
Are Student Evaluations A Valid Measure of Teaching Effectiveness: Perceptions of Accounting Faculty Members and Administrators

by *Donald Ace Morgan, John Sneed and Laurie Swinney*

Abstract

This article examines the perceptions of both administrators and faculty relating to the validity of student evaluations and the existence of possible biasing factors. The results indicate that administrators believe student evaluations measure teaching effectiveness to a greater degree than faculty, while faculty members believe their personality is the primary determinant of ratings on student evaluations. Faculty also believe that the type of course, the work load of a course, and the grade distribution of a course have a larger impact on student evaluations when compared with administrators' beliefs.

Introduction

Evaluating a faculty member's teaching ability is one of the most difficult and contentious tasks faced by accounting administrators. While teaching ability is one of the primary factors included in promotion and tenure decisions, there is little agreement on how teaching effectiveness should be measured. The ultimate result of effective teaching is student learning, where students master the content of specific courses. However, comparing student learning across courses and across disciplines is difficult. Since faculty members in accounting departments teach different courses, evaluating teaching effectiveness based on student learning is exacerbated. Many faculty members often argue the impossibility of quantitatively measuring effective teaching, yet they feel certain that they are effective teachers (Simpson, 1995). If teaching ability is to be used as a primary factor in teaching and promotion decisions, administrators must be able to validly measure effective teaching.

Student evaluations are the primary tool used by accounting administrators to evaluate teaching effectiveness (Yunker & Sterner, 1988). Most accounting administrators believe students can reliably evaluate teaching effectiveness, although they suggest using supplemental information to control for potential bias in the student evaluation process. Prior research in the education literature, however, finds that many faculty believe that student evaluations are simply a popularity contest and have no relation to effective teaching. Faculty also argue that factors beyond their control, such as course difficulty and class size can impact student ratings of faculty. If ad-

Biographical Notes

Donald Ace Morgan, University of Nebraska at Kearney, Professor of Accounting and Director of the MBA Program, College of Business & Technology, Department of Accounting and Finance, Kearney, NE 68849, USA.

John Sneed, Jacksonville State University, Professor of Accounting, College of Commerce & Business Administration, Department of Finance, Economics and Accounting, Jacksonville, AL 36265, USA.

Laurie Swinney, University of Nebraska at Kearney, Associate Professor of Accounting, College of Business & Technology, Department of Accounting and Finance, Kearney, NE 68849, USA.

Are Student Evaluations a Valid Measure of Teaching Effectiveness

ministrators are to use student evaluations as a primary measure of teaching effectiveness, faculty members should agree that student evaluations are a valid measure. The purpose of this article is to determine if accounting faculty and accounting administrators agree with what is actually being measured by student evaluations. This article should be helpful to administrators in interpreting student evaluations and in determining their appropriate use in the evaluation process. This article should also be helpful to accounting faculty in understanding what other factors could be driving the student ratings that they receive in their courses.

Prior Research

Faculty often argue that teaching effectiveness is difficult to identify and nearly impossible to validly measure, so individual faculty members should be allowed to use their subjective judgment to determine how to conduct their classes (Simpson, 1995). However, since teaching effectiveness is one of the primary factors used in promotion and tenure decisions, faculty members and administrators must find agreement on a valid method to evaluate teaching ability. Student evaluations have become the primary tool used by administrators to evaluate the teaching effectiveness of their faculty (Seldin, 1993; Yunker and Sterner, 1988).

Simpson (1995) found that student evaluations were the most consistent and most controversial source of information used to evaluate teaching effectiveness. Across time, student evaluations have changed from a voluntary tool used by individual faculty members to improve their teaching skills into a required process for all faculty, on which administrators rely to measure teaching effectiveness for promotion and tenure decisions (Stratton, Myers, and King, 1994). The increased reliance on student evaluations in promotion and tenure decisions has raised questions about whether student feedback is a valid measure of effective teaching, with many faculty arguing that there are many potential biases in the use of student evaluations (Anderson and Miller, 1997).

