

ACADEMIC QUALITY COMMITTEE
Report to the Senate Executive Committee
February 2012

The Academic Quality Committee submits the following consent item for consideration:

“The Academic Quality Committee recommends that we move our course evaluations online with Evaluation Kit in the fall of 2012. The instruments for these evaluations will be the current SITEs given in tandem with one of a number of other public domain instruments including the SEEQ (Students’ Evaluation of Educational Quality) that have already been tested for reliability and validity. Subsequent to a one year pilot testing period and an additional semester to process the results, the Faculty Senate will as early as possible during the 2013-2014 academic year recommend one instrument that will be retained for future course evaluations starting in the fall semester of 2014.”



Information Paper

Office of Institutional

Online Evaluation Vendors Overview

October 31, 2011

Vendors and Price Estimates:

EvaluationKIT: \$29,000
Online Course Eval: \$35,000-45,000
Blue/Evaluation: \$40,000
Digital Measures: \$25,000
What Do You Think: \$25,000
SIR II/eSIR: \$150,000
IDEA: \$45,000

All of the above-listed vendors offer the following features:

- Student access through BlackBoard, TopNet, and email links
- Capability of single login for all courses
- Ability to accommodate overlapping evaluation periods (bi-terms)
- Easy accommodation of cross-listed and team-taught courses
- Opt-out option for identified courses
- Reminder emails to students who haven't yet completed evaluations (without emailing students who have already completed all evaluations)
- Identification of respondents for incentive purposes (i.e., early access to grades)
- Ability to designate who can view results (instructors, department heads)
- Ability to customize with WKU logo
- Ability to block results when too few students respond
- Comments field option

With the exception of SIR II/eSIR, all vendors offer the following:

- Optional addition of an unlimited number of items by instructors/units, which can be saved by course from semester to semester

With the exception of IDEA/Student Ratings of Instruction, all vendors offer the following:

- Online report delivery, much sooner than with current SITE process

Recommendation:

IR recommends EvaluationKIT as vendor for online course evaluations. EvaluationKIT offers many user-friendly features that would be an improvement over our current process. For example, EvaluationKIT allows for immediate web access to results once the survey period has ended; the current SITE process requires several weeks from the end of the evaluation period to the distribution of reports. Reporting features allow instructors to view results in reports and bar charts, and to view summary reports for all courses taught. Instructors can even download results as raw data files.

Some other features that put EvaluationKIT above the others include user-friendly setup tools. Institutional hierarchies can be built that allow administrators to be easily added to view the appropriate results. A visual editor allows for easy setup of surveys with a variety of different question types. WKU would have complete control over dates and notifications within the system.

EvaluationKIT allows for a variety of different integration options. It can be used as a stand-alone, or integrated into Blackboard or Topnet. Students, instructors, and administrators would be able to seamlessly pass from one authenticated environment into EvaluationKIT with a single login.

While EvaluationKIT does not offer normative comparisons or a research-based group of items, those that do are limited relative to the benefits that EvaluationKIT offers. For example, while SIR II offers comparative data, the data come from a limited pool of schools using SIR II, which may not be a relevant comparison group. In addition, SIR II has rigid reporting features and limited customizability. For example, reports show how an instructor compares to others in the comparison group, but only for those items where the instructor is beyond the 90th or 10th percentile. In addition, departmental items cannot be added, and only 10 university items can be added.

In terms of cost, EvaluationKIT would provide expanded functionality and reach at equivalent (or slightly lower) cost than the current SITE process.

Vendor Details:

EvaluationKIT: <http://www.evaluationkit.com/>

- Pros:
 - Get results immediately
 - Free pilot available
 - Quoted \$28,800 cost estimate
 - Some nice features on our end – set up system so that evaluations automatically go out a certain number of days before course ends (bi-terms, summer courses, etc)
 - Variety of access options, including email and blackboard integration; can be single login
 - We can upload a file with course/student information, or integrate with blackboard (someone from IT would have to connect SIS and EvaluationKIT)
 - Reports: Can have comparisons in reports (department or university), descriptive statistics, mean, standard deviation, %, bar charts, raw data, separate reports can be made for department, university, courses

What Do You Think, College Net: http://corp.collegenet.com/products/wdyt_what.html

- Pros:
 - About \$24,000
 - Can create link on portal and have single log in
 - Implementation involves pilot
 - We submit a file containing course/instructor/student information
 - Reports: Can have department/university comparisons, can break down by demographic characteristics, can filter by “strengths” and “weaknesses”
- Cons:
 - Fewer integration options
 - Less user-friendly interface for building projects and hierarchies

IDEA, Student Ratings of Instruction: <http://www.theideacenter.org/>

- Pros:
 - Reported validity and reliability
- Cons:
 - Pay per student enrollment + per class + charge for reports (approximately \$44,000/year)
 - Instructor must fill out a form (invited by email) of their course objectives
 - Reports are sent as PDF on CD and 2 hard copies (possibly slower than current SITE process)

BLUE/Evaluation: http://www.explorance.com/prod_evaluation.asp

- Pros:
 - Connects directly to our servers for course/instructor/student information and user authentication
 - Can conduct analysis by any variable in our system
 - Reports: Customizable, various metrics reported, trend analysis
- Cons:
 - Wouldn't give a cost estimate, UofL is paying \$200,000 this year, after that \$40-50k
 - Complicated, not very user-friendly

Online Course Evaluations: <http://service.onlinecourseevaluations.com/>

- Pros:
 - Provide us with portfolio of best practices for response rate
 - Can track response rate in real time, even instructors can see response rate and go in and tell their students
- Cons:
 - \$35k-\$45k per year with 4 year commitment (includes set-up price which is about \$35k)
 - Out-dated interface
 - Complicated features that we would not want or need (tenure helper, etc)
 - Not user-friendly

Course Response, Digital Measures: <http://www.digitalmeasures.com/CourseResponse/>

- Pros:
 - WKU is already a DM client (Activity Insight)
 - Cost is about \$25,000 per year (including discounts for multiple colleges)
 - Access through email, portal, or blackboard; can be single login
 - We submit a file containing course/instructor/student information
 - Review response rates in real time
 - Reports: Can create custom reports with things like department/univ comparisons, and more
- Cons:
 - Fewer integration options
 - Less user-friendly interface for building projects and hierarchies
 - Offer customizable reports but likely would take weeks to receive

ETS—SIR II Online and eSIR: http://www.ets.org/sir_ii/administration/online_delivery

- Pros:
 - Research based, comparative data
 - Different version developed specifically for web courses
 - Online report delivery
- Cons:
 - \$150,000 per year
 - Rigid reporting features
 - Normative groups not extensive nor likely relevant

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.