

Report of Faculty Senate Committee on Student Evaluation of Teaching

2-20-09

Context:

Two workshops, titled “How Can/Do We Evaluate Teaching at LMU,” were held in the Center for Teaching Excellence (CTE) in Fall 2006 and an informal follow-up survey was conducted. Significant dissatisfaction was reported with the current Student Rating form. The director of CTE submitted a report of the workshops and survey findings to the Faculty Senate and the CAO recommending that a committee be appointed to review/revise the student rating form. This recommendation was reinforced in discussions at the 2007 Academic Leadership Workshop, in which participants expressed strong dissatisfaction with the current teaching evaluation instrument. A committee was convened in late Summer 2007 and began meeting every two weeks throughout the 2007-8 academic year and resumed meetings again this academic year.

Committee Members:

James Roe (Chair), Richard ‘Sonny’ Espinoza, Paul Humphreys, Linda Leon, Michael Miranda, Jennifer Pate, Kala Seal, with Jackie Dewar and Margaret Kasimatis serving as resource persons to the committee.

Point of View:

Because teaching is a complex human endeavor, no single measure can accurately evaluate it. Yet, Student Ratings of Teaching (SRT) are the most commonly used measure and often the only measure for evaluating teaching in higher education. Best practice calls institutions to encourage the use of multiple measures (Knapped and Cranton, 2001). Practicality prompts them to employ a single SRT form across multiple disciplines and pedagogies. Ideally, an SRT form would be aligned with the institution’s own particular mission and situation. No matter what form is adopted, it is essential to ensure that the resulting SRT data are appropriately interpreted especially when making personnel decisions (merit, tenure, promotion).

Problems with the Current Form:

- One of the major criticisms of the current form by LMU faculty is that some of the rating items are not equally applicable to all types of pedagogies, and that overall the form seems more appropriate for lecture-based courses and less appropriate, even problematic, for courses that employ group work, service learning, student presentations, laboratories, performance rehearsals, etc.
- The current form is seen as an impediment to experimentation with new pedagogies.
- The current form may discourage faculty from increasing academic rigor in their courses.
- Several of the items on the current form also fail to pass one or more of the criteria for SRTs outlined in Berk (2006) such as “The statement should be a simple sentence” and “The statement should contain only one complete behavior, thought or concept.” See Appendix 1 for the entire list of criteria.
- The yes/no question (“If your schedule allowed it, would you take another course taught by this professor, or would you recommend the professor to someone else?”) on the current form was seen as particularly problematic for several

- reasons: (1) its complex wording is often misinterpreted by students, resulting in inaccurate ratings; (2) because of its global nature, it often becomes the only question that is considered when faculty teaching performance is reviewed.
- Data from the current form is often misinterpreted and conclusions are drawn about teaching effectiveness that the data does not support.

Goals:

The Committee's goals were to develop and pilot test a new SRT form that would

- be pedagogically neutral, that is, it would not have items that were more appropriate for one type of pedagogy than another
- have items that capture dimensions of effective teaching at LMU that students are capable of assessing, and
- follow recommendations for constructing such forms (Feldman, 2007; Berk, 2006).

In addition, the Committee recognized that it was equally important to

- educate faculty and administrators about how to interpret the data obtained from any SRT, when valid conclusions could be extracted from the data and what valid conclusions could be drawn.

Process:

The volume of literature dealing with research on student ratings is immense. In 1988 Cashin noted that over 1300 books and articles existed on the topic. No doubt the number today is more than double that figure. Informed by IDEA Papers #21 & 22 (Cashin, 1989, 1990) and Chickering and Gamson's Seven Principles (1987), the committee first attempted to identify measurable characteristics of effective teaching and then formulated corresponding items for an SRT form that was piloted during the 2007-8 academic year, revised and piloted again in Fall 2008.

Both the current and the 'pilot' SRT forms were administered to 275 students in 14 classes in Fall 2007 and to 479 students in 22 classes in Spring 2008. The revised pilot SRT was piloted again in Fall 2008 with 306 students in 27 classes completing both forms. A total of 37 faculty from all six schools and colleges were involved in the pilots, and both graduate and undergraduate courses were involved. All of the Committee members during the pilot period participated in at least one of the pilots. (See Appendix 2 for a detailed breakdown.)