Substantial research exists in the education literature arguing that student evaluations are a valid and reliable measure of teaching effectiveness. Many studies argue that a properly designed student evaluation can provide reliable information about many aspects related to effective teaching (Cohen 1981; Marsh, 1984; Marsh and Ware, 1982; Marsh, 1987; Murray, 1983; Seldin, 1984). Timpson and Andrews (1997) argue that there is strong agreement in the literature that student evaluations are a valid and reliable tool when used to evaluate effective teaching.

Studies addressing faculty perceptions of the use of student evaluations, however, find that faculty members frequently question the validity of student evaluations as a tool to measure effective teaching (Kemp and Kumar, 1990; Marsh, 1987; Marsh, 1982; White and Ahmidi, 1986). Crader and Butler Jr. (1996) argue that teaching ability is not an important variable af-

fecting student evaluations and that faculty have less control over student ratings than is commonly believed. Faculty members also argue that student evaluations are simply a popularity contest and are not related to teaching effectiveness (Yunker and Sterner, 1988). Simpson (1995) found that many faculty members believe that students are unable to evaluate effective teaching and should not be allowed to do so.

In addition to believing that student evaluations simply represent a popularity contest, as opposed to being valid measures of teaching ability, faculty members also believe that ratings on student evaluations are negatively correlated with class rigour, workload, and difficulty. Many faculty members argue that they are unable to teach an intellectually challenging course and, at the same time, provide students with what is needed for higher evaluations, higher grades for less work (Simpson, 1995). Marsh and Overall (1981) found that 75 per cent of faculty believe that course difficulty has a negative impact on student evaluations while 70 per cent believe grade leniency has a positive impact on ratings on student evaluations. While research findings have documented a positive relationship between course difficulty and ratings on student evaluations (Marsh, 1987; 1984; 1982), faculty members do not believe the findings to be true.

Faculty members believe they can improve their ratings on student evaluations by lowering the standards in their courses, or giving higher grades for less work. Thus, academic administrators should realise that student evaluations are not simply a passive tool used to measure teaching effectiveness, but can result in faculty members changing their behaviour in an attempt to improve their ratings (Stratton *et.al.*, 1994). If faculty members lower course standards, assigning higher grades for less work, teaching effectiveness and student learning will be weakened. Calderon, Green, and Reider (1997) argue that if accounting faculty members believe they can improve their ratings on student evaluations by reducing the rigour in their courses, they will lower the standards of their courses. Faculty also argue that factors beyond their control (class size, student motivation) can impact ratings on student evaluations. Mason *et.l.*, (1996) argue that, if administrators do not adjust student evaluation ratings for differences in class size and course difficulty, faculty members will change their behaviour to improve their results.

While there is substantial research indicating that student evaluations are valid measures of teaching effectiveness, and that faculty members do not accept this finding, limited research has addressed the issue of administrator perceptions regarding the validity of student evaluations. Calderon and Green (1997) found that a large majority of accounting department chairs believe student evaluations are a valid tool that can be used to assess effective teaching. This is not a surprising result since they also report that 95 per cent of accounting departments are now using student evaluations to assess teaching ability.

**Are Student
Evaluations a
Valid Measure
of Teaching
Effectiveness**

Are Student Evaluations a Valid Measure of Teaching Effectiveness

Calderon *et.al.*, (1997) found that most accounting department chairs believe students are able to reliably evaluate effective teaching and that student evaluations are a valid tool for measuring teaching ability. However, the majority of the department chairs recognised that student evaluations could be biased by additional factors including course difficulty, the actual grade distribution, and the size of the class. The administrators indicated that they attempt to control for these factors by examining additional information beyond student evaluations when assessing a faculty member's teaching ability. One interesting finding is that over 70 per cent of the accounting administrators believe student evaluations are not useful in measuring student learning, the ultimate result of effective teaching. This finding suggests that accounting administrators believe that students can competently assess whether a faculty member is an effective teacher, but they cannot assess whether they have learned anything in a class.

