Analysis of the Spring-2010 Course Evaluation Results
Warren Wilson College
Office of Institutional Research, Assessment & Planning
Don Ray – May 25, 2010

Overview

- An estimated 96 percent of all spring sections, eligible for using the evaluation form, were turned in and processed. Within these, the rate of enrolled student participation was approximately 80 percent. On average, 98 percent of the items on submitted forms were successfully completed (i.e., not left blank or incorrectly marked).

- As found for the previous two semesters, median ratings for all specific instructor characteristics (Items I_01 - I_19) were 5 on a five-point scale (“Almost Always”) as was the overall instructor rating (C_22; “Excellent”). In contrast, the overall median rating for the course was 4 (“Above Average”). Self-rating of student effort was also a 4 on the same scale.¹

- Results of the of the 22 objective rating items for Spring, 2010 were analyzed and compared with those from the two previous semester evaluations. From Figure 1, which lists responses to individual items in the order presented on the form, it can be seen that there is a rather dramatic correspondence of ratings on all items on the three semesters’ data. The results support the idea that the evaluation instrument is consistently tapping into the same student perceptions whenever it is used.

- Since the 19 items dealing with specific instructor characteristics (I_01 – I_19) were all based on the same 5-point rating scale, mean differences were statistically studied. The results of these strongly suggest the presence of two general groupings of items: those that fall above a rating of 4.40 (Items I-2, -5, -7, -9, -11, -14, -15, -16, and -19) and those that fall below that rating (Items I-1, -3, -4, -6, -8, -10, -12, -13, -17 and -18). Individual item differences between groups were significant whereas, differences within were not. This same pattern was observed across all three semesters. Institutionally, the higher item group indicates the areas in which faculty generally excel and the second group points to areas for improvement. The characterization of these two groups is described later.

- The three items concerning overall perception of instructor and course excellence, and student effort are based upon a different rating scale and therefore cannot be directly compared to the first 19 items. These three items are, however, highly, positively correlated with each other. High ratings of instructor are associated with high ratings of the course and with the effort students feel themselves to expend in that course. Low ratings on all three are also found. Interestingly, the relationship between instructor and course ratings is much stronger than any involving student effort. A possible explanation is that student engagement might also be dependent on influences outside of the classroom.

- Because of the major expressed interest in the three overall rating items (C_20, C_21, C_22), results of these were plotted as a function of the student’s professed reason for enrollment in the course (R_23). Also because of the similar response patterns across semesters, these data were combined in the analysis and plotted in Figure 2. Three major findings emerge from the graph: (1) Instructors are rated more highly

¹ While the scale used for items I_1 through I_19 seems to have face validity, interpretation of the three overall rating items (C_20, C_21, C_22) might require caution. An immediate question raised is “what might ‘average’ look like?” and “where would we have to go to find it?” “Average” certainly doesn’t exist at WWC and, in its absence, “above-,” and “below-average,” along with their trusty sidekicks “poor” and “excellent” seem equally ill-defined or, at best, misplaced.
than the courses they teach and both are significantly higher than the students self-rating of effort, and (2) students are universally most pleased with courses they are taking outside of their program requirements (“Other Electives”), and in descending order, in electives in their major, in their minor/concentration, and in required courses in their major, and, least impressed with courses taken for satisfying liberal arts requirements. An identical pattern is seen for ratings of instructors teaching those respective courses. In almost all cases, the decreases at each level are statistically significant. (3) Finally, students report putting forth most effort in elective courses (within major and outside of formal programs) and least in required liberal arts courses. A possible conclusion is that students find anything labeled as a requirement to be less engaging.

○ For the third time in as many semesters, factor analysis of the data revealed the same basic clustering of items. One last time (a promise), these factors or components are generally described as:

  - **Component 1**: Instructor ability to effectively impart skills/knowledge – consisting of items I-1, -2, -3, -4, -8, -9, -10, -12, -13, -18, -19, and C-20, -21 and -22.
  - **Component 2**: Level of organization inherent in the course – consisting of items I-5, -7, -11, and -14.
  - **Component 3**: Instructors empathy toward students – consisting of items I-6, -15, -16, and -17.

The characteristics of these components have been described in previous reports. However, something that is possibly of interest is the correspondence of items within these components and the item composition of high- and low-ranking groups described above (Figure 1). Approximately three-quarters of the low-rank items are found within Component 1. An equal percentage of high-rank items are distributed across Components 2 and 3. This seems to imply that students rank our faculty somewhat low on pedagogical skills (Component 1), but relatively high on course organization and ability to relate to students.
Figure 1. Means of response for objective items on evaluation forms over last three semesters (Spring, 2009; Fall, 2009; Spring, 2010). Note the general correspondence of each item rating across semesters. Notice, also, that, collapsing across all items, Spring, 2009 ratings are significantly higher than those for Fall, 2009. A similarly higher rating was found for Spring, 2010, over Fall, 2009, but this only achieved borderline reliability.
Figure 2. Overall ratings of Course, of Student Effort, and of Instructor as a function of Primary Reason for Course Enrollment.