

COURSE INSTRUCTOR FEEDBACK (CIF)

A MAJOR UPGRADE TO THE TEACHER COURSE EVALUATION (TCE) SYSTEM

TABLE OF CONTENTS

I. Introduction	1
II. Revision of Questions/Items for the CIF System.....	1
III. Timetable for Implementation of CIF.....	4
IV. Bibliography	4
Appendix A. Online Administration of the CIF System.....	6
Appendix B. Rationale for Each New Item/Question in CIF	9
Appendix C. Discontinued TCE Items	15
Appendix D. Experience at Northwestern University	16

This initiative has been led by the

ADVISORY COMMITTEE TO THE PROVOST ON THE EVALUATION OF TEACHING

(Dennis Jacobs, Chair; Kevin Barry; Philippe Collon; Kathleen Eberhard; Pat Flynn; Tom Frecka;
John Gaski; Jill Godmilow; Mark Gunty, IR consultant; Patricia Maurice; John Robinson)

Approved by Academic Council on April 16, 2008

I. Introduction

The University of Notre Dame is committed to providing an outstanding educational experience for its undergraduate and graduate students. To advance this goal, the University strives continuously to recruit, cultivate, effectively assess, recognize and reward faculty members who are both highly effective teachers and superb scholars. In 2007, the *Advisory Committee to the Provost on the Evaluation of Teaching* (ACPET) published a set of guidelines to help departments conduct a more comprehensive evaluation of a faculty member's teaching to inform renewal, tenure and promotion decisions. This expanded review calls for faculty peers to assess in-depth a set of representative courses taught recently by the candidate and to appraise the ways in which the candidate enriches the student learning experience over and above his/her conventional teaching responsibilities.

Beyond the occasional comprehensive review described above, the University invites students each semester to report on their experiences in the classroom. Student ratings of teaching can serve three important functions:

- *Formative* (providing feedback to the instructor for the improvement of teaching)
- *Evaluative* (contributing to an overall assessment of the instructor's effectiveness as a teacher)
- *Analytical* (helping the institution determine what factors are most often associated with perceptions of effective teaching)

The Provost's Office launched the paper-based Teacher Course Evaluation (TCE) system in 1982. Since then, four different TCE forms have been used, with the latest set of changes instituted in 1997. Over the past decade, many faculty have registered complaints about the current set of TCE items; criticisms were directed at the wording of specific TCE items, difficulties in interpreting the "improvement needed" response scale, and the inability of the TCE to effectively address the diverse modes of teaching/learning at Notre Dame.

Following an extensive study, ACPET recommends that (i) the University revise the TCE questions to provide more useful feedback for instructors and more reliable and valid measures of teaching effectiveness, and (ii) that students no longer complete a paper-based TCE form in class but respond online to a series of questions customized for each course. The completely redesigned instrument is being titled *Course Instructor Feedback* (CIF) to signify the important role that students play in guiding the improvement of teaching. Appendix A outlines the benefits and potential concerns associated with transitioning the paper-based TCE to an online survey format.

II. Revision of Questions/Items for the CIF System

The web-based delivery of the CIF questionnaire enables students to view and respond to a set of questions that can be unique to each course. Balancing the evaluative and formative aspects of the CIF instrument, some items will be standardized across the university, while other items will be customizable by the instructor.

The CIF items/questions will be organized into four tiers:

- **University** – Sixteen multiple-choice items/questions will be asked of students in every course at the University. By having a common set of items across the University, the average score for each item in any given course can be compared with the average computed for all courses at the department, college, or university levels. Eleven of the items within the university tier are designed to inform the evaluation of teaching.
- **Learning Goals** – It has been shown that a student's self-reported progress in achieving course learning goals is modestly correlated with independent measures of the student's learning and the instructor's perceived teaching effectiveness (Feldman, 1996). Because each course is designed with a particular set of learning goals in mind, the CIF system will allow each instructor or academic program to select six course-specific learning goals from an extensive library of learning goals; students can then rate their progress towards each goal.
- **Section** – The section tier is optional for instructors. An instructor can preview online the questions that will be asked of his/her students. If the instructor would like to add a limited number of additional questions, s/he can select from an online question bank and/or enter customized questions (open-ended or closed-ended). The responses to section-specific questions will be reported back to the instructor, but they will not be used to inform personnel decisions (tenure, promotion, merit raise, etc.).
- **Informing Course Selection (ICS)** – In an effort to provide students with more comprehensive and accurate information at the time of course selection, a web-based ICS system will be launched in Fall 2008. Included in the ICS system will be histogram plots of the student responses to 5 closed-ended questions, drafted by students & faculty and approved by Academic Council. The responses to these items will not be used to inform personnel decisions. The 5 ICS items will appear on the CIF form for all courses.

ACPET has drafted a brief set of open-ended questions that are designed to solicit formative feedback for the improvement of teaching. ACPET has also drafted a set of closed-ended items, for the university tier of the CIF system, based on constructs that are correlated with student exam performance and overall perception of effective teaching. Appendix B provides a detailed rationale for each of the latter items along with the relevant correlation coefficients. Appendix C lists each item on the current TCE for which there is no counterpart in the new CIF system. Note that the 6-point response scale is different than that on the current TCE. Pilot testing confirmed that the distribution of student responses on the CIF is less compressed at the high end of the scale.

Open-ended Items in University Tier:

- i. Please comment on how well the activities, readings, lectures, and assignments helped you learn in this course.
- ii. Please identify what you perceive to be the greatest strengths of this instructor's teaching.
- iii. Please identify areas where this instructor could improve his/her teaching.

Multiple Choice Items in University Tier:

Rate this course on the following. Please consider your overall experience in the course, rather than isolated incidents.