The sizable gap between administrator views of student evaluations and faculty member views indicates that accounting administrators must convince faculty that their assessment of a faculty member's teaching effectiveness is not based only on the results of student evaluations. Calderon *et.al.*, (1997) argue that further research is needed to determine the extent of the differences between views of faculty members and those of administrators regarding what is actually being measured by student evaluations. If administrators fail to convince faculty that student evaluations are a valid measure of teaching effectiveness and that they adjust for the potential biases in the student evaluation process, the best result would be that faculty members simply would not accept a chair's assessment as a valid measure of their teaching effectiveness. The worst result would be if faculty changed their behaviour in an attempt to manage their ratings on student evaluations, lowering standards and assigning higher grades, which will weaken both teaching effectiveness and student learning. This seems to be a likely occurrence as accounting administrators do not seem to equate effective teaching with student learning.

Methodology

This article addresses the following research questions:

1. Do accounting faculty and administrators perceive student evaluations to be a valid unbiased measure of teaching effectiveness?
2. Do accounting faculty and administrators perceive that student evaluations are influenced by factors other than teaching effectiveness?
3. Do the perceptions of accounting administrators agree with those of accounting faculty?

4. Do the perceptions of faculty vary across rank?

By examining these questions, we will be able to determine if faculty and administrators believe that student evaluations are valid unbiased measures of teaching effectiveness. If faculty and administrators do not believe that the student evaluations are valid measures, then we will be able to determine possible sources of bias such as type of course, work load of the course, grades given in the course, and/or instructor's personality. We will also be able to determine if perceptions of student evaluation validity change as individuals move into administrative positions and/or as individuals move upward in rank.

We conducted a survey of all accounting departments at U.S. universities listed in the *Accounting Faculty Directory* (Hasselback, 1996). For each department, we included the dean, the department chair, and one faculty member at each of the three ranks (full, associate, and assistant). Some departments did not have a faculty member at all three ranks, so they received less than their allotted five surveys. This sampling procedure resulted in a total of 1,676 surveys in the initial mailing. We used two mailings to receive what we perceived to be an adequate response rate for our analysis.

A total of 583 responses were returned, for a response rate of 35 per cent. Of those returned, 532 responses (94%) were from individuals whose departments use student evaluations to evaluate faculty teaching. Table 1 reports the classification of individuals according to position and rank.

Position	n =	%	Rank	n =	%
Faculty	320	60.2	Assistant Professor	75	14.1
Chairs	130	24.4	Associate Professor	176	33.1
Deans	74	13.9	Professor	275	51.7
Missing	8	1.5	Missing	6	1.1

To address the research questions, respondents were asked to complete the following:

Do you feel the responses on student evaluations are:

**Are Student
Evaluations a
Valid Measure
of Teaching
Effectiveness**

Are Student Evaluations a Valid Measure of Teaching Effectiveness

	Always	Usually	Sometimes	Never
An accurate indication of a teacher's effectiveness?				
Indicative of the type of course rather than the teacher (tax v cost)?				
Indicative of the workload of the course rather than the teacher?				
More indicative of the teacher's personality than effectiveness?				
Correlated to grades given to students in a course?				

The respondents were also asked to indicate their rank and whether they were a faculty member, a department chair, or a dean.

Results

The frequency and percentage of responses made by all individuals is reported in Table 2. Table 2 reports that the greatest number of respondents believe that student evaluations are *usually* an accurate indication of a teacher's effectiveness. The second largest group, however, those who believed that the evaluations are only *sometimes* an accurate indication of effectiveness, followed close behind. In addition, the largest group of respondents indicated that all four of the biases occur at least *sometimes*. 45% of the respondents indicated that the evaluations were *usually* or *always* more indicative of the teacher's personality than effectiveness.

