

Office of Assessment

Fall 2006 Course Evaluations

Summary Results

Prepared by Kathryn Doherty

During the Fall 2006 semester, a course evaluation instrument was administered in nine departments of the College of Liberal Arts, the Fisher College of Science and Math, and the College of Health Professions: Biology, Chemistry, English, Environmental Science ESOL, Geography, Nursing, Psychology, and Sociology/Anthropology. This evaluation was in follow-up to the Fall 2004 and Spring 2006 course evaluation pilots. Analyses of student response and of the instrument suggest response results consistent with results from the Spring 2006 pilot and with research-based evidence on course evaluations and continue support both the validity and the reliability of the instrument itself.

Background

Using statistical analyses of the Spring 2006 results, faculty feedback, and Promotion and Tenure Committee input, the course evaluation instrument developed for the Spring 2006 pilot was revised for Fall 2006 administration. This instrument has 27 items: 15 course, instructor, and curriculum-related items; 3 overall rating items; 5 student demographic items; and 4 open-ended questions. The open-ended responses were not analyzed but were returned directly to faculty. Within the nine departments that participated, there were 242 course sections, 120 faculty, and 6113 registered students. The response rate was 78% or 4768 of the 6113 registered students. The following table presents this sample by department:

Department	Sections	Faculty	Registered	Responding	Response Rate
BIOL	6	3	244	154	73%
CHEM	5	2	100	70	70%
ENGL	191	78	4191	3279	78%
ENVS	2	2	41	39	95%
ESOL	1	1	11	9	82%
GEOG	5	5	389	284	73%
NURS	2	2	58	47	81%
PSYC	19	17	635	523	82%
SOAN	11	10	444	363	82%
TOTAL	242	120	6113	4768	78%

Analysis

Analysis of the Fall 2006 data was conducted in two parts. The first part was an analysis of the student responses. This analysis was conducted by course, by department, and by college. Summary reports are being prepared and will be distributed to the faculty who participated, to their Department Chairs, and to the CLA, FCSM, and CHP Deans. Faculty reports contain the instructor's results, overall department results, and overall college results. Chair reports contain course results for each participating course in the department, overall department results, and overall college results. The Deans' report contains overall results for each participating department and overall college results.

The second part of the analysis was a statistical analysis of the instrument itself. This consisted of factor analysis, reliability analysis, and comparison of means. The factor analysis was conducted to identify underlying variables, or factors, that explain the pattern of correlations within a set of observed variables. Reliability analysis is used to study the properties of the scales created through factor analysis and the items that make up these scales. Reliability analysis helps to determine whether responses across multiple items are consistent and also provides information about the relationships between individual items in the scale. Factor analysis and reliability analysis together help to determine the internal consistency of the instrument. Comparison of means is used to explore patterns of relationships between variables and to look at whether different groups or categories of respondents are responding in different ways. Each of these analyses is discussed below.

Factor Analysis

Factor analysis of the 18 non-demographic variables from the Fall 2006 course evaluation instrument was conducted using varimax extraction and a rotated correlation matrix solution that produced one component factor (or cluster of individual variables) containing all 18 non-demographic variables.

Factor 1: Pedagogy, Curriculum, and Student Learning		
I learned a great deal in this course		The instructor gave timely feedback on my academic performance
The instructor explained important ideas clearly		The instructor was accessible as outlined on the syllabus to help with this course
The instructor provided clear grading standards		The instructor challenged me to think in new ways
Methods for evaluating my work were applied fairly		The instructor encouraged my participation throughout the semester
Tests and/or assignments reflected the primary content of this courses as outlined on syllabus		The instructor held the class to high standards
Assignments asked me to integrate information form various sources		The instructor provided guidance in meeting those high standards
Course requirements challenged me intellectually		The course as a whole
I worked hard to meet the requirements of this course		The course content
Course objectives were clearly stated on the syllabus		The instructor

The emergence of one factor containing all 18 variables suggests that there is consistent correlation among the variables and that these variables together explain the variation in student response.

