UNIVERSITY LEADERSHIP COUNCIL

Gathering and Applying Meaningful Data from Student Course Evaluations

Custom Research Brief • April 28, 2010

RESEARCH ASSOCIATE
Anne Lewandowski

RESEARCH MANAGER
Keely Bielat

TABLE OF CONTENTS

THE ADVISORY BOARD COMPANY
WASHINGTON, D.C.

I. Research Methodology

Project Challenge:

Leadership at a member institution approached the Council with the following questions:

- *What instruments do peer institutions use for course evaluation? Do different divisions within the institution use different instruments, and are these instruments developed internally or by a third party?*
- *What are the objectives for evaluations? How do different parties within the university use the information obtained from student course evaluations?*
- *What processes do institutions use to ensure high response rates and also protect confidentiality when administering evaluations?*
- *How do faculty view the evaluation process? To what extent are evaluations used in tenure and promotion decisions?*

Project Sources:

- Advisory Board's internal and online (www.educationadvisoryboard.com) research libraries
- *The Chronicle of Higher Education* http://chronicle.com
- National Center for Education Statistics (NCES) http://nces.ed.gov/
- Council for the Advancement of Standards in Higher Education, CAS Professional Standards for Higher Education (6th Edition)

Research Parameters:

- Per the requesting member's guidelines, the Council targeted its outreach to private institutions with student populations between 5,000 and 15,000.
- Contacts generally serve as vice provost of faculty affairs.

I. Research Methodology

A Guide to Institutions Profiled in this Brief			
Institution	**Location**	**Enrollment (Total / Undergraduate)**	**Classification**
University A	South City: Midsize	14,500/12,200	Private: Research Universities (high research activity)
University B	Midwest Suburb: Large	8,300/5,800	Private: Master's Colleges and Universities (larger programs)
College C	Midwest City: Large	12,500/11,900	Private: Master's Colleges and Universities (medium programs)
University D	West City: Large	8,300/7,100	Private: Research Universities (high research activity)
University E	Northeast Suburb: Large	12,300/8,300	Private: Doctoral/Research Universities
College F	Northeast Suburb: Large	3,300/2,200	Private: Master's Colleges and Universities (larger programs)
University G	Northeast City: Large	10,100/6,600	Private: Doctoral/Research Universities
University H	West City: Large	11,200/5,800	Private: Master's Colleges and Universities (larger programs)
University I	Midwest City: Large	5,300/3,200	Private: Master's Colleges and Universities (larger programs)
University J	Midwest City: Large	5,700/3,300	Private: Master's Colleges and Universities (larger programs)

Source: National Center for Education Statistics

II. Executive Summary

Key Observations:

- **Across contact institutions, the overarching objective of course evaluations is to assess teaching quality.** However, student course evaluations never count for a certain weight of a faculty member's review and are always considered alongside peer reviews, observations by supervisors, and self-assessments for promotion and tenure decisions. Contacts also note that course evaluations are not used to assess administrators or staff.

- **In general, contact institutions use home-grown evaluation instruments instead of externally-developed instruments in order to collect data that appropriately reflects the individual institution's unique learning objectives.** Although third parties may offer products that guarantee a certain response rate or provide national benchmarking data, most contacts are satisfied with their home-grown evaluation instruments.

- **While many institutions are attempting to standardize evaluation instruments across different divisions and courses, contacts note that it is important for departments and schools to have the ability to customize course instruments.** To gather data that can be compared across the institution while reflecting the distinct objectives and pedagogical styles of different disciplines, many institutions use instruments with both a standardized set of questions and a customizable section.

- **Many institutions are exploring using online evaluations to reduce processing time and paper waste; however, contacts uniformly report difficulties in ensuring that online evaluations will yield the same high response rates as paper evaluations.** Improved evaluation design, more consistent communication from faculty and administrators, and clearer incentives for completing course evaluations may increase student participation.

