; 46%), followed by the "video-taping of classes" (26/100; 26%). Other responses included discussions at faculty meetings, summer workshops, and paid trips to professional meetings. Faculty participation in virtually all development activities was voluntary. Only five departments (5%)

required any development activities of teaching faculty. Other responding chairs noted referring junior faculty with low student evaluations to a college center for instructional development. Such college-level centers were noted by eight respondents with two others noting that centers were in the process of being developed in their institutions.

## Conclusions

The survey results indicate that political science departments are aware of the importance of quality teaching. With variation from department to department, the majority of departments give significant weight to teaching in their personnel decisions. Moreover, undergraduate departments give more weight to teaching than do graduate departments. Most departments have clearly defined systems of summative teaching evaluation. However, these systems generally lack diversity in the sources of data used to evaluate teaching, depending primarily upon student assessments. When such assessments are relied upon for summative evaluation purposes, particular emphasis must be given to the reliability and validity of the system used for the assessment of instruction. These assessments are used primarily as a means to identify potential problems in the classroom and, therefore, tend to measure the lowest-level characteristics of teaching quality.

The use of the evaluation systems for formative evaluation purposes is more encouraging. The widespread emphasis upon narrative-style questions in student assessments of instructional quality facilitates the use of these assessments to improve teaching. Moreover, the survey indicates that a number of departments are using systems of peer evaluations and mentoring for the development of instruction.

Despite these positive indications, the mechanisms designed to help faculty who wish to improve their teaching are generally insufficient. As political science depart-

ments increase their emphasis on quality instruction as a part of tenure and promotion, it is reasonable for faculty to expect that adequate provision be made to provide them with opportunities to improve their teaching abilities in response. Moreover, in an environment in which schools are increasingly competing for students, students can be expected to press for even more improvements in the instruction that they receive—with commensurate economic and political pressures for departments to do so. While this survey indicates sustained effort by political science departments to improve the quality of instruction (typified by the fact that 44% of the departments reported having formally reexamined their systems of evaluation and development of instruction within the last year), it also demonstrates how far we have to go.

## Notes

1. The authors would like to thank the staff of the Virginia Commonwealth University Survey Research Laboratory and the VCU Committee on Roles and Rewards for their assistance in the completion of this project.

2. We use the common terminology presented in the teaching evaluation literature. A summative evaluation is one designed and used as a tool for administrative decision making on matters such as hiring and firing, promotion, salary, or tenure. Formative evaluation, on the other hand, serves as a tool for instructional development, acting as a channel of feedback to faculty members for the improvement of their classroom performance.

Of course, these evaluative functions are not mutually exclusive. For reasons of convenience, departments often use a single evaluation technique or instrument for both summative and formative purposes. Moreover, a faculty member's regard for improving her or his summative evaluations often serves as the incentive to improve instruction—thus encouraging the formative use of the same evaluation process.

3. The actual number may be much higher. This number represents only those departments that volunteered this information in the open-ended questions or in the question concerning how often visitations occur.

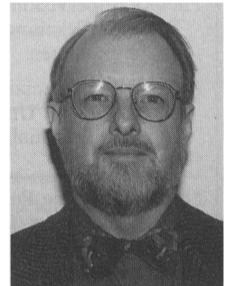
4. For an explanation of summative and formative evaluations see note 2.