- A. Excellent
- B. Very Good
- C. Good
- D. Satisfactory
- E. Poor
- F. Very Poor
- G. Not Applicable

1. Overall organization of the course
2. Availability of appropriate help or learning resources outside class
3. Helpfulness of required assignments (readings, projects, etc.) in facilitating my learning
4. Usefulness of the feedback I received concerning my work in the course
5. Instructor's preparation for each class
6. Instructor's clarity of communication
7. Instructor's fairness and impartiality in conducting the class
8. Instructor's effort to help students develop mastery of the course material
9. Instructor's stimulation of my interest in the subject matter
10. Overall promotion of my creative, analytical or critical thinking
11. Overall teaching effectiveness of this instructor
12. The degree of intellectual challenge I experienced in the course was
 - A. Extremely high
 - B. Very high
 - C. Somewhat high
 - D. Moderate
 - E. Somewhat low
 - F. Very low
13. The percentage of this course's regular class meetings I attended this semester was approximately
 - A. 50% or less
 - B. 60%
 - C. 70%
 - D. 80%
 - E. 90%
 - F. 95%+
 - G. Not applicable
14. I spent approximately ___ hours per week, on average, doing work for this course outside regularly scheduled class time.
 - A. 0-1
 - B. 2-3
 - C. 4-5
 - D. 6-7
 - E. 8-10
 - F. 11-14
 - G. 15 or more
15. I am taking this course as an
 - A. Explicit requirement with my major, minor or concentration
 - B. Elective that fulfills a requirement within my major, minor or concentration
 - C. College requirement
 - D. University requirement
 - E. Free elective
16. I anticipate that my final grade in this course will be
 - A. A
 - B. A-
 - C. B+
 - D. B
 - E. B-
 - F. C+
 - G. C or lower

Rate your progress toward the following learning goals, taking course level into account.

17 - 22. *Items 17-21 represent different learning goals relevant to the course.*

- A. Excellent
- B. Very Good
- C. Good
- D. Satisfactory
- E. Poor
- F. Very Poor
- G. Not applicable

III. Timetable for Implementation of CIF

The University has contracted Gap Technologies to administer the new online Course Instructor Feedback (CIF) system using a product called Online Course Evaluations. University implementation of the CIF system consists of instrument development followed by deployment of CIF within Gap's software.

INSTRUMENT DEVELOPMENT

<u>Time frame</u>	<u>Action</u>
Summer 2007	Limited pilot testing of University tier (16 sections) With in-classroom follow-up interviews
Fall 2007	Refinement of University tier Second pilot test in 90 sections (2460 students), dual administration method (paper and online)
Spring and Summer 2008	Pilot testing with online-only administration University tier: multiple choice and open-ended items Limited testing of Learning Goals tier

GAP DEPLOYMENT

<u>Time frame</u>	<u>Action</u>
Spring 2008, end of semester	Limited pilot test of CIF online
Summer 2008	Full pilot test of CIF online
Fall 2008	Full-scale implementation of CIF online

IV. Bibliography

- Arreola, Raoul. 2000. *Developing a Comprehensive Faculty Evaluation System, Second Edition*. Bolton, MA: Anker Publishing.
- Berk, Ronald A. 2006. *Thirteen Strategies to Measure College Teaching*. Sterling, VA: Stylus Publishing.
- Brasskamp, L. A. and J. C. Ory. 1994. *Assessing Faculty Work: Enhancing Individual and Institutional Performance*. San Francisco: Jossey-Bass.
- Brookfield, Stephen D. 2006. *The Skillful Teacher: On Technique, trust, and Responsiveness in the Classroom, Second Edition*. San Francisco: Jossey-Bass.
- Bush, L., Maid, B., and Roen, D. 2002. In: Wehlburg, C. and S. Chadwick-Blossey, editors, *To Improve the Academy*. Bolton, MA: Anker Publishing, pp. 312-315.
- Cashin, William E. 1995. "Student Ratings of Teaching: The Research Revisited." The Center for Faculty Evaluation and Development, Kansas State University. IDEA Paper No. 32.

- Centra, John A. 1993. *Reflective Faculty Evaluation: Enhancing Teaching and Determining Faculty Effectiveness*. San Francisco: Jossey-Bass.
- Centra, John A. 1996. "Identifying Exemplary Teachers: Evidence from Colleagues, Administrators, and Alumni." *New Directions for Teaching and Learning* 65:51-64.
- Chism, N. 1999. *Peer Review of Teaching: A Source Book*. Bolton, MA: Anker Publishing.
- Feldman, Kenneth A. 1988. "Effective College Teaching from the Students' and Faculty's View: Matched or Mismatched Priorities?" *Research in Higher Education*, 28(4): 291-344.
- Feldman, Kenneth A. 1996. "Identifying Exemplary Teaching: Using Data from Course and Teacher Evaluations." *New Directions for Teaching and Learning*, 65: 41-50.
- Fink, L. Dee. 2006. "Improving the Evaluation of College Teaching." University of Oklahoma Program for Instructional Innovation.
- Gillmore, Gerald M. 2001. "What Student Ratings Results Tell Us about Academic Demands and Expectations." Office of Educational Assessment, University of Washington.
- Hardy, Nedra. 2003. "Online Ratings: Fact and Fiction." *New Directions for Teaching and Learning* 96: 31-38.
- Heines, Jesse M., and David M. Marting, Jr. 2004. "Development and Deployment of a Web-based Course Evaluation System: Trying to Satisfy the Faculty, the Students, the Administration, and the Union." University of Massachusetts Lowell.
- Leadership Development Program, University of California Berkeley. 2005. "Improving the Process of Course Evaluation: The Online Alternative for Berkeley."
- Marks, Ronald B. 2000. "Determinants of Student Evaluations of Global Measures of Instructor and Course Value." *Journal of Marketing Education*, 22(2): 108-119.
- McKeachie, Wilbert. "Student Ratings: The Validity of Use." *American Psychologist*, 52(11): 1218-1225.
- Nuhfer, Edward B. 2003. "Of What Value Are Student Evaluations?" Center for Teaching and Learning, Idaho State University.
- Sorenson, D. Lynn and Trav D. Johnson, editors. 2003. *Online Student Ratings of Instruction, New Directions for Teaching and Learning*. San Francisco: Jossey-Bass.
- Theall, Michael. 1999. "New Directions for Theory and Research on Teaching: A Review of the Past Twenty Years." *New Directions for Teaching and Learning*, 80: 29-52.
- Weinberg, Bruce A., Belton M. Fleischer, and Masanori Hashimoto. 2007. "Evaluating Methods for Evaluating Instruction: The Case of Higher Education." National Bureau of Economic Research, working paper 12844.
- Young, C., and L. Young. 1999. *Journal of Excellence in College Teaching*. 10:63,76.

