

## **Provost's Advisory Committee on Teaching and Learning (PACTL) Report to Senate**

### **REVIEW OF NEW UNIVERSITY-WIDE TEACHING EVALUATION FORM**

#### Introduction

At its June, 1996 meeting, Senate approved a new Instructor and Course Evaluation form (see Exhibit 1) for university-wide use beginning in the 1996-97 academic year. The distinctive features of the new teaching evaluation form include numerical ratings of a broad range of teacher and course characteristics, written comments from students on instructor and course, assessment of student characteristics such as class attendance and expected grade, and use of a 7-point poor-outstanding rating scale rather than the traditional 5-point agree-disagree scale. The teacher and course characteristics included on the new teaching evaluation form were selected according to the following criteria: (1) observable by students, (2) under the control of the instructor, (3) applicable to all or nearly all forms of teaching, and (4) positively related to student learning.

This report summarizes a review of the new teaching evaluation form conducted, at the request of Senate, by a subcommittee of the Provost's Advisory Committee on Teaching and Learning (PACTL). The subcommittee consisted of Harry Murray (Psychology), Michael Atkinson (Educational Development), Colin Baird (Chemistry), Debra Dawson (Educational Development), and Jeff Tennant (French). PACTL's review of the new evaluation form was divided into three parts: (1) survey of student opinion, (2) survey of faculty opinion, and (3) statistical analysis of selected aspects of reliability and validity. These are reported in turn below, following a brief summary of previous research on student evaluation of university teaching.

#### Previous Research on Student Evaluation of Teaching

Over 2000 references on student evaluation of college and university teaching have appeared in the research literature since 1970. Much of this research has focused on two key questions: (1) Does student evaluation provide reliable and valid information on quality of teaching? and (2) Does student evaluation lead to improvement of teaching? Although individual studies sometimes report diverging results, there is surprising consensus in research findings relating to these two questions, as summarized in Exhibits 2 and 3 respectively. The weight of research evidence indicates: (1) that student evaluation is reasonably reliable and valid as a measure of teaching, and (2) that student evaluation does in fact lead to improvement of teaching. Evidence supporting student evaluation as a reliable and valid source of information on quality of teaching includes the fact that evaluations of a given instructor tend to be consistent across years and courses, and the fact that students taught by highly evaluated teachers tend to perform better on objective measures of amount learned than students taught by less highly evaluated teachers. Evidence supporting the view that student evaluation leads to improvement of teaching includes the fact that instructors receiving mid-term student feedback tend to be evaluated more positively at end-of-term than teachers receiving no mid-term feedback, and the fact that the mean teacher rating for an academic unit as a whole tends to improve gradually across successive years following the introduction of a student evaluation program in that unit. It was expected that results similar to those summarized in Exhibits 2 and 3 would be found for the new teaching evaluation form adopted at UWO.

#### Student Survey

Student opinion of the new UWO teaching evaluation form was surveyed by a one-page questionnaire administered in 22 different undergraduate classes representing 16 different academic disciplines. Students responded to 11 questions on a 5-point agree-disagree rating scale, and were given the opportunity to provide supplementary written comments. A total of 1123 out of 1615 student registrants completed the questionnaire, giving a return rate of 69.5%. Female students and Social Science students were overrepresented in the sample relative to their frequency in the UWO student population as a whole.

Table 1 shows the percentage of students who agreed, disagreed, or were undecided for each of the 11 survey items. For purposes of this table, "agree" and "strongly agree" responses have been combined into one category, as have "disagree" and "strongly disagree" responses. It may be noted that students were generally very positive about the new teaching evaluation form. Approximately 85% of respondents said that the new evaluation form assesses a sufficiently wide range of characteristics, whereas 81% considered the form to be applicable to the style of teaching to which they were accustomed, 92% thought the written comments section of the evaluation form was useful, 77% believed the new 7-point rating scale was more appropriate than the previous 5-point scale, and 60% agreed that, all things considered, the new evaluation form was superior to previous forms with which they were familiar. Students were strongly opposed to deleting the written comments section or making it optional for individual faculty members (90% and 84% disagreement respectively). On the other hand, only 20% of respondents thought that sufficient weight was placed on student evaluation of teaching in faculty personnel decisions at UWO, only 14% were aware that teaching evaluations are published on the UWO Web site, and only 4% said they used published evaluations in selecting courses. The latter two results may be due to the fact that last year's teaching evaluations were not available on the Internet until after most students had already registered for courses.