Only after the initial piloting of the form in Fall 2007 did the Committee encounter the most recent research and recommendations of Feldman (2007) on the dimensions of teaching that correlate with learning achievement and Berk (2006) on best practice in construction of SRT forms. The Committee re-examined the pilot form in light of these new findings. This "review" was also informed by a meeting of the Southern California Faculty Developers Learning Community devoted to evaluating teaching via SRT held at LMU on January 11, 2008. We found through Theall and Feldman (2007) and Abrami, Rosenfield, and Dedic (2007) and Abrami, d'Apollonia and Rosenfield (2007) that Feldman's original dimensions (see Feldman, 1998) remained valid. These dimensions

listed in order of importance (from #1 as most important) to learning (from Feldman, 1998) follow.

Instructional Dimension	Ranked Importance with Learning Achievement	Addressed by these Items on the Current form	Addressed by these Items on the Pilot form
Preparation and organization	1	2,3,5,8,9	1, 2
Clarity and understandability	2	5	1, 2
Perceived outcome or impact, including opportunity to succeed	3	4,6,7	5, 6
Stimulation of interest in content	4		7
Encouragement and openness	5.5		3
Availability and helpfulness	5.5	1	3, 4

The Committee then reviewed the pilot form it had constructed to see that all six of these dimensions were addressed by nine ratings items on the form and noted that the current form heavily weights the Preparation and organization dimension and the Perceived outcome dimension and fails to address two dimensions: Stimulation of interest in content and Encouragement and openness.

The committee had recognized stimulation of interest in content as a component of effective teaching and included it as item #7, but also realized that a students' initial motivational level is an important covariant factor, that is, responses to #7 must be interpreted in light of responses to the last demographic item, "I had interest in taking this course." This provides a context for interpreting the results. In fact, Cashin (1990) recommends that the students' motivation level be taken into consideration when interpreting any student rating data because it shows higher correlations with other student rating items than any other variable. This means that instructors are more likely to get higher ratings in classes where students had a prior interest in the subject matter or were taking the course as an elective. Hence the committee felt that it was essential to include the last demographic item.

Of particular concern was whether the newly constructed form had items that addressed the top two dimensions: #1 Preparation and organization and #2 Clarity and understandability. Because of its goal to produce a pedagogically neutral form, the Committee had decided not to include any item with language that dealt directly with organization of classroom presentations such as, "The classroom presentations proceeded in an organized manner," opting instead to use items that rated clarity of learning goals and whether those goals were effectively addressed by course activities or experiences (Items #1 and 2 on the newly constructed form). However, because the current form contained so many items related to the first two dimensions (organization and clarity of presentations) as compared to the pilot form, the Committee felt it would be important to determine if the pilot form still captured those dimensions.

Statistical analysis showed significant positive correlation between #2 on the pilot form and items #2 (.86), #4 (.79) and #5 (.85) on the current form. These correlations are extremely high; in social science research a correlation of .30 is considered large. The committee concluded from this analysis that the pilot form constructed to be pedagogically neutral still captured key instructional dimensions (organization and clarity) that were more obvious (but possibly pedagogically dependent) on the old form.

A statistical analysis was performed to test the internal consistency of the pilot form. As a measure of reliability, Cronbach's coefficient alpha was used to examine internal consistency of the pilot form. Coefficient alpha splits the scale into two halves and correlates the subjects' scores on the two halves. The test then repeats this procedure for all possible splits of the scale into two halves. The value of coefficient alpha is the average of these correlations. A value of .70 is the generally agreed upon acceptable lower limit. Using data from the first two pilots (739 valid responses to the nine items of the pilot form); a coefficient alpha of .864 was found. This indicates that the pilot form possesses a high degree of internal consistency.

Berk (2006) provided criteria for writing rating items (see Appendix 1). These criteria had been not easily accessible in a single source until recently. Some items on most commercial forms, including those that tout their psychometric reliability and validity, can't pass these criteria and some items on LMU's current form do not pass these criteria. The revised pilot form (see Appendix 3) passes all of these criteria.