Appendix A. Online Administration of the CIF System

In considering how best to administer the survey through which students provide feedback on courses, there are two natural choices: (i) paper forms distributed and collected in class, and (ii) a web-based form that students complete online. There are several benefits and potential concerns associated with transitioning from paper-based to online administration of the TCE or CIF.

BENEFITS

The University community may realize the following benefits:

- The CIF can be administered online without sacrificing 10-20 minutes of class time for distribution, completion and collection of the paper-based TCEs.
- Online administration of the CIF would prevent the possibility that an instructor could refuse to participate. In Fall 2006, for example, TCEs were not collected in 181 class sections that were designated for evaluation.
- NetID-limited access to online CIF submission would guarantee that only students enrolled in a class could submit a CIF for that course. Currently, any student who happens to be present on the day that the TCEs are administered can complete one or more forms; conversely, any student who is absent that day does not have the opportunity to complete a paper-based TCE.
- Online administration of the CIF will eliminate the need for a student from each class to collect the TCE forms and deliver them to a drop box. Removing this step from the process will prevent forms from being lost and will guarantee that no student from the class can review or alter the responses of his/her classmates.
- Online administration of the CIF will eliminate the need for 60,000 paper TCEs to be (i) hand-counted, packaged and mailed to instructors three weeks before the semester ends, (ii) collected and hand-sorted by IR staff after the TCEs have been completed, and (iii) sent off campus to a 3rd-party scanner.
- The period over which students can complete the online CIF will span from ten days before the last day of class through the last reading day.
- Students can complete the free-response questions to the online TCE without concern that their handwriting will be recognized by the instructor.
- Students would not have to repeatedly answer five demographic questions for every class, because this information could be accessed from institutional data.
- An online CIF system can incorporate a customized set of questions that would complement the items utilized across the University. Because a portion of the questions will be tailored to selected characteristics of the course, each instructor will receive more relevant feedback about his/her courses. Similarly, students will appreciate responding to a set of CIF questions that are better matched to the course they have just experienced.

CONCERNS

Two concerns that are most often expressed with regards to online versus paper-based administration of student ratings are (i) the fear of low response rates and of the bias this may introduce, and (ii) the potential loss of confidentiality for students who complete the CIF.

In reflecting upon the concern of low **response rates**, it is important to keep in mind that response rates with the paper-based TCE system average about 82% in courses for which forms are received and about 78% overall, if we include all sections that were designated for evaluation.

There is no precise threshold for acceptable response rates, but it is important that two conditions are met. First, response rates should bolster overall reliability by minimizing random error. That is, enough students should respond so that faculty and administrators have confidence that the sampling error is small. Second, response rates should bolster overall validity by minimizing systematic error. That is, there should be little to no difference between the kinds of students who do and do not respond so as to not introduce response bias. If, for example, students who were dissatisfied with a course were less likely to respond, the results would be artificially inflated.

Fortunately, we have data from several sources to indicate that changes in response rate as Notre Dame moves online are unlikely to undermine either reliability or validity. Several top-20 research universities have moved successfully to an online student feedback system, including Harvard, Northwestern, UC Berkeley, Stanford, Vanderbilt, Washington University and Yale. Many have enjoyed 70-85% response rates. At Stanford, response rates online actually increased compared to their paper system, and they were achieving high 80s by their third round of online administration. Strategies have been identified by these institutions (see below) to enhance response rates. Moreover, during the pilot test of the online CIF in fall 2007 at Notre Dame, in which 90 sections participated, the average response rate was 73%. This rate was achieved without any incentives, after the students had already been asked to complete paper TCEs in class, and in a shorter time frame (just seven days) than we will use when CIF is implemented.

With regard to **response bias**, we again have data from other institutions and from our own pilot test. Research conducted at Northwestern (Hardy, 2003: 34) when they moved online indicated that, comparing paper and online feedback in the same classes taught by the same instructors using the same questions, 42% of the ratings were higher online, 46% of the ratings were higher on paper, and 12% were mixed. Hardy's conclusion: "The ratings may be lower or higher or the same." The report from Northwestern (Appendix D) provides succinct and useful insights into the very concerns expressed by many faculty at Notre Dame.

Notre Dame's fall 2007 pilot data does not allow for a direct comparison of the paper and online results because different questions were asked in the two modes, and the response scales were also different. The two global items (TCE Q17 and CIF Q11) had a correlation of 0.87.

Institutional Research will monitor the response rates for individual sections and employ various incentives campus-wide to enhance student response rates. Some of the **successful strategies** that other institutions have adopted to ensure high response rates include:

1. Promote an ethic where students take seriously their responsibility to provide feedback to their instructors (All institutions)
2. Students receive frequent reminder emails from the dean, until they have completed their course evaluations. (Vanderbilt)
3. Students receive points in their courses for completing their course evaluations (Washington Univ.)
4. Students are entered into a lottery when they complete their course evaluations. (Georgia Tech gives away iPods; Stanford uses iTunes giveaways).
5. Students can see their grades online only after they have completed their course evaluations (Yale and Stanford)
6. A student can view the student ratings compiled for other courses only if she has completed the evaluations for all of her own courses from the past semester (Northwestern)

Notre Dame plans to adopt a hybrid between the first, second and sixth strategies. Those students who complete their CIFs in a given semester gain unrestricted access to the *Informing Course Selection* reports during the following registration period. Noncompliant students can view only the Instructor- and Registrar-provided portions within the *Informing Course Selection* reports. Discussions with the Academic Affairs Committee of Student Government during Spring 2008 indicate strong support for all three of these strategies, along with consideration of the possibility of a modified version of regulating early access to grades dependent on completion of CIFs. Evidence from the fall 2007 pilot also indicates the more actively that faculty encourage participation, the higher the response rate.