Table 2 summarizes students' supplementary comments regarding the new evaluation form. A total of 203 individual comments were received, which could be grouped into 24 categories. Only those categories with a response frequency of 10 or higher are reported in Table 2. It may be noted that one common theme in the supplementary comments was lack of emphasis on student evaluation in personnel decisions or for improvement of teaching, another was preference for more evaluation of the course as opposed to the instructor, and a third was the need to make students more aware of the reasons for student evaluation of teaching, and the uses of evaluation results.

### Faculty Survey

Faculty opinion of the new teaching evaluation form was assessed by a questionnaire sent by campus mail to all full-time faculty members who had been at UWO for at least three years and had been evaluated during the 1996-97 academic year on the new form. The questionnaire consisted of 18 items answered on a 5-point agree-disagree rating scale, plus a section for supplementary written comments. Of 826 faculty members surveyed, completed questionnaires were received from 273, giving a return of 33.1%. With such a low return rate, the possibility exists that the opinions of respondents were not representative of the UWO faculty as a whole. On the other hand, it should be noted that the distribution of faculty ranks, genders, and faculty affiliations in the sample of respondents did not differ significantly from that of the faculty as a whole.

Table 3 shows the percentage of faculty respondents who agreed, disagreed, or were undecided for each of the 18 questionnaire items. In comparison to students, the reaction of faculty respondents to the new teaching evaluation form could be described as "luke warm". Approximately 73% of faculty members said that the new teaching evaluation form assessed a sufficiently wide range of teacher characteristics (vs.85% for students), whereas 72% believed that the new form was applicable to their style of teaching (vs.81% for students), and 72% agreed that the written comments section was a useful supplement to numerical ratings (vs.92% for students). Also, 69% of faculty members stated that the new teaching evaluation form was suitable for use in salary, promotion, and tenure decisions, 64% thought that it provided useful feedback for improvement of teaching, and 60% believed that the written comments section had provided them with useful diagnostic feedback. On the negative side, only 44% of faculty respondents agreed that the new teaching evaluation form was superior to previous forms used in their departments, and only 42% thought that the new 7-point rating scale was more useful for use in personnel decisions than previous 5-point scales, although for both of these questions a large percentage of respondents was undecided (31% and 27% respectively), which may possibly reflect a "wait and see" position.

Another cause for concern in the faculty survey was the surprisingly large number of respondents (28%) who reported having received abusive, obscene, or otherwise inappropriate written comments from students in their 1996-97 teaching evaluations. In the absence of direct inspection of student comments (which was not undertaken because of confidentiality of data), it is impossible to know how many of these objectionable

comments were actually abusive or obscene, and how many were inappropriate for other reasons (for example, they were silly or childish or unfounded). Also, contrary to the initial expectation that reports of abusive written comments would be more frequent for female than for male faculty members, breakdown of survey responses according to gender indicated approximately equal frequencies for females and males (27% and 29 % respectively). Finally, it should be noted that although abusive comments were reported by a sizeable minority of respondents, the majority of faculty believed that written comments provided beneficial feedback (60%), were a useful supplement to numerical ratings (72%), should be retained in the new evaluation form (72%), and should not be optional for individual faculty members (65%).

