

Summary of Major Findings of the Three-Year Pilot of the SEEQ

Prepared by Dr. Dean McKay

Faculty Evaluation Committee

April 10, 2003

The SEEQ (Student Evaluation of Education Quality) has been used at Fordham for the past three years in a pilot period to determine whether the instrument may be viably administered on a permanent basis. The SEEQ was chosen because it has been adopted at many other selective institutions of higher education, as well as for its clarity, structure, and ease of administration. The University is committed to establishing objective evaluation data regarding its faculty for several reasons. First, faculty benefit from the feedback provided in structured and objective form, as courses are updated, or new courses developed. Second, student evaluations, when obtained from objective ratings, may be legitimately used as part of personnel and promotion decisions. And third, students have recently sought evaluative data on faculty, which is an advance over the widespread use of biased sources such as word of mouth or information from an unrepresentative sample of students. This summary is based on the full technical report provided by Thanos Patelis, Ph.D., of the College Board, and presented to the Faculty Evaluation Committee. Interested faculty may consult the full technical report by contacting the Faculty Evaluation Committee or the Dean's office.

An Overview of the SEEQ

The SEEQ was developed from the underlying premise that teaching is a complex and multifaceted activity. Accordingly, there are nine primary dimensions evaluated in the SEEQ. These are: Learning/Value, Instructor Enthusiasm, Organization/Clarity, Group Interaction, Individual Rapport, Breadth of Coverage, Examinations/Grading, Assignments/Reading, and Workload/Difficulty. A total rating is also provided. Supplemental items, both quantitative and qualitative, may be included that are germane to individual departments or instructors.

Summary of the Pilot Study:

Over the three-year pilot period of the SEEQ, there were between 18,000 and 20,000 responses, across 1100 to 1200 courses for the entire University. Overall, students gave high ratings to faculty across the nine dimensions of the SEEQ. Before summarizing the key aspects of the pilot data, a brief explanation of the key technical terms to be presented is in order:

1. The *reliability* of a measure is the extent that the items are related to one another, and therefore measuring the same concept. Reliability can be rated as poor, adequate, good, or excellent for each of the nine dimensions of the SEEQ.
2. The *validity* of an instrument is an indication that the measure evaluates what it is supposed to measure. The SEEQ dimensions were identified in the pilot data period in support of the validity of the measure.
3. The *mean* is equivalent to the average. Means are presented for each area of the SEEQ.

4. The *standard deviation* goes with the mean. This number represents the amount of variability around the mean. That is, for each mean, the ratings vary around that score at a rate that is expressed by the standard deviation.
5. Each of the nine dimensions of the SEEQ can be placed into percentiles based on the scores obtained within a larger group. These are called *norms*. In this case, each faculty score can be compared to their representative group, or norm group (that is, the entire full-time faculty of the University).
6. The information obtained from questionnaires may be examined to determine relations between ratings. *Correlations* are a standard way of showing how two scores are related; in this case, each of the nine dimensions of the SEEQ are related to each other.
7. Each dimension of the SEEQ is derived from a set of items that relate to that dimension. *Factor analysis* is a statistical procedure designed to test whether those items correctly group together to form the nine dimensions described in the test.

Test results of the SEEQ

Stability and norms

For the period of Fall 1999 through Fall 2002, faculty ratings for all nine areas assessed with the SEEQ were stable. That is, there was little fluctuation in overall assessment of faculty effectiveness for each semester of the pilot period. As indicated earlier, the ratings for all nine dimensions were, in general, high (Means are approximately 7 on a nine point scale, with higher scores indicating higher quality; the exception is workload/difficulty, which was stable at approximately 6 on a nine point scale, with high scores indicating greater difficulty and 5 indicating 'just right').

In general, ratings on the SEEQ that were culled from graduate courses were slightly higher than the evaluations from undergraduate courses. Because there is a significantly smaller graduate student body, and because graduate courses are not offered in all departments, it was recommended that ratings from graduate students be excluded from the normative data.

Correlations among the nine dimensions

While on the one hand ratings for the faculty remained high across all semesters of the pilot period, there were only modest correlations among the dimensions. Interestingly, the highest correlations were found between workload and the other dimensions, with greater workload leading to higher positive ratings for the course. It should be noted that correlations are not equivalent to causation, and we might expect that there is a point of diminishing return on the relation between workload and positive evaluation. There were also exceptions to this in some departments where there was no relation between workload and evaluation.

The SEEQ includes an item that requests students to forecast the grade anticipated in a course. This student expectation was not correlated with the quality of the course; that is, students expecting a higher grade did not rate courses higher than students who anticipated a lower grade.

Factor Analysis of the SEEQ

As a test of validity, factor analysis was used in analyzing the SEEQ pilot data. The results of this analysis showed that the proposed areas of assessment of the SEEQ were confirmed.

Reliability of the SEEQ dimensions

The reliability of the SEEQ in the pilot study was excellent for all nine dimensions. There was little variability in reliability across departments on 8 dimensions. The workload dimension was shown to have good reliability for all departments as well.

Supplemental Items

The SEEQ, as it is currently constructed, is a reliable and valid instrument assessing nine areas of instruction. Faculty or departments are encouraged to consider developing questions for supplemental inclusion that would provide feedback that is germane to their areas of specialty. This feature allows departments and faculty to customize the SEEQ to meet their individual needs.