- **While confidentiality is a continual concern, contact institutions with paper evaluations all report success with traditional methods of evaluation collection (i.e., requiring a student to collect and turn in all paper evaluations in a sealed envelope).** Contact institutions also place restrictions on who may view evaluation results; in most cases, only faculty members and their department heads may view final results, which are often presented in aggregate. For online evaluations, only the system administrator typically has access to identifying information and individual responses are encrypted.

- **Although evaluations with low response rates may suffer from selection bias (i.e., only students with strong feelings about a faculty member respond), these evaluations can still provide useful information, such as suggestions for improving the course and records of extremely positive or negative performances.** Questions regarding students' motivations for taking the course may provide context for interpreting these students' responses, and some institutions publish disclaimers on evaluation results with low response rates to encourage decision-makers to use data with caution.

- **Faculty frequently resist changes to course evaluation procedures due to concerns that evaluations will inaccurately reflect their teaching quality and will outweigh other factors in personnel decisions.** Institutions can assuage faculty concerns by improving communication about evaluation content, working to raise student response rates to provide a more thorough data sample, and clarifying how student course evaluations factor into tenure and promotion decisions.

III. Structure of Evaluation Instruments

Structuring Effective Evaluations

Although most contacts hope to standardize course evaluations, many institutions currently administer multiple course evaluations in different formats for different divisions. The table below outlines basic characteristics of course evaluations across contact institutions.

Overview of Evaluations at Contact Institutions			
Institution	**Third Parties Provide Evaluation**	**Uniform Evaluation Across Institution**	**Evaluations Administered Online**
College C	In process of adopting third-party evaluations	✓	✓
University D	In process of adopting third-party evaluations	✓	For online courses
College F	IDEA*	✗	✗
University G	IDEA*	✗	For online courses
University A	✗	✓	✗
University B	✗	✓	For online courses
University E	✗	✓	✗
University H	✗	✗	✗
University I	✗	✓	Piloting online evaluations
University J	✗	✓	✓

In the above chart, a "✓" indicates that the institution uses the specified practice; an "✗" indicates that the institution does not use the practice.

**These institutions use the IDEA Center survey (described in more detail in Section IV) in addition to internally-developed evaluation instruments.*

III. Structure of Evaluation Instruments

Homegrown Evaluation Instruments

Despite the growing prominence of evaluation instruments developed by third-party vendors, most contact institutions use internally-developed instruments (and sometimes multiple instruments across the institution) to conserve resources and to tailor evaluations more closely to unique institutional needs.

Collecting Meaningful Data with Homegrown Evaluations

The biggest challenge faced by institutions with homegrown evaluation tools is collecting data that can be compared for faculty members and courses across the institution. Easily comparable data yields clear insights into faculty members' teaching over time, trends in different departments and divisions (i.e., whether students consistently give math courses lower ratings than humanities courses), and context for individual evaluations. Strategies for collecting more useful, comparable data from student course evaluations are outlined below.

Quantify Data Whenever Possible

Likert and multiple-choice responses can be quantified more easily than open-ended responses, facilitating comparisons across faculty members, disciplines, and semesters. Multiple-choice and Likert questions can also obtain student feedback directly pertaining to specific criteria for measuring teaching and course quality (e.g., communication skills, clarity and fairness of grading, etc.). However, contacts warn that some faculty, particularly in writing-intensive disciplines, may resist increased collection of quantitative data. To address these faculty concerns and collect additional information, most evaluation instruments include some qualitative, open-ended questions in addition to Likert and multiple choice questions.

Collect Contextual Information

To provide more context for data gathered from course evaluations, **College C**'s current evaluation instrument includes a section addressing student motivation for taking the class. Questions in this section explore why a student chose to enroll in the course and students' approximate attendance rate. Meanwhile, course evaluations at **University J** collect information on student GPA. These questions help to profile individual students in the course, providing helpful context when analyzing students' impressions of teaching quality.