Reliability Analysis

The second analysis that was conducted on the evaluation data was a reliability analysis. Reliability analysis on the scale of the 18 non-demographic variables showed very high levels of reliability, using Cronbach's alpha to demonstrate internal consistency. The following table presents the alpha reliability.

Scaled Variable	Alpha Reliability
Pedagogy, Curriculum and Student Learning (all variables)	.95
Overall Items (three overall variables)	.94
Pedagogy, Curriculum and Student Learning without 3 overall items (15 variables)	.91

Alpha reliability results close to 1 indicate a strong relationship among individual variables within the scale and consistency across response on these items. Analysis of individual variables within each scale showed that in no case did the alpha reliability increase with the removal of any variable from the scale (with the exception of the item “I worked hard to meet the requirements of this course,” whose deletion increased the alpha reliability by 1/100th of a point); thus, the scaled variable was stronger by keeping each variable originally identified through factor analysis than by removing any one variable. These results support the reliability of the course evaluation instrument.

Comparison of Means

The third analysis that was conducted on the pilot course evaluation result was a comparison of mean response by demographic group. Results from these analyses support intuitive, anecdotal, and research-based analyses of expected student response, and thus support the validity of the evaluation instrument. Five demographic variables were identified as independent variables through which to compare student response to the three overall items on the evaluation: instructor, course, and content. Results from each are summarized below.

	Course	Content	Instructor
Overall Mean	4.07	4.01	4.25
Someone recommended course or instructor	4.45	4.43	4.62
I was interested in the topic	4.38	4.33	4.44
It was a requirement for my major	4.08	4.02	4.27
It filled a Gen Ed requirement	4.00	3.90	4.18
It fit my schedule	3.89	3.98	4.10

Looking at the mean response, where 5.00 is high (strongly agree, excellent, etc.), results using reason for enrollment as an independent variable highlight the differences in student evaluation of a course or instructor based on the reason the student is taking the course. All results are statistically significant at $p \leq 0.000$. As these results suggest and the research supports, students enrolled because of a recommendation or because of interest in the topic of the course were statistically more likely to give a favorable evaluation of the course, the content and the instructor than students who enrolled to fulfill a requirement or fit their schedule. These results should be considered when interpreting student response within individual courses.

When comparing means on the three overall variables and using current year in college as the independent variable, the following results (mean response) were evident:

	Course	Content	Instructor
Overall Mean	4.08	4.01	4.25
First Year Students	3.97	3.87	4.14
Sophomores	4.08	4.04	4.25
Juniors	4.19	4.13	4.33
Seniors	4.13	4.08	4.35

As with the results for main reason for enrollment, these results are not unexpected. Looking at the overall mean response for each overall variable (course, content, and instructor), first year student responses are less favorable than upper-class student response ($p < 0.000$), There are many ways that these data might be interpreted (e.g., student experience, expectations, understanding, etc.) and this should be taken into consideration when analyzing course evaluation results. These considerations become especially important in light of the number of first year students enrolled in general education courses.

The third independent variable used to compare the mean response on the three overall variables was expected grade in course. Again, results match anecdotal experience and are summarized in the following table:

	Course	Content	Instructor
Overall Mean	4.08	4.01	4.25
Grade of "A"	4.36	4.23	4.50
Grade of "B"	4.03	3.97	4.21
Grade of "C"	3.52	3.60	3.81
Grade of "D"*	2.89	3.26	3.32
Grade of "F"*	3.20	3.40	2.80

*Please note that the sample size for "D" and "F" grades is small compared to the overall sample so results should be retested in future administrations of the course evaluation.

Using an overall mean for each of the three overall variables, students expecting a grade of A are statistically more likely ($p \leq 0.000$) to respond favorably than students expecting a lower grade. These results should be considered when looking at overall course evaluation results for courses in which students typically receive lower grades (e.g., upper level courses, courses in specific disciplines, etc.).