Several items in the faculty survey asked respondents to comment on possible changes or improvements in the administration and use of the new teaching evaluation form. Responses to these items showed reasonable consensus. For example, 90% of respondents agreed that annual teaching evaluation results should be reported to faculty members by June 1, 91% thought that data from the four student information items (eg., class attendance, expected grade) should be included in annual feedback to faculty, 84% said that feedback to faculty should include comparative norms for different departments, course types, and class sizes, 74% believed that written comments from students should be used solely as feedback to the instructor, and not for salary, promotion, and tenure purposes, and 53% believed that designation of teacher evaluation items as "not applicable" should be decided in advance by the instructor rather than being left to the judgement of individual students.

Table 4 summarizes a content analysis of faculty respondents' supplementary written comments on the new evaluation form. A total of 317 individual comments were received, which could be divided into 38 categories. Only categories with a response frequency of 10 or higher are reported in Table 4. The most common theme in the supplementary comments was that the new evaluation form emphasized the performance of the teacher rather than the quality of the course, or favoured a particular lecture-style of teaching and thus was not suitable for use in some contexts. Other themes included the view that student ratings are not sufficiently valid to be used for evaluation of teaching, and the argument that students should be required to sign their names on teaching evaluation forms.

#### Statistical Analysis of Reliability and Validity

Tables 5, 6, and 7 report statistical analyses of the reliability (i.e., consistency or stability) and validity (i.e., accuracy or veridicality) of instructor mean ratings on the new teaching evaluation form for selected departments only. Most of these analyses are based on instructor mean ratings for undergraduate classes taught in the Department of Psychology in the 1995-96 and 1996-97 academic years. Supplementary analyses are also reported for the Departments of English and Mathematics.

Table 5 summarizes the results of several tests of the reliability or consistency of instructor mean ratings on the new evaluation form. As has been reported in previous research, these results indicate that instructor mean ratings exhibit: (1) high reliability or consistency across randomly selected subgroups of student raters (mean correlation = .84) ; (2) moderate to high reliability (correlations of .67 to .95) across different items or sets of items on the evaluation form; and (3) moderate reliability (correlation = .62) across different courses taught in the same academic year. In addition, results reported at the bottom of Table 5 show that, at least for psychology instructors, mean ratings on the new, 20-item university-wide teaching evaluation form in the 1996-97 academic year were consistent with mean ratings obtained in the 1995-96 academic year on the 10-item teaching evaluation form previously used in the Department of Psychology (.70 for same course, .72 for all courses combined.). The consistency between old and new evaluation forms was particularly strong (.95) for a sample of 25 instructors who, for research purposes only, were evaluated on both forms in the same course and same year (i.e., 1995-96). In summary, these results show that an instructor's teaching evaluation scores tend to be reasonably stable or consistent across different groups of raters, sets of items, courses, and teaching evaluation forms, indicating that alternative sources of data on teaching effectiveness tend to yield converging results.

Table 6 shows correlations between various teacher and course characteristics and instructor mean ratings on the overall effectiveness item (#19) of the new teaching evaluation form. These results are based on data for 103 psychology classes taught in the 1996-97 academic year. It may be noted that, on average, overall teacher effectiveness ratings were significantly higher for senior faculty (eg., full professors) than for junior faculty (eg., instructors); significantly higher for senior (eg., Year 4) courses than for junior (eg., Year 1) courses; significantly higher for courses with higher attendance levels on the day of evaluation; significantly higher for courses with higher reported attendance for the course as a whole (student information item #1); significantly higher for courses with higher mean expected final grades (student information item #2); and significantly higher for courses with higher mean ratings of initial level of student interest (student information item # 4). On the other hand, overall teaching effectiveness ratings did not differ significantly, on average, for female vs. male faculty members, for Fall Term vs. Spring Term administration of the teaching evaluation form, for small vs. large classes, or for required vs. optional courses. Results similar but not identical to those in Table 6 were obtained when the same analyses were repeated for two other departments with large undergraduate enrollments, namely English (N=160 classes) and Mathematics (N=85 classes). For example, course status (required vs. optional) correlated significantly ( $r = -.34$ ) with instructor mean ratings in Mathematics, but not in English or Psychology, whereas instructor rank correlated significantly with ratings in Psychology but not in English or Mathematics.