The concern about **student confidentiality** can be addressed in more concrete ways than can concerns about response rates. Because the CIF system must match students with courses, complete anonymity is not possible. A student will login to the CIF system by entering his/her netID and password. The student will then be presented with a list of all the courses in which he/she is enrolled. Clicking on each course link will deliver a customized CIF questionnaire for the student to complete online. When all of the student's forms have been completed or the system is closed for the semester, the identifier is de-linked from the data base. *No unique identifying information for the student will be stored with the student's CIF responses for a course.* Gap will run the analysis to determine which students have completed all of their CIFs and will return to Notre Dame a data base with a Yes/No marker for whether each student has met the participation criteria. That file will be used to determine each student's level of access to the *Informing Course Selection* views.

Appendix B. Rationale for Each New Item/Question in CIF

The rationale for each multiple-choice item in the CIF university-tier includes the following:

- Explanation of why a student’s response to the item might be coupled to the student’s achievement in the course and to his/her perception of the instructor’s effectiveness.
- Correlation coefficients reveal the strength of association between the students’ rating of the specific item/construct and either their exam performance in the course or their perception of the instructor’s teaching effectiveness. The Pearson’s correlation coefficient (r) measures the strength and direction of a linear relationship between two variables. An r value of zero indicates no correlation between the two variables, whereas an r value of exactly +1 indicates a perfect positive linear fit.
- Intended uses for results gathered for a specific CIF item may include the following:
 - *Formative* (providing feedback to the instructor for the improvement of teaching)
 - *Evaluative* (contributing to an overall assessment of the instructor’s effectiveness as a teacher)
 - *Analytical* (helping the institution determine what factors are most often associated with perceptions of effective teaching)

1: Overall organization of the course		
Relationship to Student Exam Performance	When students are aware of the organization of a course they can better see the big picture and gauge their own progress. To the extent that learning in a particular area needs to be sequential, course organization facilitates development of conceptual understanding and skills.	$r = .57$ ¹
Relationship to overall perception of effective teaching	A well-organized course, as perceived by the students, helps them see the instructor's plan and is evidence of a significant preparation effort.	$r = .73$ ²
Intended Use of Item	Formative , Evaluative , Analytical	
Similar item(s) on former TCE	#2 - The Course is well organized and its goals clearly established.	

¹ Through a meta-analysis across several studies and involving thousands of students, Feldman measured correlation coefficients (r) between student ratings for specific constructs and student performance as measured by a final examination for the same course. Feldman, K. "Identifying Exemplary Teaching: Using Data From Course and Teacher Evaluations." in Svinicki, M.D. and Menges, R.J. (eds.) *Honoring Exemplary Teaching*, New Directions for Teaching and Learning, No. 65, Spring, 1996, San Francisco: Jossey-Bass Publishers, Inc.

² Correlation coefficients are derived from online pilot of new items conducted at Notre Dame in Fall 2007 (N = 2460 students). The correlation coefficient (r) indicates the relationship between the students’ responses to this specific item and their overall perception of the instructor’s teaching (Item #11).

2: Availability of appropriate help or learning resources outside class		
Relationship to Student Exam Performance	Students are more likely to succeed when they have access to appropriate resources and help from the professor or TAs as needed.	r = .36 ¹
Relationship to overall perception of effective teaching	When students perceive that they are provided the help and resources that maximize their success they are more likely to consider themselves well taught.	r = .62 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#8 – Help is available to students outside of class.	

3: Helpfulness of required assignments (readings, projects, etc.) in facilitating my learning		
Relationship to Student Exam Performance	Clearly thought-out assignments that engage students in meaningful tasks and connect with course goals help students perform at a high level.	---
Relationship to overall perception of effective teaching	When students recognize the contribution of the assignments to their learning they don't feel like they are being asked to do "busy work" and are likely to see the work as part of a well designed learning environment.	r = .67 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	No similar item	

4: Usefulness of the feedback I received concerning my work in the course		
Relationship to Student Exam Performance	Prompt informative feedback extends the learning experience of students beyond the initial completion of an assignment, increases engagement, and focuses time on task.	r = .23 ¹
Relationship to overall perception of effective teaching	Useful feedback is an important part of the learning process and values the work that students have done. When promptly and consistently received, it contributes to the students' sense of being well taught.	r = .62 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#9 – The instructor's evaluations of my work are helpful.	

5: Instructor's preparation for each class		
Relationship to Student Exam Performance	A well-prepared instructor can maximize the impact of time spent with students and of their progress toward successful achievement of course goals.	r = .57 ¹
Relationship to overall perception of effective teaching	When the instructor is well prepared, it sends the message that teaching is considered to be an important activity and that students' time is valuable.	r = .73 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#1 – The instructor is well-prepared for each class.	

6: Instructor's clarity of communication		
Relationship to Student Exam Performance	Clear communication facilitates student understanding of what is expected of them as well as their learning/development of the ideas, concepts or skills associated with the course.	r = .56 ¹
Relationship to overall perception of effective teaching	If communication is not clear students will be frustrated in their attempts to learn and this will reflect negatively on their perception of effective teaching. Many students interpret lack of clarity with lack of concern about the entire teaching activity. Particularly in lecture settings, clarity is the strongest component of the student-teacher relationship.	r = .81 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#3 – The class material is clearly presented.	

7: Instructor's fairness and impartiality in conducting the class		
Relationship to Student Exam Performance	This construct affects student achievement more in the negative than the positive; i.e., unfairness and partiality are likely to discourage students' best efforts.	r = .26 ¹
Relationship to overall perception of effective teaching	Students perception of effective teaching will be reduced both by perceptions of unfairness (e.g. tests that don't reflect preparation) and by perceptions of partiality by the instructor.	r = .65 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#10 – The instructor's dealings with students are fair and impartial.	