Table 2 Frequency and Percentage Responses		
	Frequency	Percentage
An accurate indication of a teacher's effectiveness?		
Always	8	1.5
Usually	259	48.7
Sometimes	244	45.9
Never	12	2.3
Missing	9	1.7
Indicative of the type of course rather than the teacher (tax v cost)?		
Always	15	2.8

**Are Student
Evaluations a
Valid Measure
of Teaching
Effectiveness**

Usually	145	27.3
Sometimes	336	63.2
Never	27	5.1
Missing	9	1.7
Indicative of the workload of the course rather than the teacher?		
Always	26	4.9
Usually	147	27.6
Sometimes	324	60.9
Never	26	4.9
Missing	9	1.7
More indicative of the teacher's personality than effectiveness?		
Always	50	9.4
Usually	196	36.8
Sometimes	264	49.6
Never	12	36.8
Missing	10	1.9
Correlated to grades given to students in a course?		
Always	40	7.5
Usually	160	30.1
Sometimes	287	53.9
Never	38	7.1
Missing	7	1.3

Perceptions by Position

The frequency and percentage of responses made by individuals according to position is reported in Table 3. Table 3 reports that the greatest number of administrators believe that student evaluations are *usually* an accurate indication of a teacher's effectiveness. In contrast, the greatest number of faculty believe that student evaluations only *sometimes* indicate a teacher's effectiveness. Further, in the case of every bias investigated, the total per-

Are Student Evaluations a Valid Measure of Teaching Effectiveness

percentage of faculty responding *usually* or *always* was higher than the total percentage of administrators responding *usually* or *always*.

Table 3				
Frequency and Percentage Responses by Position				
	Frequency		Percentage	
	Administrators	Faculty	Administrators	Faculty
An accurate indication of a teacher's effectiveness?				
Always	7	1	3.4	.3
Usually	111	147	54.4	45.9
Sometimes	82	157	40.2	49.1
Missing	3	8	1.5	2.5
Never	1	7	.5	2.2
Indicative of the type of course rather than the teacher (tax v cost)?				
Always	4	11	2.0	3.4
Usually	36	107	17.6	33.4
Sometimes	151	181	74.0	56.6
Missing	13	13	6.4	4.1
Never	0	8	0.0	2.5
Indicative of the workload of the course rather than the teacher?				
Always	5	21	2.5	6.6
Usually	46	100	22.5	31.3
Sometimes	139	180	68.1	56.3
Missing	13	12	6.4	3.8
Never	1	7	.5	2.2
More indicative of the teacher's personality than effectiveness?				
Always	15	35	7.3	10.9
Usually	64	130	31.4	40.6
Sometimes	115	146	56.4	45.6
Missing	8	2	3.9	.6

Are Student Evaluations a Valid Measure of Teaching Effectiveness

Never	2	7	1.0	2.2
Correlated to grades given to students in a course?				
Always	10	30	4.9	9.4
Usually	50	108	24.5	33.8
Sometimes	123	160	60.3	50.0
Missing	21	16	10.3	5.0
Never	0	6	0.0	1.9

We used t-tests to determine whether the perceptions of administrators were significantly different than the perception of faculty members. To analyse the data we coded the responses as follows: always (3), usually (2), sometimes (1), and never (0). Thus, a higher mean score for a group indicates a higher level of agreement for the statement investigated. The mean scores for each group and statement are reported in Table 4.

	Means		Sig of F
	Administrators	Faculty	
An accurate indication of a teacher's effectiveness?	1.60	1.45	.004
Indicative of the type of course rather than the teacher (tax v cost)?	1.15	1.37	.000
Indicative of the workload of the course rather than the teacher?	1.21	1.42	.000
More indicative of the teacher's personality than effectiveness?	1.43	1.63	.001
Correlated to grades given to students in a course?	1.24	1.48	.000

The statistical analysis showed that the administrator's perception is significantly greater ($p=.004$) than the faculty's perception that student evaluations are an accurate indication of teacher's effectiveness. Further, the perception that factors exist which bias the student evaluation is significantly stronger ($p<.001$) for the faculty group than the administrator group in all biases investigated. These results indicate that although administrators believe that student evaluations are good indicators of teacher effectiveness, faculty do not share this belief. Further, faculty recognise that factors exist which bias the evaluations to a greater degree than their administrative counterparts.