Provide Limited Customizable Options

As institutions become more centralized and pursue more interdisciplinary academic activities, many universities are trending away from allowing divisions to use unique evaluation instruments. Instead, institutions with diverse academic offerings, including **University A** and **University G**, are increasingly creating standardized evaluations that have customizable options. Divisions with a unique pedagogical style (e.g., schools of music) may add or alter questions to more specifically address distinct teaching methods or learning outcomes for the schools; this option ensures utility to a specific program while still allowing for the collection of comparable data across the institution.

III. Structure of Evaluation Instruments

Common Evaluation Design Flaws

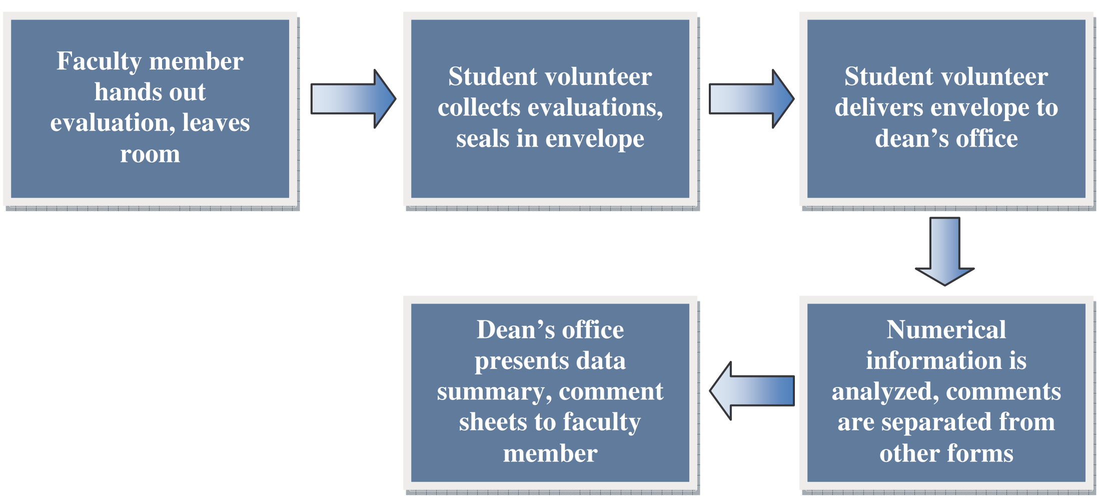
Many evaluation instruments contain avoidable design flaws that reduce the usefulness and completeness of the data. Three of the most common flaws are highlighted below:

- **Too many questions:** Too many questions on the course evaluation form may frustrate students and result in rushed, careless answers. **College C** is working to reduce the length of its online evaluation instrument from the current 32 questions in order to reduce the burden on students' time.
- **Too many formats:** In addition to yielding data that is difficult to compare across units, multiple evaluations with different formats can confuse or frustrate students, particularly when evaluations are administered online. Contacts emphasize that students are more likely to fill out evaluations thoughtfully and completely when instruments are easy to use.
- **Too many scoring scales:** When evaluations use inconsistent scoring scales (i.e., some questions ask for a rating from 1-4, others from 1-10), rankings become more difficult to compare. All evaluations across the institution should use one uniform scoring scale that remains constant over time to facilitate comparisons of data.

Typical Procedures for Hard Copy Evaluation Administration

Evaluations are generally administered in the two weeks of courses directly before final exams and final calculations of grades; contacts emphasize that all divisions within an institution should administer evaluations within the same timeframe to avoid confusion among students. This is especially true for online evaluations, which give rise to additional concerns about timing (see page 10).

Most institutions which administer hard copy evaluations to students in class use the following procedure to ensure confidentiality of information:



While this process generally provides sufficient confidentiality for students to make comments, students in small classes may feel uncomfortable providing written comments that their professors will see. To ensure confidentiality for these students, staff at **University B** and **University I** type all hand-written comments. At **College C**, faculty do not receive comment sheets for classes where fewer than five students comment.