8: Instructor's effort to help students develop mastery of the course material		
Relationship to Student Exam Performance	Students are motivated by the instructor's concern for their learning. This motivation, along with the additional support that might come from a concerned instructor, can lead to higher levels of achievement.	r = .30 ¹
Relationship to overall perception of effective teaching	Demonstrable interest in student learning is a key factor when students evaluate teaching effectiveness.	r = .77 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#12 – The instructor shows care for students' learning.	

9: Instructor's stimulation of my interest in the subject matter		
Relationship to Student Exam Performance	Students with higher interest levels are likely to be more engaged and spend more time on task leading to higher levels of achievement.	r = .38 ¹
Relationship to overall perception of effective teaching	Students understand that some subjects will not interest them as much as others, but see the instructor's efforts to stimulate interest as part of an overall effort to fully engage the students.	r = .77 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#11 – The instructor is enthusiastic about the subject matter.	

10: Instructor's promotion of my creative, analytical or critical thinking		
Relationship to Student Exam Performance	Creative, analytical or critical thinking lead to deeper and more persistent learning than does rote memorization.	r = .25 ¹
Relationship to overall perception of effective teaching	Particularly in settings other than lectures, students equate the promotion of deeper thinking with the essence of good teaching.	r = .77 ²
Intended Use of Item	Formative, Evaluative , Analytical	
Similar item(s) on former TCE	#4 – The instructor stimulates creative or analytical thinking.	

11: Overall teaching effectiveness of this instructor		
Relationship to Student Exam Performance	Students are more committed to studying for classes in which they believe they are being well-taught.	---
Relationship to overall perception of effective teaching	This item allows the student to communicate an overall perception of effective teaching that may reflect constructs beyond those explicitly identified in items #1-11.	r = 1.0 ²
Intended Use of Item	Formative , Evaluative , Analytical	
Similar item(s) on former TCE	#17 – Now please evaluate only the quality of the instructor’s teaching.	

12: The degree of intellectual challenge I experienced in the course was ...		
Relationship to Student Exam Performance	Students study harder and report learning more in courses that provide an intellectual challenge.	r = .25 ¹
Relationship to overall perception of effective teaching	Students respect instructors who challenge them and who express high expectations for their learning.	r = .31 ²
Intended Use of Item	Formative , Analytical	
Similar item(s) on former TCE	No Similar Item	

13: The percentage of this course's regular class meetings I attended this semester was approximately ...		
Relationship to Student Exam Performance	If participating in class adds great value to student learning, then excessive absences may reduce student achievement in the course.	---
Relationship to overall perception of effective teaching	If a student finds an instructor to be a particularly ineffective teacher, s/he is more likely to skip class.	r = .16 ²
Intended Use of Item	Formative , Analytical	
Similar item(s) on former TCE	No Similar Item	

14: I spent approximately ___ hours per week, on average, doing work for this course outside regularly scheduled class time.		
Relationship to Student Exam Performance	If assignments outside of class are designed to promote effective learning, then time spent outside of class should correlate with student achievement	r = .09 ¹
Relationship to overall perception of effective teaching	The relationship between time students spend on the course, and evaluations seems to depend on the nature of the time and motivations involved. If a student likes the course content, or if the student expects an A or B in the course, more time is generally associated with a higher overall perception of effective teaching.	r = -0.02 ²
Intended Use of Item	Formative , Analytical	
Similar item(s) on former TCE	#15 – In comparison with other courses I have taken at Notre Dame, the amount of time I spend on this course is:	

15: I am taking this course as a ... (requirement, elective, etc.)		
Relationship to Student Achievement	The reason why a student is taking a course may influence the student's commitment to learning in the course.	---
Relationship to overall perception of effective teaching	A student's perception of an instructor's teaching effectiveness may depend on the extent to which the student sees the course as integral or beneficial to his/her chosen program of study.	---
Purpose of Item	Formative , Analytical	
Similar item(s) on former TCE	Course is: (a) in major and required (b) in major but not required (c) not in major but required (d) not in major and not required.	

16: I anticipate that my final grade in this course will be ...		
Relationship to Student Exam Performance	A student's self-assessment of his/her performance in a course is often similar to the grade assigned by the instructor.	---
Relationship to overall perception of effective teaching	Students who perform well in a course are more likely to judge the instructor to be an effective teacher.	r = .19 ²
Purpose of Item	Formative , Analytical	
Similar item(s) on former TCE	I anticipate that my final grade in this course will be ...	

17-22: Rate your progress toward the following learning goals, taking course level into account. (<i>Items 17-22 represent different learning goals relevant to the course.</i>)		
Relationship to Student Exam Performance	A student's self-assessment of his/her own learning is reasonably correlated with an objective assessment of the student's learning gains.	r = .46 ¹
Relationship to overall perception of effective teaching	Students who believe they have learned a great deal in a course are more likely to judge the instructor to be an effective teacher.	---
Intended Use of Item	Research indicates that students' perception of progress toward learning goals is an important factor in their overall perception of teaching. Notre Dame's Office of Institutional Research will conduct detailed analyses of the CIF data from Fall 2008 to determine whether the results from the learning goals tier (Items 17-22) have characteristics useful for informing the evaluation of teaching at Notre Dame. Until the Academic Council reviews the analyses, the learning goals tier should be used only for Formative purposes.	
Similar item(s) on former TCE	#19. How would you rate your progress toward this goal as a result of this course?	

Appendix C. Discontinued TCE Items

The following TCE items are being discontinued and have no comparable item included in the university-tier of the CIF:

TCE Item

#5 – The examinations appropriately test my mastery of the course material.

Explanation for omission – Faculty peers are in a better position than students to judge the appropriateness of exams. See ACPET Guidelines – <http://provost.nd.edu/academic-resources-and-information/>

#7 – When asked questions the instructor satisfies the students.

Explanation for omission – The respondent should not be asked to judge whether other students are satisfied.

#14 – On at least one occasion, I had a discussion with my teacher outside class.