Are Student Evaluations a Valid Measure of Teaching Effectiveness

Perceptions by Rank

The frequency and percentage of responses made by faculty who were not administrators is reported in Table 5 according to rank. Table 5 reports that at all three ranks, the largest group believed that student evaluations are only *sometimes* an accurate indication of a teacher’s effectiveness. The second largest group, however, those who believed that the evaluations are *usually* an accurate indication of effectiveness, was close behind at all three ranks. In addition, for two of the four biases investigated, the total percentage of professors responding *usually* or *always* was lower than the total percentage of associate professors or assistant professors responding *usually* or *always*. Professors are less convinced than associates or assistants that student evaluation responses are indicative of the type of course and that student evaluation responses are indicative of the workload of the course. Professors, however, are more convinced than associates or assistants that student evaluation responses are indicative of the teacher’s personality and that the responses are correlated to grades given to students in the course.

Table 5 Frequency and Percentage Responses of Faculty by Rank						
	Frequency			Percentage		
	Prof	Asso	Asst	Prof	Asso	Asst
An accurate indication of a teacher’s effectiveness?						
Always	1	0	0	.8	0	0
Usually	54	59	33	43.2	47.2	47.8
Sometimes	63	60	34	50.4	48.0	49.3
Never	4	3	1	3.2	2.4	1.4
Missing	3	3	1	2.4	2.4	1.4
Indicative of the type of course rather than the teacher (tax v cost)?						
Always	6	4	1	4.8	3.2	1.4
Usually	33	49	25	26.4	39.2	36.2
Sometimes	79	60	41	63.2	48.0	59.4
Never	4	8	1	3.2	6.4	1.4
Missing	3	4	1	2.4	3.2	1.4
Indicative of the workload of the course rather than the teacher?						
Always	11	7	3	8.8	5.6	4.3

Are Student Evaluations a Valid Measure of Teaching Effectiveness

Usually	31	41	27	24.8	32.8	39.1
Sometimes	75	69	36	60.0	55.2	52.2
Never	5	5	2	4.0	4.0	2.9
Missing	3	3	1	2.4	2.4	1.4
More indicative of the teacher's personality than effectiveness?						
Always	14	14	7	11.2	11.2	10.1
Usually	53	47	29	42.4	37.6	42.0
Sometimes	56	58	32	44.8	46.4	46.4
Never	0	2	0	0.0	1.6	0.0
Missing	2	4	1	1.8	3.2	1.4
Correlated to grades given to students in a course?						
Always	14	12	4	11.2	9.6	5.8
Usually	47	41	19	37.6	32.8	27.5
Sometimes	59	62	39	47.2	49.6	56.5
Never	4	7	5	3.2	5.6	7.2
Missing	1	3	2	.8	2.4	2.9

We used oneway analysis of variance to determine whether the differences in perceptions between professors, associate professors, and assistant professors were significant. Once again, a higher mean score for a group indicates a higher level of agreement for the statement investigated. The mean scores for each group and statement are reported in Table 6.

	Means			Sig of F
	Professor	Associate Professor	Assistant Professor	
An accurate indication of a teacher's effectiveness?	1.43	1.46	1.47	.840
Indicative of the type of course rather than the teacher (tax v cost)?	1.34	1.41	1.38	.685
Indicative of the workload of the course rather than the teacher?	1.39	1.41	1.46	.828

Are Student Evaluations a Valid Measure of Teaching Effectiveness

More indicative of the teacher's personality than effectiveness?	1.66	1.60	1.63	.822
Correlated to grades given to students in a course?	1.57	1.48	1.33	.092

The statistical analysis showed that no significant differences existed between any of the three groups relating to the respondent's perception that student evaluations are an accurate indication of a teacher's effectiveness. Further, there were no significant differences in perceptions that factors exist which bias the student evaluation between any of the three groups in four out of the four possible biases investigated. These results indicate that an individual's perceptions relating to the value of student evaluations and possible biases do not change as the individual moves upward in rank. Apparently, faculty recognise early in their academic careers the pitfalls of the student evaluation process and their views do not change as they gain experience in the classroom.