Student perceptions of the new form were gathered in two ways: (1) Two student focus groups (one with 8 grad students; one with 8 undergrad students) were held in early Spring 2008 to review and comment on the two forms (NOTE: these students did not actually use the forms); (2) An open-ended question on the 'pilot' SRT form asked students the following: "In the space below please COMMENT on this evaluation form as compared to the current LMU course evaluation form, if you are able to make a comparison." The results were: 665 students (out of 1060) expressed opinions that could be coded as to preference. Of those 60 % indicated a preference for the pilot form, 14 % for the current form and 26 % expressed indifference or equal satisfaction with the two forms. The responses from the first two pilot tests were also coded for reasons given in support of a preference. Far and away, the most common reason given (45%) was "better content." The focus groups provided mixed results with the majority of the graduate students preferring the new form but the undergraduates being very mixed.

Faculty perceptions of the new form were gathered by an on-line survey of piloters, excluding any piloters on the SET Committee. When asked if they had a preference for one of the forms 62.5% said they preferred the pilot form, while only 6.25% said they preferred the existing form. The remaining 31.25% indicated no preference for either form.

Guidelines for interpreting SRT data based on Pallet (2006) and Theall and Franklin (1991) were developed by the Center for Teaching Excellence and disseminated by the Committee to all faculty (August 2008), and chairs, deans, and members of the

Committee on Rank and Tenure (Academic year 2007-8). These guidelines are posted at <http://www.lmu.edu/Page35750.aspx> and found in Appendix 4.

Summary of Results:

- The Committee has developed (through an iterative process involving three pilot tests) a form that is pedagogically neutral.
- The proposed form has items capturing important dimensions of effective teaching associated with student learning (and which students are capable of assessing).
- All items meet criteria for survey construction.
- The proposed form includes an item related to academic rigor.
- The proposed form includes a demographic item regarding the student's initial desire to take the course as one perspective for interpreting the data.
- The proposed form allows instructors or departments to add two items of their own choosing.
- Distributed guidelines for interpreting student evaluation form data to all faculty, deans, chairs and the Committee on Rank and Tenure and received positive feedback from a number of individual faculty.
(See <http://www.lmu.edu/Page35750.aspx>).
- Pilot tested the form 63 sections of courses, both grad and undergrad, taught by 37 faculty representing all 6 schools and colleges in the university (Total number of students = 1060).
- Statistical analysis shows that certain questions on the pilot form revised to be pedagogy-neutral still capture key instructional dimensions that were on the old form (such as organization and clarity).
- A qualitative analysis of the 754 forms administered in 37 classes during the 2007-2008 academic year indicates a very strong preference for the new form by student users of the form for a variety of reasons, most commonly "better content."
- The results of the faculty satisfaction survey (n = 16 respondents) showed, among those who indicated a preference, a 9 to 1 preference for the pilot form.
- Disseminated reports of the Committee's work to all faculty through three information/discussion sessions (Nov 4, 5, and 10, 2009) to the Faculty Senate Fall 2008 and Spring 2009, and to Deans Council Fall 2008 and Spring 2009.

Recommendation made to Faculty Senate February 26, 2009

Whereas the purpose of the Student Evaluation Committee was to review and revise the current evaluation form utilizing the latest research on using student rating forms to evaluate teaching, taking into consideration the most important dimensions of effective teaching while creating a pedagogically-neutral instrument and following best practice in survey construction

Be it resolved that the Faculty Senate endorse the new Student Rating of Teaching form and recommend adoption beginning in Fall 2009.

References cited:

- Abrami, P. C., d'Apollonia, S., and Rosenfield, S. (2007). The dimensionality of student ratings and instruction: What we know and what we do not. *in* R.P. Perry and J.C. Smart, (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 385-546.
- Abrami, P. C., Rosenfield, S. and Dedic, H. (2007). The dimensionality of student ratings and instruction: An update on what we know, do not know, and what we need to do. *in* R.P. Perry and J.C. Smart, (eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 446-456.
- Berk, R. A. (2006). *Thirteen Strategies to Measure College Teaching*. Sterling, VA: Stylus.
- Cashin, W. E. (1990). Student ratings of teaching: recommendations for use: Kansas State Univ. Center for Faculty Evaluation and Development, Idea Paper n. 22. Available September 21, 2008 from <http://www.idea.ksu.edu/resources/Papers.html>.
- Cashin, W. E. (1989). Defining and evaluating college teaching: Kansas State Univ. Center for Faculty Evaluation and Development, Idea Paper n. 21. Available September 21, 2008 from <http://www.idea.ksu.edu/resources/Papers.html>.
- Cashin, W. E. (1988). Student ratings of teaching: a summary of the research: Kansas State Univ. Center for Faculty Evaluation and Development, Idea Paper n. 20. Available September 21, 2008 from <http://www.idea.ksu.edu/resources/Papers.html>.
- Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics: a review and synthesis: *Research in Higher Education*. 24 129-213.
- Feldman, K. A., (1998). Identifying exemplary teachers and teaching: evidence from student ratings: *in Teaching and Learning in the College Classroom* 2nd ed., K. A. Feldman and M. B. Paulsen, (Eds.), Needham Heights, MA: Simon & Schuster. 391-414.
- Feldman, K. A. (2007), Identifying exemplary teachers and teaching: evidence from student ratings. *in* R.P. Perry and J.C. Smart, (Eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 93-129.
- Nuhfer, E. B. (2008). The affective domain and the formation of the generator: Educating in fractal patterns XXIII part 1: National Teaching and Learning Forum, 18/2 8-11. and part 2, National Teaching and Learning Forum, 18/3 9-11. Available from any computer on the CSUCI campus at <http://www.ntlf.com/restricted/>.
- Nuhfer, E. B. and Dewar, J. (2008). "Guidelines for Creating Good Items for Student Ratings of Professors." SoCal Faculty Developers Learning Community. Los Angeles, CA: <http://www.lmu.edu/pagefactory.aspx?PageID=41603>.
- Pallett, W. (2006) "Uses and abuses of student ratings." In *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Peter Seldin (ed.). Bolton, MA: Anker Publishing.
- Theall, M. and Franklin, J. (eds.) (1991). *Effective practices for improving teaching*. New Directions in Teaching and Learning, no. 48, San Francisco: Jossey-Bass.

Theall, M., and Feldman, K. A., (2007) Commentary and update of Feldman's (1997) "Identifying exemplary teachers and teaching: evidence from student ratings." in R.P. Perry and J.C. Smart, (Eds.), *The Scholarship of Teaching and Learning in Higher Education: An Evidence-Based Approach*. Springer, 130-143.

Appendix 1: Criteria for writing items for SRT (Berk, 2006)

1. The statement should be *clear and direct*.
2. The statement should be *brief and concise*.
3. The statement should contain only *one complete behavior, thought, or concept*.
4. The statement should be a *simple sentence*.
5. The statement should be at the *appropriate reading level*.
6. The statement should be *grammatically correct*.
7. The statement should be *worded strongly*.
8. The statement should be *congruent with the behavior* it is intended to measure.
9. The statement should accurately measure a *positive or negative behavior*.
10. The statement should be *applicable to all respondents*.
11. The respondents should be in the *best position to respond* to the statement
12. The statement should be *interpretable in only one way*.
13. The statement should NOT contain a *double negative*.
14. The statement should NOT contain *universal or absolute terms*.
15. The statement should NOT contain *nonabsolute, warm-and-fuzzy terms*.
16. The statement should NOT contain *value-laden or inflammatory words*.
17. The statement should NOT contain *words, phrases, or abbreviations* that would be *unfamiliar to all respondents*.
18. The statement should NOT tap a *behavior appearing in any other statement*.
19. The statement should NOT be *factual* or capable of being interpreted as factual.
20. The statement should NOT be *endorsed* or given one answer by *almost all respondents or by almost none*.

Appendix 2: Breakdown of pilot courses by school and college, and level.

School/College # sections piloted

BCLA	17
CBA	10
CFA	10
SCSE	17
SFTV	1
SoE	8

Course Level

Lower division	23
Upper division	27
Graduate	13

Appendix 3: LMU Course Evaluation Forms

Final version of revised pilot for vote

Revised Pilot administered in Fall 2008

Original Pilot Form administered in Fall 2007 and Spring 2008

Current Course Evaluation Form

Loyola Marymount University Course Evaluation Form

Course Title _____ Term _____

Instructor _____ Course/Section _____

MARKING INSTRUCTIONS

Please use a No. 2 pencil or black or blue ball point pen.