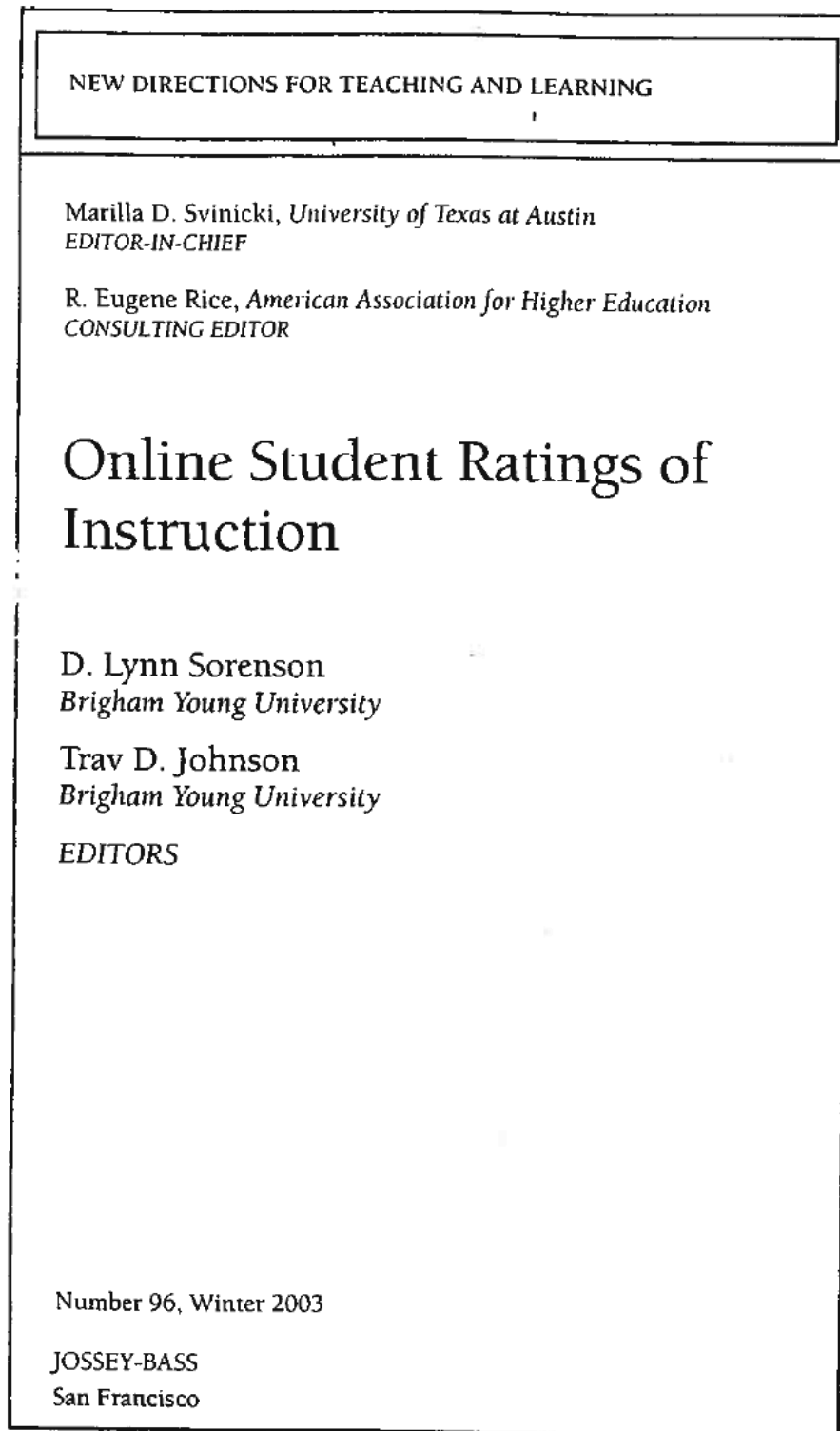
Explanation for omission – The response to this item may be more indicative of the student's initiative, or lack thereof, than the instructor's openness.

#16 – In evaluating any course, you can consider the course content alone or what the instructor does with the material. Please evaluate the course content alone.

Explanation for omission – Faculty peers are in a better position than students to judge the appropriateness of course content. See ACPET Guidelines – <http://provost.nd.edu/academic-resources-and-information/>

Appendix D. Experience at Northwestern University

The following book chapter outlines Northwestern's experience moving from paper to online student evaluations. Many faculty at Northwestern had concerns similar to those expressed by Notre Dame faculty.



improved this process; evaluation information is now available to instructors as soon as they have submitted their grades for the quarter to the registrar's office. Through the years, numerical scores and student comments have been made available to the university community—in the past, with an annual publication the students could purchase and most recently by online postings.

For years, students have used evaluation data during registration to assist them in selecting their courses and instructors. Now, using the online system, students sit at their computers to register for their classes with one window open to the registration site and another open to the evaluation site. In addition, instructors rely heavily on their student evaluations to improve their teaching and restructure their courses; administrators rely on the evaluations to assist in curricular changes and in making promotion and tenure decisions.

Although the initial idea of online student evaluations sent waves of fear across campus, there were also strong advocates who wanted to see online student ratings eventually replace the cumbersome paper-collection method. One of the most appealing aspects of the online system was the immediate availability of results. Many instructors also welcomed recovering class time formerly used in administering evaluations.

Amid both support and resistance, Northwestern University began collecting most evaluations online for the spring quarter of the 1999–2000 academic year. At that time, new policy dictated that all courses be evaluated online unless instructors specifically requested the traditional paper evaluation forms for their classes. The new online evaluation form mirrors the paper evaluation form; it has the same questions and format, which allows integration of the evaluation reports from both systems. Since implementation of the new system, students have become convinced that online evaluations may actually be more anonymous than submitting their handwritten comments on the paper forms. Students seem to prefer submitting evaluations online rather than writing them on paper.

Student Rating Scores

During the early implementation phase of the Web-based system, perceptions (and misperceptions) surfaced about online ratings of instruction, some of which are discussed below.

PERCEPTION: Results will be (or are) lower for online ratings than for paper ratings.

REALITY: The ratings may be (somewhat) lower.