Conclusion

The purpose of this article was to empirically investigate the perceptions of both administrators and faculty relating to the validity of student evaluations and the existence of possible biasing factors. Prior research describes student evaluations as the most consistent and controversial source of information used to evaluate teaching effectiveness (Simpson, 1995). Further, measuring teaching effectiveness is crucial not only for instructor feedback on their own performance but is also vital for course/curriculum design and modification, for facilitation of administrative decision making, and for provision of information for academic counselling (Marsh, 1982). The use of student evaluations by 95% of respondents surveyed by Calderon and Green (1997) speak to the administrators' perception of the validity of the instrument. The results of Calderon and Green are consistent with Timpson and Andrew's (1997) statement that there is strong agreement in the literature that student evaluations are a valid and reliable tool when used to evaluate the effectiveness of teaching.

Other research, however, has examined the perceptions of faculty relating to the use of student evaluations as a measure of teaching effectiveness. Faculty members frequently question the validity of student evaluations (Kemp and Kumar, 1990). Faculty members believe that student evaluations are influenced by personality (Yunker and Sterner, 1988), by the workload of the course (Simpson, 1995), the leniency of grading (Marsh and Overall, 1981), and the type of course taught (Morgan *et.al.*, 1998). Calderon and Green (1997) call for further research to determine the extent of the differences between the views of faculty members and those of administrators regarding the validity of student evaluations and possible biasing factors.

**Are Student
Evaluations a
Valid Measure
of Teaching
Effectiveness**

The results of our research are consistent with the results of prior research. We also found that the largest group of respondents (48.7%) believed that student evaluations are an accurate indication of a teacher's effectiveness. The second largest group (45.9%), however, believed that student evaluations are only sometimes an accurate indication of a teacher's effectiveness. We also found perceptions of the existence of biasing factors. A large percentage of our respondents believe that evaluations are at least sometimes biased by type of course (63.2%), workload of the course (63.2%), teacher's personality (49.6%), and grades given in the course (53.9%). Approximately a fourth of our respondents made a stronger statement by indicating that these factors usually bias the student evaluations.

Prior studies did not compare the perceptions of faculty and administrators relating to the validity of student evaluations and the potential existence of biasing factors. Our research showed that while administrators believe that student evaluations are usually a valid indicator of teacher effectiveness, faculty believe that student evaluations are only sometimes a valid indicator of teacher effectiveness. The views of administrators and faculty were significantly different. Further, faculty responded with stronger agreement that student evaluations are indicative of the biasing factors discussed above. Once again, in the case of all four biasing factors, the views of administrators were significantly different than the views of faculty.

Further, prior studies did not compare the perceptions of faculty at different ranks relating to the validity of student evaluations and the potential existence of biasing factors. Our research showed that although the percentage of faculty who believe that student evaluations are usually an accurate indication of a teacher's effectiveness decreased slightly as individuals moved upward in rank, the views of faculty at different ranks were not significantly different. The results were mixed for perceptions of the existence of biasing factors. The total percentage of professors responding usually or always was lower than the total percentage of associate professors or assistant professors responding usually or always for only two out of the four biases investigated. Professors are more convinced than associates or assistants that student evaluation responses are indicative of the teacher's personality and that the responses are correlated to grades given to students in the course. Professors, however, are less convinced than associates or assistants that student evaluation responses are indicative of the type of course and that student evaluation responses are indicative of the workload of the course. The differences in perceptions, however, were not statistically significant.

A singular conclusion that can be drawn from the reported results is that the differences in perceptions between the views of faculty and administrators relating to the validity of student evaluations cannot be attributed to the administrator's higher rank. Our results suggest that an individual's perception of validity increases not when the individual is promoted, but rather when the individual takes on an administrative hat. These results indicate

**Are Student
Evaluations a
Valid Measure
of Teaching
Effectiveness**

that since faculty believe that student evaluations are indicative of factors other than teaching effectiveness, faculty are likely to attempt to manage their ratings on student evaluations by lowering standards and assigning higher grades. In order to avoid this likely outcome, administrators must either convince faculty that the biases are not as great a problem as perceived or take into consideration these potential biases when making their own assessment of the effectiveness of the faculty member's teaching.