III. STRUCTURE OF EVALUATION INSTRUMENTS

Exploring Online Evaluation Instruments

While most institutions currently administer online evaluations only for online courses, many institutions are increasingly seeking to implement online evaluations for all courses for the following reasons:

- **Online results are easier and quicker to process.**
- **Online evaluations are cheaper and "greener" than paper.**
- **Online formats can be easier for faculty to customize.**

However, faculty may resist adopting online evaluations because of low response rates; several institutions reports response rates of 20 to 30 percent for online evaluations (as opposed to 85 percent or higher response rates for hard copy evaluations administered in class). As a result of faculty reservations, several institutions are piloting online evaluations for divisions or certain courses before implementing them institution-wide. One institution's pilot program is highlighted below.

Spotlight: Online Evaluation Pilot Programs

University I began an online evaluation pilot program last spring to ease the institution's transition into using online evaluations for all courses. Administrators hope that the change will reduce paper usage and reduce the amount of time spent analyzing evaluation (results from hardcopy evaluations currently require six weeks to be returned to faculty members).

Faculty initially resisted the implementation of online evaluations due to concerns that response rates would drop from their currently high levels (over 90 percent response rate) to 70 percent or lower. To win faculty support, the administration agreed to pilot online evaluations with willing faculty members; when pilot program evaluations consistently receive a 79 percent response rate, faculty will approve the transition to institutional online evaluations.

Although faculty buy-in for the pilot program has been slow, faculty interest is increasing because of the promise of more timely evaluation results; additionally, the online evaluations give individual faculty members opportunities to customize questions to their particular courses.

III. Structure of Evaluation Instruments

Solving the Response Rate Problem

Although hard copy evaluations administered in class rarely yield a 100 percent response rate (due to class absences and students' leaving questions blank), in-class paper evaluations consistently yield higher response rates than online evaluations.

The design of online evaluations frequently creates barriers that make filling out evaluations more difficult and time-consuming for students. Furthermore, poor communication and few incentives for participation will drive down online response rates. Common problems with online course evaluations are described below.

Response Rate Reducers

- **Online evaluations are too hard to navigate.** For example, **College C**'s initial attempt at an online evaluation form did not feature a "back" button, making it difficult for students to edit or rethink responses.
- **Online evaluations are too long.** As previously mentioned, students might not be willing to spend their own time filling out a lengthy questionnaire. While these students may turn in partially-complete paper evaluations administered in class, they may not submit an online evaluation at all.
- **Evaluations may open or close at different times.** When divisions administer online evaluations at different times (i.e., one division may require all evaluations to be submitted before final exams, while another division does not require evaluations to be submitted until after exams), students are more likely to miss submission deadlines.
- **Students may have greater confidentiality concerns.** When students watch another student seal course evaluations in an envelope, they can feel reasonably secure in the confidentiality of their responses. Online evaluations, however, might strike students as less secure, particularly if evaluations must be submitted before final grades are reported.
- **Faculty members may fail to notify students of evaluations.** When faculty are no longer responsible for administering evaluations in class, they may forget to remind students that evaluations are open and that their feedback is important.

III. Structure of Evaluation Instruments

Although the above factors may have a dramatic negative impact on response rates, these issues are not insurmountable. Contacts suggest that improved communication, streamlined Web design, and stronger incentives for students are key to improving response rates for online evaluations.

Improved Communication

Since online evaluations require students to use their own time, consistent reminders are essential to high response rates. The dean's office or center for teaching will frequently take the responsibility for sending email reminders and evaluation links to students to encourage participation.

Streamlined Web Design

As contacts at **University B** and **College C** note, response rates will be higher when evaluations are integrated with online learning portals (i.e., Blackboard). In addition to reducing a barrier to access, these portals can deliver pop-up reminders to students who have not completed evaluations.

