ABSTRACT

Student evaluations of teaching (SET) need to be reliable, valid, and accurate because they are frequently used for high-stakes summative evaluation decisions about instructors, such as promotion, tenure, and merit pay. Although widely used by administrators, SET are often criticized by faculty as being inadequate measures of instructional effectiveness. However, the majority of researchers believe SET are generally reliable, valid, and worthwhile means of evaluating teaching despite evidence that some factors may influence SET independent of instructional effectiveness. One of these potential biasing factors is delivery method, and the literature contains indications of a SET bias against online instruction compared to face-to-face instruction. If such a bias exists, SET of online courses cannot be equitably compared with those of face-to-face courses. This case study used qualitative research techniques to answer the following research question: What are the differences in SET between online and face-to-face courses as evidenced by a thematic analysis of open-ended questions? Participants were students enrolled in 82 class sections taught by 41 instructors, one online and one face-to-face class section for each instructor, at Regent University, Virginia, during academic year 2004-05. Responses to open-ended SET questions were content analyzed and divided into 1,742 text segments. Each text segment was classified in two different sets of categories: appraisal and topical. Crosstabulation resulted in no significant difference in the proportion of appraisal text segments by delivery method. However, there were significant differences in the proportion of text segments for topical themes and topical categories by delivery method. Online students considered the course topical theme and organization and materials topical categories more important than face-to-face students.
Face-to-face students considered the instructor theme and person and knowledgeable categories more important than online students. MANOVA conducted on responses to closed-ended questions of overall evaluations of the instructor and course found no significant differences between delivery methods. Finally, responses to open- and closed-ended questions were correlated. Classes that provided negative criticism tended to give lower overall evaluations, and classes that commented about the instructor grading unfairly tended to give lower overall evaluations. Implications for research and practice are presented.