## References

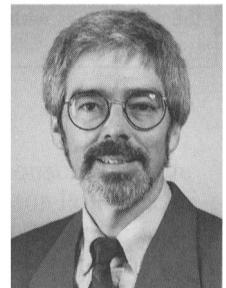
- Aleamoni, Lawrence M. 1981. "Student Ratings of Instruction." In *Handbook of Teacher Evaluation*, ed. Jason Millman. Beverly Hills: Sage Publications.
- Centra, John A. 1993. *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*. San Francisco: Jossey-Bass Publishers.
- Centra, John A. 1979. *Determining Faculty Effectiveness: Assessing Teaching, Research, and Service for Personnel Decisions and Improvement*. San Francisco: Jossey-Bass.
- Erickson, Glenn R., and Bette L. Erickson. 1979. "Improving College Teaching: An Evaluation of a Teaching Consultation Procedure." *Journal of Higher Education* 50:670-683.
- Ferren, Ann S., and Louise G. White. 1977. "Models for Faculty Development." *Teaching Political Science* 5:23-38.
- Franklin, Jennifer, and Michael Theall. 1990. "Communicating Student Ratings to Decision Makers: Design for Good Practice." In *Student Ratings of Instruction: New Directions for Teaching and Learning*, No. 43., ed. Michael Theall and Jennifer Franklin. San Francisco: Jossey-Bass Publishers, 1990.
- Hennessy, Bernard. 1975. "The Evaluation of Teaching: A Collegial Enterprise." *Teaching Political Science* 3:3-17.
- Julian, F., Don Chamberlain, and Robert Seay. 1991. "A National Status Report on Outcomes Assessments by Depart-

## About the Authors

**J. Clifford Fox** is a senior research associate of the Center for Public Policy's Survey Research Laboratory at Virginia Commonwealth University. He also teaches classes in American politics and public policy in the Department of Political Science and Public Administration.



**Scott Keeter** is the chair of the Department of Political Science and Public Administration at Virginia Commonwealth University.



**The Teacher**

ments of Political Science." *PS: Political Science and Politics* 24(2):205-8.

Marsh, H. W. 1987. "Student Evaluations of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research." *International Journal of Educational Research* 11:253-388.

Miller, Richard I. 1987. *Evaluating Faculty for Promotion and Tenure*. San Francisco: Jossey-Bass Publishers.

Ory, John C. 1990. "Student Ratings of Instruction: Ethics and Practice." In *Student Ratings of Instruction*. New Directions for Teaching and Learning, no. 43, ed. Michael Theall and Jennifer Franklin. San Francisco: Jossey-Bass Publishers.

Scriven, Michael. 1981. "Summative Teacher Evaluation." In *Handbook of Teacher Evaluation*, ed. Jason Millman. Beverly Hills: Sage Publications.

Sherman, Thomas M., et al. 1987. "The Quest for Excellence in University Teaching." *Journal of Higher Education*, 58:66-84.

Weimer, Maryellen. 1990. *Improving College Teaching: Strategies for Developing Instructional Effectiveness*. San Francisco: Jossey-Bass Publishers.

Winer, B. J. 1962. *Statistical Principles of Experimental Design*. New York: McGraw-Hill.

**Political Science Programs To Prepare Ph.D. Students To Teach Undergraduates**

Teacher Preparation Formats	First Quartile Ranked Departments*		Second Quartile Ranked Departments*		All Other Departments*	
	Required for Ph.D. Students	An Option for Ph.D. Students	Required for Ph.D. Students	An Option for Ph.D. Students	Required for Ph.D. Students	An Option for Ph.D. Students
A Course or Seminar Offered by the Department to Prepare Students to Teach Undergraduates	36%	16%	32%	14%	24%	13%
A Teaching Assistantship with Experience Teaching Undergraduates	56%	40%	32%	55%	29%	50%
Faculty Supervision of Graduate Students Teaching Undergraduates	64%	28%	55%	18%	55%	29%
A Course or Special Presentation Offered by the University Devoted to Training Teaching Assistants**	24%	—	45%	—	8%	15%

\*Rankings of the Political Science Departments are from *Research-Doctorate Programs in the United States: Continuity and Change*, Marvin L. Goldberger, Brendan A. Maher, and Pamela Ebert Flattau, Eds. Washington, D.C.: National Academy Press, 1995, Table M, pp. 366-69. Ninety-eight departments are ranked. All 25 departments in the first quartile responded to this survey and 22 of the 24 departments (92%) in the second quartile responded. Of the 72 other departments, 62 (86%) responded to the survey. Data are reported from among responding departments only. Departments were asked to indicate if they are considering increasing the teaching preparation given to graduate students. Fifty-two percent of the departments ranked in the first quartile, 63% in the second quartile, and 52% of the other departments answered affirmatively.

\*\*These university courses are available in 60%, 45% and 47% of the departments ranked in the top quartile, second quartile, and in all of the universities with Ph.D. programs.