Correct Mark



Incorrect Marks



Major(s): _____

Class Year: Freshman Sophomore Junior Senior Graduate Other

LMU Cumulative GPA: Under 2.0 2.0-2.49 2.5-2.99 3.0-3.49 3.5-4.0 Not Applicable

Your Class Attendance: Rarely Occasionally Usually Almost Always Always

Was this course in your major department? Yes No

Was this course required by your major? Yes No

Are you taking this course to fulfill Core requirements? Yes No

I had a strong interest in taking this course: Strongly Disagree Disagree Uncertain Agree Strongly Agree

Instructions: Please MARK the response which MOST ACCURATELY reflects your opinion and include any comments or explanations to the following questions.

Strongly Disagree Disagree Uncertain Agree Strongly Agree

1. Learning outcomes for the course were clearly stated. SD D U A SA

Comments:

2. The learning outcomes were effectively addressed in this course. SD D U A SA

Comments:

3. There were constructive interactions between the instructor and the students. SD D U A SA

Comments:

4. The instructor was accessible for discussions about the course. SD D U A SA

Comments:

5. I received feedback that improved my learning in this course. SD D U A SA

Comments:

Strongly Disagree Disagree Uncertain Agree Strongly Agree

6. The course challenged me to do my best work.

SD D U A SA

Comments:

7. My experience in the course increased my interest in the subject matter.

SD D U A SA

Comments:

8. How would you rate the overall effectiveness of the instruction in this course?

Very Poor Poor Fair Good Very Good

Comments:

9. *OPTIONAL question for faculty or departmental use.*

SD D U A SA

Comments:

10. *OPTIONAL question for faculty or departmental use.*

SD D U A SA

Comments:

11. What did you find to be most beneficial about the course?

12. What would have made this course more effective for you?

	Strongly Agree	Agree	Uncertain	Disagree	Strongly Disagree
6. I received feedback that helped improve my learning in this course.	5	4	3	2	1
Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7. The course challenged me to do my best work.	5	4	3	2	1
Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8. My experience in the course increased my interest in the subject matter.	5	4	3	2	1
Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

	Very Good	Good	Fair	Poor	Very Poor
9. How would you rate the overall effectiveness of the instruction in this course?	5	4	3	2	1
Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Questions 10 and 11 will be on the final version of this form; DO NOT answer them at this time.

10. What did you find to be most beneficial about the course?

11. What would have made this course more effective for you?

Please ANSWER the following:

12. In the space below please COMMENT on this evaluation form as compared to the standard LMU course evaluation form, if you are able to make a comparison.

	Definitely True	Somewhat True	Uncertain	Somewhat False	Definitely False
6. I received frequent feedback on my comprehension of or ability to apply course material.	5	4	3	2	1

Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
-----------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

7. The course challenged me to do my best work.	5	4	3	2	1
---	---	---	---	---	---

Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
-----------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

8. The instructor increased my interest in the subject matter.	5	4	3	2	1
--	---	---	---	---	---

Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
-----------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

	Very Good	Good	Fair	Poor	Very Poor
9. How would you rate the overall effectiveness of the instructor?	5	4	3	2	1

Comments:	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
-----------	-----------------------	-----------------------	-----------------------	-----------------------	-----------------------

Questions 10 and 11 will be on the final version of this form; DO NOT answer them at this time.

10. What did you find to be most beneficial about the course?

11. What would have made this course more effective for you?

Please ANSWER the following:

12. In the space below please COMMENT on this evaluation form as compared to the standard LMU course evaluation form, if you are able to make a comparison.

Course Title _____ Term _____

Instructor _____ Course/Section _____

MARKING INSTRUCTIONS

Please use a No. 2 pencil or black or blue ball point pen.

Correct Mark



Incorrect Marks



Please make no marks in this Area



Major: _____

Class Year: Freshman Sophomore Junior Senior Graduate Other

LMU Cumulative GPA: Under 2.0 2.0-2.49 2.5-2.99 3.0-3.49 3.5-4.0 N/A

Your Class Attendance: Always Almost Always Usually Occasionally Rarely N/A

Was this course in your major department? Yes No N/A

Was this course required by your major or the University core? Yes No N/A

Instructions: Please MARK the response which MOST accurately reflects your opinion, and explain your responses to items 1-10.