Using attendance record for the course as the fourth independent variable when comparing mean response on the three overall variables, results indicated that students who attend class more often are statistically more likely ($p \leq 0.000$) to respond favorably on course evaluation items. These results are summarized below:

	Course	Content	Instructor
Overall Mean	4.08	4.01	4.25
Never missed a class	4.11	4.05	4.26
Missed 5 or more classes	3.81	3.80	4.08

Attendance results suggest that students who attend more often are more likely to offer positive evaluation of the course. These results should be considered in light of existing research and experience, as well as within attendance policy requirements and attendance records for individual courses.

The final independent variable used to compare mean response on the three overall variables was preparation time for the course. The following table summarizes these results:

	Course	Content	Instructor
Overall Mean	4.08	4.01	4.25
7 or more hours	3.93	3.92	4.08
3-4 hours preparation per week	4.17	4.11	4.31
0-1 hours preparation per week	3.75	3.68	4.00

When looking at student response, students who spend about 3-4 hours preparing for the course each week respond more favorably ($p \leq 0.000$) when evaluating a course. Students who spend 7 or more hours per week preparing for a course respond less favorably. This may suggest that students who are spending more hours than average studying for the course are possibly struggling with the course and may thus evaluate the course or the instructor less favorably. Students who spend 1 hour or less preparing for the course give the lowest response. Here, one might interpret these results as indicating students who are less engaged in a course are less likely to offer positive evaluations of the course. Additionally, 3-4 hours of preparation time may be a time commitment that fosters engagement without producing the overload of preparation time (7 hours or more) that may be associated with less student satisfaction with the course and the instructor.

Comparison with Spring 2006 Results (Mean Response)

The comparison of mean response on the Fall 2006 course evaluation supports anecdotal and intuitive understanding of the ways in which students respond on course evaluations and is consistent with student response for the Spring 2006 evaluations.

	Course Spring 06	Course Fall 06	Content Spring 06	Content Fall 06	Instructor Spring 06	Instructor Fall 06
Overall Mean	4.14	4.08	4.11	4.01	4.36	4.25
Someone recommended course or instructor	4.53	4.45	4.56	4.43	4.77	4.62
Gen Ed requirement	3.90	4.00	3.88	3.90	4.21	4.18
First Year students	3.91	3.97	3.96	3.87	4.15	4.14
Seniors	4.34	4.13	4.29	4.08	4.47	4.35
Grade of "A"	4.40	4.36	4.35	4.23	4.54	4.50
Grade of "C"	3.68	3.52	3.79	3.60	4.15	3.81
Never missed a class	4.22	4.11	4.21	4.05	4.40	4.26
Missed 5 or more classes	3.72	3.81	3.84	3.80	4.10	4.08
0-1 hour per week preparing	3.89	3.75	3.91	3.68	4.10	4.00
3-4 hours per week preparing	4.28	4.17	4.25	4.11	4.51	4.31

These results also support the research and the literature on course evaluations, which suggests that the ways in which students respond may be impacted by student demographic factors (i.e., reason for enrolling, expected grade, etc.). Results should be considered when reviewing the evaluation results from individual courses.

Summary

Nine departments, 120 faculty, 242 sections, and 4768 registered students participated in the Fall 2006 course evaluation project. Analysis of the data supports the validity and reliability of the instrument and the intuitive, anecdotal, and research-based experience related to student response on course evaluations. Factor analysis produced one strong cluster of variables that demonstrated high reliability and internal consistency. Comparison of means analyses point to statistically significant differences in student response based on reason for enrolling in the course, year in college, expected grade, attendance record, and hours of preparation per week. This course evaluation will be administered again during the Spring 2007 semester. Results will be analyzed and compared with the Spring 2006 and Fall 2006 findings.