Stronger Incentives

Institutions have begun to offer stronger incentives to students for participation in online evaluations. Allowing students to see course grades early is the most common incentive. Some institutions are considering releasing some evaluation results to students who have participated, but faculty generally oppose this idea.

Minimum Valid Response Rates

Low response rates are generally viewed as problematic because selection bias frequently distorts results of these evaluations: when few students participate, it is likely that only the students who feel very strongly about a teacher respond. Despite the problem posed by low response rates, no contacts reported throwing out evaluation results where only a very small percentage of students responded.

Contacts generally suggest that a response rate between 60 and 70 percent would be considered acceptable for evaluation. Many contacts note that since student evaluations are never the only criterion used to assess a faculty member's teaching quality, potentially distorted data will ultimately be viewed in context with peer and supervisor observations and other measures of teaching quality. To ensure that evaluations with low response rates are considered fairly, **University E** publishes a warning on result reports where the response rate was very low or where there were few students in a class.

IV. Considering Third-Party Vendors

Weighing the Benefits of Third-Party Instruments

Although most contact institutions rely on internally-developed evaluation instruments, instruments from third-party vendors are increasing in popularity because they can reduce the time and energy institutions must expend to collect and analyze data. Externally-developed instruments used by many different institutions can also facilitate teaching quality comparisons across different institutions, providing a new resource to institutional research offices. However, many institutions hesitate to adopt third-party instruments because of concerns that these instruments may not be sufficiently customizable or may lack specificity. Institutions that are satisfied with their homegrown instruments rarely see reason to purchase an instrument.

> *"Is the standardized instrument going to be that much better that it pays to break with your current system? Obviously there's going to be a break and obviously you're not going to be able to track the same data the same way that you were tracking it before. There's going to need to be a whole new education program put into place so that the faculty, deans, and other personnel committees understand it, so the question always is, is the standardized instrument so much better to be worth it?"*
>
> -Council Interview

Several contact institutions are in the process of adopting or piloting externally-developed instruments. Promising instruments and their potential benefits are described in the following table (please note that **College C** and **University D** are still in the process of considering these instruments and have not yet implemented them).

Third-Party Instruments Used or Considered by Contact Institutions

Instrument	Institution	Characteristics
OnlineCourse Evaluations.com	• University D	• Guarantees minimum response rate (Web site reports that customers average a 72% response rate) • Surveys can be designed to ask follow-up questions • Tracks and surveys students who drop courses • Compatible with BlackBoard
CoursEval 3.0	• College C	• Provides simple, easy-to-understand result reports • Compatible with online information environment (Jenzabar); single log-in for students • Easily customizable to different courses • Can allow students who fill out evaluations to see grades earlier and have access to select evaluation results
IDEA Center Survey	• College F • University G	• Measures learning objectives to assess teaching quality • Customizable to different departments' learning objectives • Data comparable with national database • Contacts at University G report that the tool is ill-suited to measure learning outcomes for non-traditional disciplines (e.g., music, fashion, etc.)

IV. Considering Third-Party Vendors

As described in *The Chronicle of Higher Education*'s article, "Rating Your Professors: Scholars Text Improved Course Evaluations," a major advantage of externally-developed instruments, such as the IDEA system, is that users of these instruments may access national data. In addition to serving a benchmarking purpose, aggregate data from institutions over many years (since 1975 in the case of the IDEA system) can be used to correct for potential biases in evaluations. For example, IDEA analysts have determined a method to adjust teacher evaluations based on whether a course is required of non-majors.[1]

Alternate Method of External Evaluations

All programs in **College F**'s school of health sciences are accredited by national accrediting bodies; as a result, evaluations for individual programs are tailored to the particular accrediting body's needs. To supplement student and peer evaluations of program quality, the institution regularly administers evaluations to alumni so they can assess program impact on their careers. These evaluations enable the institution to adjust programming to students' professional needs, instead of responding solely to student impressions of teacher quality. Faculty who are hesitant to rely heavily on student evaluations may prefer the use of these alumni surveys as a means to improve teaching quality and learning outcomes.