One professor on campus contacted CTEC with a request for data to do a study in which he might show that online ratings are lower than paper-pencil ratings. The CTEC office complied and supplied him with data collected from online and from paper forms for fall quarter 1999 through fall

Online Ratings: Fact and Fiction

Nedra Hardy

When Northwestern University (Evanston, Illinois) put forth the idea of using the Web to collect student ratings of instruction, the proposal was met with fear and skepticism by most segments of the university community. Some felt that the time-honored course evaluations would be desecrated by using the Web. Students seemed certain they would lose the anonymity they had always presumed when completing their evaluations on paper. Instructors and administrators worried that students would not complete their evaluations unless they were held captive in class to do so. Across campus, there was a generalized faculty fear that the only students who would respond online were those with strong negative opinions of a class or instructor.

Background

At Northwestern University, evaluations have been collected by the Course and Teacher Evaluation Council (CTEC) for more than twenty years. Although these evaluations can be customized (for example, by instructor, department, or college), the following items are common to all questionnaires: five core (general) questions, four open-ended questions (for example, strengths, weaknesses, or suggestions), and the "summary comments" in which students summarize their responses to the four open-ended questions. These paper rating forms were used by Northwestern University before 1999, when experimentation with online collection began. By that time, data entry had progressed from manual entry to the use of scan forms. Processing the paper evaluations—including scanning numerical responses, typing the summary comments, and reporting results—was a cumbersome process that took from ten to twelve weeks. The new online system has

quarter 2000 (the first four quarters in which online ratings were used). The study was conducted during the summer of 2001 and included data from 5,112 classes (2,457 classes had used paper forms, and 2,655 classes had used online forms).

The comparison of the data indicated that the overall online scores were 0.25 of a point lower than the paper scores on the six-point scale of 1 (low) to 6 (high). This professor then requested that identifiers be included in the study because he believed the difference in scores would become substantially larger if further analyses were done comparing online versions with paper versions from the exact same classes (same instructor, same course), including evaluations done both ways. The effect of the second study was exactly the same: 0.25 of a point lower (on a six-point scale) for online evaluations.

REALITY: The ratings may be lower or higher or the same.

To further examine the relationship of online and paper ratings, the CTEC conducted its own comparison studies—one at the end of fall quarter 2001 and another after spring quarter 2002. For this study, CTEC examined twenty-six classes in which the same instructor taught the same class multiple times, using both paper and online evaluations. The comparison involved the five core questions on the evaluation form, which are general questions asked about every class at Northwestern. The group of classes used for the study is representative of all the undergraduate schools within the university, with class sizes ranging from four to 634 students.

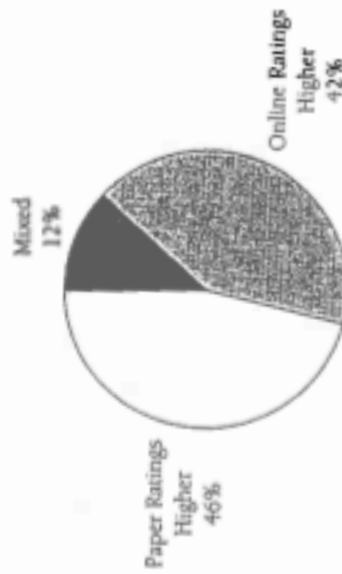
The twenty-six classes in the comparison study were taught a total of 274 times between fall quarter 1996 and spring quarter 2002. Paper forms were used to evaluate 135 classes, and online evaluations were conducted on 139 classes. Each class was evaluated with only paper or only online evaluations; collection methods were never mixed within the same class. Every course was evaluated at least twice—and some up to thirty-seven times—using one or the other collection method.

Among the courses included in the study, eleven average scores from online evaluations were higher than the average scores from paper evaluations (on the five core items). Twelve courses' average scores from paper evaluations were higher than the average scores from the online ratings (on the five core items). For three of the courses, the average scores were mixed, with some of the items rated higher on paper and some rated higher online (see Figure 3.1).

Student Comments

PERCEPTION: "[I prefer using paper forms because] people are more likely to write comments [on paper], and the response rate is everyone in class. . . ." [Comment by a teacher of a large introductory course, as quoted in the student newspaper]

Figure 3.1. Comparison of Results for Online and Paper Ratings, Same Class Taught by Same Instructor



REALITY: One large introductory course registered little difference between the number and length of student comments, whether online or on paper.

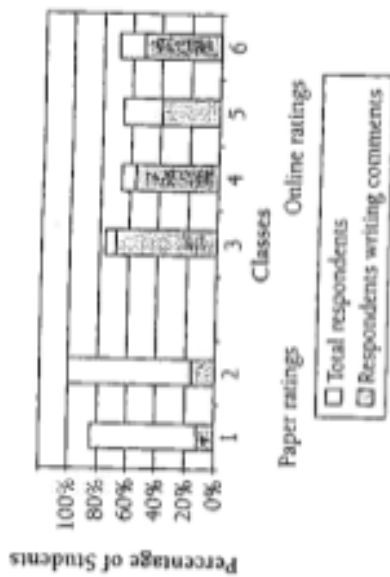
In a review of seven sections of this large introductory course (ranging in size from 432 to 634 students), five of the classes were evaluated using paper, and two were evaluated online. Both the number of students commenting and the length of student comments were comparable for paper and online evaluations. In fact, the above-mentioned instructor usually elicited extensive favorable comments, regardless of the method of data collection. (The overall response rate for paper evaluations in this teacher's course was 54 percent, and the overall response rate for online evaluations was 48.5 percent.)

REALITY: Substantially more written comments are submitted online.

In another study, six classes with similar content were examined within the same department. The enrollment of each class was between thirty-nine and forty-three students. Four of the classes were evaluated online, and two were evaluated using paper forms. First, the overall response rates of all classes were compared. (The overall response rate included all students who responded to any portion of an evaluation for the class.) Then, the number of student comments was compared. The summary comments seem to be the most important to students because they are posted on the Web for university-wide review. A compilation of these summary comments is also included with the numerical report that is returned to the instructor, the department chair, and the dean.