For most of the instructor and course characteristics listed in Table 6, it is difficult to know whether a significant correlation with instructor mean rating should be interpreted as evidence of "bias" or "error" in student evaluation of teaching, or as a valid reflection of factors that contribute to effective teaching. For example, a significant positive correlation between instructor rank and teacher rating could be interpreted to mean that students are biased against younger teachers, or alternatively, as a tendency for increased teaching effectiveness resulting from age and experience to be validly reflected in student ratings. Similarly, a positive correlation between mean expected grade and mean teacher rating could reflect a tendency for students to "reward" lenient-grading teachers with high ratings, or alternatively, could mean that students actually learn more in courses taught by more effective teachers, and this higher level of learning is reflected both in higher grades and in higher ratings of the instructor. Even in cases where correlation with instructor and course characteristics can be unambiguously interpreted as "bias" or "error", it is important to bear in mind that because of intercorrelation of bias factors, all sources of bias in combination typically account for only 10 to 15% of the total variance in instructor mean ratings. Also, it is possible to eliminate the impact of most sources of bias through the use of statistical adjustments or separate norm groups for different types of courses.

Table 7 compares the frequency distribution of psychology instructors' mean ratings on the overall effectiveness item of the new teaching evaluation form (1996-97 data) to the distribution of ratings on the corresponding item of the previous departmental teaching evaluation form (1995-96 data). It may be noted that overall mean ratings on old vs. new evaluation form were 3.68 and 5.51 respectively, both of which correspond to a verbal rating mid-way between "good" and "very good". Similarly, the percentage of instructors rated as "satisfactory" or better and as "good" or better was approximately the same on old and new evaluation forms. In combination with the old-new correlation data summarized in Table 5, these results suggest that teacher ratings obtained with the new and old evaluation forms in the Department of Psychology were similar in terms of general level of rating, distribution of ratings, and rank ordering of instructors. Finally, the data in Table 7 suggest some possible advantages for the new teaching evaluation form, including increased dispersion (standard deviation) of instructor mean ratings, reduced negative skewness (i.e. less leniency bias), and reduced kurtosis (i.e., a more normal or less peaked distribution).

## Conclusions

1. The research literature indicates that, as a general rule, student evaluation of teaching provides reliable and valid information on teaching effectiveness, and leads to measurable improvement in quality of teaching.
2. UWO students appear to be generally satisfied with the new teaching evaluation form, although they are concerned that insufficient emphasis is placed on student evaluation of teaching in faculty personnel decisions

and in teaching improvement.

3. Despite some reservations, UWO faculty members also appear to be generally satisfied with the new teaching evaluation form. Among other things, faculty members are concerned about abusive written comments from students and want earlier feedback of evaluation results and comparative norms to aid interpretation of student evaluations.

4. Statistical analyses indicate that the new teaching evaluation form shows reliability and validity data similar to that reported in the research literature. Specifically, instructor ratings on the new evaluation form appear to be reliable (stable) across raters, items, and courses, weakly correlated with extraneous variables, and consistent with ratings obtained on previously used evaluation forms.

---

#### Exhibit 2

Does student evaluation provide reliable and valid information on quality of teaching?

1. Student evaluations of a given instructor are reasonably consistent across raters, rating forms, courses, and time periods (reliability coefficients = .70 or higher).

2. Student evaluations agree with evaluations of the same instructors made by other, independent judges, such as colleagues and alumni (correlations generally .50 or higher).

3. Student evaluations show small but significant correlations with extraneous factors such as class size, strictness of grading, course level, academic discipline, and required vs. optional course status (correlations generally .30 or less), but all of these factors in combination account for less than 15% of the total variance in teaching evaluations, and in most cases it is possible to control the impact of extraneous factors through the use of separate norm groups or statistical adjustments.

4. Student evaluations correlate moderately with more objective indicators of teaching effectiveness, such as amount learned by students as measured by class mean performance on a common final exam in a multi-section course (mean correlation = .45), and student motivation for further learning as measured by frequency of enrollment in advanced courses (mean correlation = .70).