References

- Andersen, K. and E. Miller (1997), "Gender and Student Evaluations of Teaching", *PS: Political Science & Politics*, June, pp.216-219.
- Calderon, T. and B. Green (1997), "Use of Multiple Information Types in Assessing Accounting Faculty Teaching Performance", *Journal of Accounting Education*, 15, pp.221-237.
- Calderon, T., B. Green, and B. Reider (1997), "Perceptions and Use of Student Evaluations by Heads of Accounting Departments", *Accounting Educators' Journal*, 9, pp.1-27.
- Cohen, P.A. (1981), "Student Ratings of Instruction and Student Achievement: A Meta-Analysis of Multisection Validity Studies", *Review of Educational Research*, 51, pp.281-304.
- Crader, K. and J. Butler Jr. (1996), "Validity of Students' Teaching Evaluation Scores: The Wimberly-Faulkner-Moxley Questionnaire", *Educational and Psychology Measurement*, 56, pp.304-314.
- Hasselback, J. (1997), *Accounting Faculty Directory*, New York: Prentice Hall.
- Kemp, B. Wayne and G.S. Kumar (1990), "Student Evaluations: Are We Using Them Correctly?", *Journal of Education for Business*, 65, November/December, pp.106-111.
- Marsh, H.W. (1982), "Factors Affecting Students' Evaluations of the Same Course Taught by the Same Instructor on Different Occasions", *American Education Research Journal*, 19, pp.485-497.
- Marsh, H.W. (1984), "Students' Evaluations of University Teaching: Dimensionality, Reliability, and Validity, Potential Biases, and Utility", *Journal of Educational Psychology*, 76, pp.707-754.
- Marsh, H.W. (1987), "Students' Evaluation of University Teaching: Research Findings, Methodological Issues, and Directions for Future Research", *International Journal of Educational Research*, 11, pp.253-388.
- Marsh, H.W., and J.U. Overall (1981), "Validity of Students' Evaluations of Teaching Effectiveness", *Journal of Educational Psychology*, 72, pp.468-475.
- Marsh, H.W. and J.E. Ware, Jr. (1982), "Effects of Expressiveness, Content Coverage, and Incentive on Multidimensional Student Rating Scales: New Interpretation of the Dr. Fox Effect", *Journal of Educational Psychology*, 74, pp.126-134.

**Are Student
Evaluations a
Valid Measure
of Teaching
Effectiveness**

Are Student Evaluations a Valid Measure of Teaching Effectiveness

Mason, P., J. Steagall, and M. Fabritius (1995), "Student Evaluations of Faculty: A New Procedure for Using Aggregate Measures of Performance", *Economics of Education Review*, 14, pp.403-416.

Murray, H.G. (1983), "Low-Inference Classroom Teaching Behaviors and Student Ratings of College Teaching Effectiveness", *Journal of Educational Psychology*, 75, pp.856-865.

Seldin, P. (1984), *Changing Practices in Faculty Education*, San Francisco: Jossey-Bass.

Seldin, P. (1993), "The Use and Abuse of Student Ratings of Professors", *The Chronicle of Higher Education*, July, p.A40.

Simpson, R. (1995), "Uses and Misuses of Student Evaluations of Teaching Effectiveness", *Innovative Higher Education*, Fall, pp.3-5.

Stratton, R., S. Myers, and R. King (1994), "Faculty Behavior, Grades, and Student Evaluations", *Research in Economic Education*, Winter, pp.5-15.

Timpson, W. and D. Andrew (1997), "Rethinking Student Evaluations and the Improvement of Teaching", *Studies in Higher Education*, 22, pp.55-65.

White, C.S. and M. Ahmadi (1986), "A Novel Approach to the Analysis of a Teacher Evaluation System", *Journal of Education for Business*, 61, October, pp.24-27.

Yunker, P.J. and J. Sterner (1988), "A Survey of Faculty Performance Evaluation in Accounting", *Accounting Educators' Journal*, Fall, pp.63-71.