One class using paper forms had an enrollment of forty-three students. This class had a response rate of 100 percent; all forty-three students responded to at least some part of the rating form. However, only seven of the

Figure 3.2. Response Rates for Classes Evaluated on Paper and Those Evaluated Online



forty-three responding students (16 percent) wrote comments. These seven students wrote a total of nine lines of text about the class. The second class using paper evaluation had responses from thirty-three of the forty students enrolled, for an overall response rate of 83 percent. Only four of those thirty-three students (12 percent) wrote summary comments, generating a total of six lines of text.

Then the classes using paper rating forms were compared with the classes evaluated online. The overall response rate in the classes evaluated online was somewhat lower. However, students wrote many more comments when they rated courses online. One of the online classes with an enrollment of forty students had responses from twenty-nine students (73 percent). Nineteen of these twenty-nine student respondents (66 percent) completed the summary comments, generating sixty-three lines of text. The other three classes evaluated online had similar results. The overall response rates of the online classes ranged from 64 through 69 percent. The number of students who completed the summary question ranged from 37 percent to 48 percent, resulting in proportionate increases in the amount of text written. On average, classes evaluated online had more than five times as much written commentary as the classes evaluated on paper, despite the slightly lower overall response rates for the classes evaluated online (Figure 3.2). The more extensive comments collected online provide a valuable resource for instructors who want to improve teaching and learning in their classes.

PERCEPTION: In comparison with comments written on paper rating forms, students write more negative comments when they submit responses online.

REALITY: Students submit positive, negative, and mixed comments online.

Students tend to be more expansive in their comments when they enter them online at their convenience. However, there is a division of thought regarding the nature of these comments. Some faculty members believed that online rating systems encourage students to be less hurried and more contemplative when entering their comments; others felt that students' isolation and the anonymity of the online environment would promote negative student comments.

The subjective nature of student comments makes them more difficult to measure than numerical ratings. To quantify comments, CTEC researchers classified each comment as positive, negative, or mixed. The comments were classified using each student's response in its entirety to the summary question. Student comments were classified as positive or negative when everything written tended in that direction. Responses classified as mixed contained both positive and negative elements. For example, a student may have commented favorably about most elements of the class but indicated one area that needed improvement. Compared with courses that were evaluated using paper forms, comments written about courses that were evaluated online showed no tendency toward more negative or mean-spirited responses.

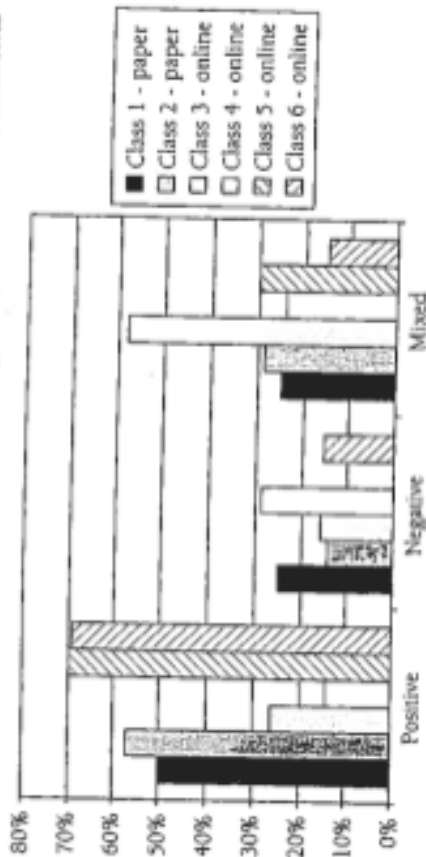
In comparing the six courses evaluated online and on paper, the number of positive, negative, and mixed comments was similar. In fact, two of the online classes had a higher percentage of positive comments than the classes evaluated on paper, which contradicted the preconception about online ratings. One of these classes using online evaluations was taught by a first-year instructor. The students seemed to be sensitive to the teacher's nature and to offer suggestions for improvements in the future. Although one of the classes evaluated online had a higher percentage of negative than positive comments, this would be expected within any sampling of classes (see Figure 3.3 for a distribution of the subjective quality of summary essay comments).

Conclusion

After decades of paper-pencil evaluations, Northwestern University implemented a system for online student ratings of instruction in 1999. Since then, students have evaluated instruction online and have obtained a certain portion of the compiled reports through the Web. More portions of the evaluation report are made available online to faculty and administrators.

Although some skeptics were resistant to the change to online course ratings, many (mis)perceptions have been dispelled by the realities of the Northwestern online ratings experience; for example, average numerical scores for the online ratings have been shown to be about the same as for the paper ratings. Although response rates have been somewhat lower using the online system, those students who do respond write more detailed

Figure 3.3. Distribution of Positive, Negative, and Mixed Comments



comments online. These additional comments are a valuable resource to instructors who want to improve teaching and learning in their classes. Overall, the reality of the online ratings experience at Northwestern has motivated administrators, faculty, staff, and students to continue online data collection and online reporting for student ratings of instruction.

Questions for Further Investigation

Online student ratings of instruction are in their infancy and invite a great deal of study and improvement. As CTEC continues to collect evaluations, predominantly through the electronic process, important questions will need to be answered.

So far at Northwestern, response rates for the online evaluations have been somewhat lower than those for paper evaluations. Is it possible to collect evaluations electronically with close to the same response rate as those collected using paper rating forms? Under what conditions might this be achieved? Might this goal be achieved if every student sat in class with a laptop computer? How high must response rates be to establish validity?

Online ratings yield more student comments, and these comments appear to be more useful to faculty and students. Does this increase in student comments compensate for a somewhat lower response rate from the online ratings? In other words, which is more valid and useful, evaluations from a greater number of students (some of whom will write almost anything just to have it finished and leave the classroom) or collecting evaluations from fewer students who have definite ideas concerning the class and who provide more written feedback?

Do online evaluations generally yield lower scores than paper evaluations? Research conducted so far at Northwestern is inconclusive on this

point. If further research shows that online ratings are lower, what are the implications? What would need to happen for the university community to accept an adjusted standard of ratings results?

These and other questions need to be answered as Northwestern and other universities explore and move forward with online student ratings of instruction.