5. Student evaluations are predictable from trained observers' reports of the frequency of occurrence of specific classroom teaching behaviours (mean multiple correlation = .80), indicating that instructors receiving high evaluations do in fact teach differently than instructors receiving lower evaluations.

#### Reference

Marsh, H. W. and Dunkin, M. J. (1992). Students' evaluations of university teaching: A multidimensional perspective. In J. C. Smart (Ed.), Higher Education:

Handbook of Theory and Research. Volume 8. New York: Atherton.

---

#### Exhibit 3

Does student evaluation of teaching lead to improvement of teaching ?

1. Surveys of faculty opinion at various colleges and universities indicate that, on average, 74% of faculty members believe student ratings provide useful feedback for improvement of teaching, and 69% believe student ratings have in fact led to improved teaching.

2. Field experiments comparing faculty members randomly assigned to either receive or not receive mid-term feedback from students showed that feedback alone produced a small but significant (.10) gain in end-of-term

student ratings, whereas feedback supplemented by expert consultation produced a much larger (.40) gain in end-of-term ratings.

3. Longitudinal analyses of mean teacher ratings over periods of 3 to 25 years following the introduction of student evaluation of teaching in a given academic unit have generally (but not always) found gradual improvement in perceived quality of teaching across years.

4. Undesirable educational practices such as grade inflation, watering down of academic standards, and entrenchment of traditional "hierarchical" methods of teaching are sometimes attributed to student evaluation of teaching, but there appears to be no systematic empirical evidence to support these claims.

#### Reference

Murray, H. G. (1997). Does evaluation of teaching lead to improvement of teaching? *International Journal of Academic Development*, 2 , 8-23.

---

Table 1

#### Results of Student Survey

##### Questionnaire Item (paraphrased) and Percent Responses

1. The new teaching evaluation form assesses a sufficiently wide range of instructor and course characteristics.

Agree: 84.8

Disagree: 6.2

Undecided: 9.0

2. The items on the new evaluation form are generally applicable to the style of teaching to which I am accustomed.

Agree: 80.9

Disagree: 3.9

Undecided: 15.2

3. The separate course evaluation item is a useful feature of the new form.

Agree: 85.4

Disagree: 3.3

Undecided: 11.3

4. The written comments section allows me to comment on aspects of teaching not covered by numerical items.

Agree: 91.8

Disagree: 1.7

Undecided: 6.5

5. Individual faculty members should decide whether students complete the written comments section of the evaluation form.

Agree: 5.6

Disagree: 84.3

Undecided: 9.2

6. The written comments section should be deleted from the new evaluation form.

Agree: 3.3

Disagree: 90.0

Undecided: 6.7

7. I was aware that instructor and course evaluations are published on the UWO Web site.

Agree: 14.4

Disagree: 78.5

Undecided: 7.1

8. I used published teaching evaluations on the UWO Web site in selecting courses for this year.

Agree: 4.4

Disagree: 89.4

Undecided: 5.2

9. The 7-point rating scale on the new evaluation form is more appropriate for evaluation of teaching than the 5-point scale on previous evaluation forms.

Agree: 76.7

Disagree: 7.1

Undecided: 16.2

10. All things considered, the new teaching evaluation form is superior to other evaluation forms with which I am familiar.

Agree: 59.7

Disagree: 2.9

Undecided: 37.4

11. Sufficient weight or emphasis is given to student evaluation of teaching in decisions on faculty salary, tenure, and promotion at UWO.