1. The professor was available for consultation during office hours.

Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:

2. The classroom presentations appeared organized with a sense of direction to the instruction.

Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:

3. Class time was used effectively.

Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:

4. Out-of-class assignments or projects were relevant to defined course content.

Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:

5. The professor's presentations facilitated learning of the course content.

Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:

6. Tests were representative of the course content.

Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:



7. The professor provided helpful and timely oral or written evaluation regarding tests, assignments and/or performances.

- Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:

8. The professor provided a description of grading standards and procedures at the beginning of the course.

- Constantly Most of the time Some of the time Seldom Almost never N/A

COMMENTS:

9. Required texts were both used in and useful during the course, or were useful as a reference.

- Always Usually Sometimes Infrequently Never N/A

COMMENTS:

10. If your schedule allowed it, would you take another course taught by this professor, or would you recommend this professor to someone else?

- Yes No N/A

Yes COMMENTS:

NO COMMENTS:

11. What did you find to be most beneficial about the course?

12. What would have made this course more effective for you?

Appendix 4: Guidelines for Interpreting Student Evaluations

Student teaching evaluations are the most commonly used measure for evaluating teaching in higher education. There are at least two purposes for evaluating teaching: to improve the teaching and to make personnel decisions (merit, retention, promotion). When using student teaching evaluations for either of these purposes, it is essential to follow certain guidelines to ensure valid interpretation of the data. The following guidelines are adapted from Theall and Franklin (1991) and Pallett (2006).¹

#1. *Sufficient Response Ratio*

There must be an appropriately high response ratio.² For classes with 5 to 20 students enrolled, 80% is recommended for validity; for classes with between 21 and 50 students, 75% is recommended. For still larger classes, 50% is acceptable. Data should not be considered in personnel decisions if the response rate falls below these levels.

#2. *Appropriate Comparisons*

Because students tend to give higher ratings to courses in their majors or electives than they do to courses required for graduation, the most appropriate comparisons are made between courses of a similar nature. For example, the Bellarmine College of Liberal Arts average would *not* be a valid comparison for a lower division American Cultures course.

#3. *When Good Teaching is the Average*

When interpreting an instructor's rating, it is more appropriate to look at the actual value of the rating instead of comparing it to the average rating. In other words, a good rating is still good, even when it falls below the average.

#4. *Written Comments*

Narrative comments are often given great consideration by administrators, but this practice is problematic. Only about 10% of students write comments (unless there is an extreme situation), and the first guideline recommends a minimum 50% response threshold. Thus decisions should not rest on a 10% sample just because the comments were written rather than given in numerical form! Student comments can be valuable for the insights they provide into classroom practice and they can guide further investigation or be used along with other data, but they should not be used by themselves to make decisions.

#5. *Other considerations*

Class-size can affect ratings. Students tend to rank instructors teaching small classes (less than 10 or 15) most highly, followed by those with 16 to 35 and then those with over 100 students. Thus the least favorably rated are classes with 35 to 100 students.

There are disciplinary differences in ratings. Humanities courses tend to be rated more highly than those in the physical sciences.

#6. *One Final Point*

Teaching is a complex and multi-faceted task. Therefore the evaluation of teaching requires the use of multiple measures. In addition to teaching evaluations, the use of *at least* one other measure, such as peer observation, peer review of teaching materials (syllabus, exams, assignments, etc.), course portfolios, student interviews (group or individual), and alumni surveys is recommended. Contact the Center for Teaching Excellence (310-338-2772) if you need assistance in adopting one of these alternate measures or have any questions about these guidelines.

¹ Pallett, W. "Uses and abuses of student ratings." In *Evaluating faculty performance: A practical guide to assessing teaching, research, and service*. Peter Seldin (ed.). Bolton, MA: Anker Publishing, 2006.

Theall, M. and Franklin, J. (eds.) *Effective practices for improving teaching*. New Directions in Teaching and Learning, no. 48, San Francisco: Jossey-Bass, 1991.

² The following describes how to compute the response ratio for a given set of forms from one section of one course. First, note the number (n) of forms returned and the number (N) of students in the class, compute the ratio n/N, and then multiply by 100% to convert the ratio to a percent. Then, for each question under consideration, from this percent subtract the percent of blank and "Not Applicable" responses. The result is the response ratio for that particular question. If the result does not meet the threshold recommended in Guideline #1 above, the data from that question should not be considered.