[1] David Glenn, "Rating Your Professors: Scholars Test Improved Course Evaluations," in *The Chronicle of Higher Education*, 25 April 2010.

V. Applying Course Evaluation Results

Using Evaluations to Improve Student Learning

Well-designed and -administered course evaluation instruments can yield a wealth of information for institutional decision-makers. However, to avoid overwhelming both students and the office of institutional research, institutions must limit information gathered through semesterly course evaluations. No contact institutions use course evaluations to assess administrators or departmental staff (although other evaluations may be used for this purpose), and few use course evaluations to assess student learning outcomes. Although many contacts express interest in including more student learning outcomes-related content in future evaluations, student self-assessments in evaluations would only serve as indirect measures of outcome achievement.

The Role of Student Course Evaluations in Faculty Reviews

Changes to course evaluations and their use in promotion and tenure proceedings frequently draw the ire of the faculty. A minority of professors view students as illegitimate judges of teaching quality (for example, Harvard professor Harvey C. Mansfield has said student evaluations "subject the wise to the judgment and scrutiny of the unwise"[2]). More commonly, faculty worry that they will suffer from selection bias when evaluations yield low response rates.

Contacts report that faculty concerns over low response rates are valid, as students with strong opinions (particularly negative) are more likely to respond. Furthermore, several contacts note that student evaluations generally highlight very good and very poor teachers but rarely differentiate among teachers of similar quality. As a result of these concerns, **no contact institution counts student course evaluations as a certain percentage of a faculty member's review.** Although institutions still consider student evaluations to be meaningful indicators of teacher quality, many factors play a role in promotion and tenure decisions, including self-assessments, peer reviews, and observations by supervisors. Examples of how contact institutions ultimately use student course evaluations are highlighted below.

University E

- Center for teaching excellence provides feedback on indicators where faculty request extra help, videotape classes
- Numerical results made public, increasing student buy-in
- Comments not released publicly
- Faculty may choose which comments to present to personnel

University G

- Evaluations reviewed by faculty members, deans, and department chairs
- Used minimally in personnel decisions
- Chairs discuss consistently poor evaluations with faculty members

University D

- Deans receive course evaluations and present them to faculty, must report if faculty reviewed evaluations
- Faculty who are not responsive to poor evaluations are recommended for professional development
- Faculty not required to include course evaluations in personnel dossiers

[2] Quoted in "Faculty Passes Q Reforms" by Maxwell L. Child and Christian B. Flow in *The Harvard Crimson*, 13 February 2008.

V. Applying Course Evaluation Results

Contacts note that communication with faculty about the format and use of course evaluations is essential to making successful changes to evaluations and incorporating these assessments into the institution's culture.

> *"I think a lot of [faculty pushback] is based on fear that they [faculty] do not really know how evaluations are being used to make decisions about tenure and promotion, so one of the things I'm really going to talk about as we move into the new system is the primary reason why course evaluations exist is to give individual feedback to faculty about individual courses. It is used to make other decisions, such as hiring decisions for part-time faculty and tenure and promotion, but these are departmental decisions. I am going to put on our Web site, 'If you want to know more about how these are being used, you need to talk to your department chair.'"*
>
> -Council Interview

Professional Services Note

The Advisory Board has worked to ensure the accuracy of the information it provides to its members. This project relies on data obtained from many sources, however, and The Advisory Board cannot guarantee the accuracy of the information or its analysis in all cases. Further, The Advisory Board is not engaged in rendering clinical, legal, accounting, or other professional services. Its projects should not be construed as professional advice on any particular set of facts or circumstances. Members are advised to consult with their staff and senior management, or other appropriate professionals, prior to implementing any changes based on this project. Neither The Advisory Board Company nor its programs are responsible for any claims or losses that may arise from any errors or omissions in their projects, whether caused by the Advisory Board Company or its sources.