Agree: 20.4

Disagree: 38.7

Undecided: 40.9

---

Table 2

Student Survey: Content Analysis of Supplementary Comments

---

Category	Frequency
Student evaluation of teaching is not taken seriously in personnel decisions or improvement of teaching.	51
There should be more items evaluating the course as opposed to the instructor (e.g., course quality, reading materials, work load, grading system).	26

Students are not made aware of the uses or purposes of student evaluation of teaching.	22
Some items on the new evaluation form are inappropriate (i.e., unclear, inapplicable, or too general).	20
The written comments sections of the new evaluation form are useful or valuable.	14
There should be increased opportunities for student evaluation of teaching (e.g., more time available, mid-term evaluation, alternative date for those who missed class).	10

---

Note: Results are reported only for comment categories with response frequencies of 10 or higher.

---

### Table 3

#### Results of Faculty Survey

#### Questionnaire Item (paraphrased) and Percent Responses

1. Results of student information items should be made available to faculty.

Agree: 91.0  
 Disagree: 1.5  
 Undecided: 7.5

2. Instructor ratings should be adjusted to account for factors such as expected grade and percent attendance.

Agree: 70.3  
 Disagree: 14.9  
 Undecided: 14.8

3. The new teaching evaluation form assesses a sufficiently wide range of instructor characteristics.

Agree: 73.4  
 Disagree: 15.0  
 Undecided: 11.6

4. The items on the new evaluation form are generally applicable to my style of teaching.

Agree: 72.1  
 Disagree: 16.6  
 Undecided: 11.3

5. The separate course evaluation item is a useful feature of the new form.

Agree: 72.1  
Disagree: 13.2  
Undecided: 14.7

6. The new 7-point rating scale provides more useful data for salary, promotion, and tenure decisions than the previous 5-point scales.

Agree: 42.0  
Disagree: 30.7  
Undecided: 27.3

7. Designation of items as "not applicable" should be decided in advance by the instructor rather than by individual students.

Agree: 53.4  
Disagree: 24.4  
Undecided: 22.1

8. Written comments provide a useful supplement to numerical ratings.

Agree: 72.3  
Disagree: 19.8  
Undecided: 7.9

9. Written comments provide useful feedback for improvement of teaching.

Agree: 59.8  
Disagree: 25.8  
Undecided: 14.4

10. I received abusive, obscene, or inappropriate written comments in last year's teaching evaluation.

Agree: 28.2  
Disagree: 65.0  
Undecided: 6.8

11. Written comments should be used as feedback to the instructor but not for decisions on faculty salary, promotion, and tenure.

Agree: 74.4  
Disagree: 18.4  
Undecided: 7.1

12. Individual faculty members should decide whether students complete the written comments section of the evaluation form.

Agree: 22.0  
Disagree: 64.9  
Undecided: 13.1

13. The written comments section should be deleted from the new evaluation form.

Agree: 16.1  
Disagree: 72.0

Undecided: 11.9

14. Faculty members should receive the results of annual teaching evaluations by June 1.

Agree: 90.4

Disagree: 5.0

Undecided: 4.6

15. Feedback to faculty should include teaching evaluation norms for different departments, course types, and class sizes.

Agree: 83.7

Disagree: 6.1

Undecided: 10.2

16. The new teaching evaluation form provides useful feedback for improvement of teaching and courses.

Agree: 64.2

Disagree: 16.4

Undecided: 19.4

17. The data provided by the new teaching evaluation form are suitable for use in salary, promotion, and tenure decisions.

Agree: 69.2

Disagree: 18.5

Undecided: 12.1

18. All things considered, the new teaching evaluation form is superior to evaluation forms used previously in my department.

Agree: 43.8

Disagree: 25.5

Undecided: 30.7

---

Table 4

Faculty Survey: Content Analysis of Supplementary Comments

---

Category Frequency

---

The new evaluation form focuses too much on instructor "classroom performance", and not enough on course quality and amount learned by students.	30
Some items on the new evaluation form are inappropriate for certain types of courses.	27
Feedback to instructors should occur earlier, especially for first-term courses.	26
Written comments from students are sometimes abusive,	19

rude, irresponsible, or contradictory.

Students should be required to sign their names on teaching evaluation forms. 17

Student evaluation of teaching is biased, invalid, or influenced too much by teacher "popularity". 13

The new teaching evaluation form favours a transmissive, lecture-style method of teaching, and thus is not suitable for use in all departments or all types of courses. 10

---

Note: Results are reported only for comment categories with response frequency of 10 or higher.

---

Table 5

Reliability of New Teaching Evaluation Form

(Sample: 116 psychology classes in 1995-96 academic year, 103 psychology classes in 1996-97 academic year; unit of analysis: class mean ratings)

---

Reliability Test Coefficient	Reliability
------------------------------	-------------

---

Mean interrater reliability (split-half) for instructor evaluation items #5 to 19 (N= 103)	.84 *
--	-------

Mean intercorrelation of instructor evaluation items # 5 to 18 (N=103)	.67*
--	------

Mean correlation of instructor evaluation items # 5 to 18 with overall evaluation item # 19 (N=103)	.80 *
---	-------

Correlation of overall instructor rating (item #19) with mean rating for all other items combined (N=103)	.95 *
---	-------

Correlation of overall instructor rating (item # 19) across different courses taught by same instructor in same academic year ( N= 87)	.62 *
Correlation of mean instructor rating on old vs. new teaching evaluation forms (all items combined): Same Course, Same Year (N= 25)	.95 *
Same Course, Successive Years (N=50)	.70 *
All Courses Combined, Successive Years (N= 36)	.72 *

---

\* Statistically significant at .05 level

---

Table 6

Correlation of Course and Instructor Characteristics with Overall Instructor Rating

(Sample:103 psychology classes, 1996-97 academic year; unit of analysis: class mean ratings)

---

Course or Instructor Characteristic Pearson Coorelation

---

Instructor Rank ( 1= Instructor...4= Full Professor)	.22 *
Instructor Gender ( 1= Male, 2= Female)	-.11
Course Level ( Year 1 ...Year 4)	.32 *
Time of Evaluation ( 1= Fall Term, 2= Spring Term)	-.07
Class Size ( Range = 5 to 1200)	-.03
Percentage of class present for evaluation	.39 *
Mean Percentage of classes attended (Item 1, 1= less than 20...5=more than 90)	.44 *
Mean Expected Grade in Course (Item 2, 1=F...5=A)	.40 *
Course Status (Item 3, 1=Optional, 2= Required)	-.06
Mean Initial Level of Interest in Course (Item 4, 1= Low, 2= Medium, 3=High)	.24 *

---

\* Statistically significant at .05 level

Table 7

## Distribution of Overall Instructor Ratings for Old vs. New Evaluation Forms

(Sample: 116 psychology classes in 1995-96 academic year, 103 psychology classes 1996-97 academic year; unit of analysis: class mean ratings)

## Old Form /New Form

(5-point scale, item #10, N=116) (7-point scale, item #19, N=103)

Rating	Frequency (%)	Rating	Frequency (%)
1.00 to 1.50 Poor	0.0	1.00 to 1.50 Very Poor	0.0
1.51 to 2.00	1.7	1.51 to 2.00	0.0
2.01 to 2.50 Satisfactory	4.3	2.01 to 2.50 Unsatisfactory	0.0
2.51 to 3.00	9.5	2.51 to 3.00	0.0
3.01 to 3.50 Good	19.0	3.01 to 3.50 Borderline	1.0
3.51 to 4.00	37.9	3.51 to 4.00	3.9
4.01 to 4.50 Very Good	22.4	4.01 to 4.50 Satisfactory	5.8
4.51 to 5.00 Outstanding	5.2	4.51 to 5.00	11.6
		5.01 to 5.50 Good	27.2
		5.51 to 6.00	28.2
		6.01 to 6.50 Very Good	15.5
		6.51 to 7.00 Outstanding	6.8
Mean	3.68	5.51	
Median	3.80	5.60	
Standard Deviation	0.64	0.74	
Skewness	-0.69	- 0.46	
Kurtosis	0.38	0.18	
% "satisfactory" or better	98.3	95.1	
% "good" or better	84.5	77.7	