Online Student Ratings of Teaching Effectiveness: Analysis of Data from Select Semesters (2009-2010)

Prepared for the Committee on Faculty Affairs of the University Faculty Senate
The Pennsylvania State University
by
Angela R. Linse, Schreyer Institute for Teaching Excellence

Fall 2010

Introduction

This report provides a detailed comparative analysis of data from paper and online administration of Penn State's university student ratings instrument, the Student Rating of Teaching Effectiveness (SRTE). The report includes results of analyses from early in Penn State's three-year implementation plan transitioning from use of paper ratings forms to online administration. The data are drawn from seven academic units and four campuses at Penn State. The first three campuses to participate fully in the online project are Penn State DuBois, Penn State Erie-The Behrend College, and Penn State Great Valley. At the University Park campus, the College of Health and Human Development was joined by the College of Agricultural Sciences and the College of Earth and Mineral Sciences. The School of Nursing includes all courses offered at the University Park campus and the Milton S. Hershey Medical Center.

The Committee on Faculty Affairs of the University Faculty Senate requested analyses of data aggregated into University Park and Commonwealth Campuses. Each analysis compared Online SRTE results to the most recent results from in-class administration of SRTEs using paper forms. Fall and spring semesters were compared separately to reduce the possibility that differences could be attributed to curricular differences between the two semesters.

The first part of this report provides the context for the statistical analysis by beginning with an overview of the SRTEs at Penn State. Over the past 25 years, many faculty and administrators have joined Penn State and may not be familiar with the development and implementation history. This section includes the legislative and administrative process establishing the SRTEs, the format of the SRTEs, and subsequent revisions to the guidelines for administering the SRTEs. The next section provides the background and chronology of the Online SRTE Pilot Project and the development of the Online SRTE Implementation Plan.

The report then details the concerns that faculty members have expressed about the transition to an online delivery system. These concerns were collected over the course of the three-year pilot and during the early stages of expansion and implementation of the Online SRTE project. The concerns with the highest priority include changes in average scores and changes in the response rates. The remainder of the report describes the data sample, statistical analyses, and results.

History of SRTEs at Penn State

Legislation and Administration

University student ratings were proposed in an Advisory and Consultative Report from the Senate Committee on Faculty Affairs. The report was discussed, minor changes proposed, and approved by the full Senate on April 30, 1985. A corrected copy of the legislation, including corrections and changes approved at the April 30, 1985 Senate meeting, was published in the February 25, 1986 Senate Agenda (see http://senate.psu.edu/agenda/srte/srte_record22586.pdf).

Recommendation 2 of the report calls for a Statement of Practices for the Evaluation of Teaching for Promotion and Tenure to be adopted as a supplement to promotion and tenure policies and procedures.

The original wording of the Senate legislation states "The Office of the Executive Vice President and Provost of the University shall be responsible for coordinating and administering the student evaluation system in consultation with individual units" (item A.4). By September 1986, the official name of the system had been decided, Student Ratings of Teaching Effectiveness. The Statement of Practices document (cited above) had been updated with this name and the units designated to administer and process the SRTEs. The Office of the Vice Provost for Academic Affairs is the administrative unit and the Schreyer Institute for Teaching Excellence is the processing unit; the Institute absorbed the original processing unit in 2002.

On February 11, 1986, Vice Provost Carol Cartwright requested that Academic Deans, Campus Executive Officers, and Directors of Academic Affairs communicate to Department and Division Heads that they should offer all faculty members the opportunity to review and comment on the proposed fixed pool of questions for the SRTEs. [1]

By May 9, 1986, Vice Provost Cartwright announced that the fixed pool of questions had been finalized "through widespread consultation with faculty members."[2] Dr. Cartwright's memo clarified that the final stage of the process is for the faculty in the academic departments to select questions from the fixed pool. Dr. Cartwright's memo specified that Commonwealth Campus faculty members were to be involved in development of the departmental forms. Each unit was allowed from one to three different versions of the SRTEs "to suit different types or levels of instruction in their discipline."[2] The departments were to have made their final selections by October 15, 1986, that deadline was later extended to February 6, 1987.[3]

The SRTEs were first administered in spring 1987.[4]

---

[1] Memo on file at the Schreyer Institute for Teaching Excellence dated 2-11-1986 from Carol A. Cartwright, Dean for Undergraduate Programs and Vice Provost, to the Council of Academic Deans, for Transmittal to Department and Division Heads.

[2] Memo on file at the Schreyer Institute for Teaching Excellence dated 5-9-1986 from Carol A. Cartwright, Dean for Undergraduate Programs and Vice Provost, to the Council of Academic Deans for Transmittal to Department Heads (including heads of divisions and programs).

[3] Memo on file at the Schreyer Institute for Teaching Excellence dated 11-3-1986 from Carol A. Cartwright, Dean for Undergraduate Programs and Vice Provost, to the Council of Academic Deans for transmittal to Department Heads, and to Campus Executive Officers for transmittal to Directors of Academic Affairs.

[4] Memo on file at the Schreyer Institute for Teaching Excellence dated 7-27-1987 from Carol A. Cartwright, Dean for Undergraduate Programs and Vice Provost, to the Council of Academic Deans and Campus Executive Officers.

Over the next few years, a number of faculty raised issues about the normative data included on SRTE instructor reports and how these normative data were being interpreted by administrators and review committees. On February 21, 1989, the senate voted to remove all references to "norms" and "normative data" from The Statement of Practices for the Evaluation of Teaching Effectiveness and to eliminate reporting of the median and standard deviation (http://senate.psu.edu/agenda/srte/srte_2-21-89.pdf, p. 2).

The last substantial change in the Statement of Practices was made on September 16, 2003 (http://www.psu.edu/ufs/agenda/sep16-03agn/sep16-03agn.html#AppendixB) and addressed a variety of issues including clarification of the frequency of reviews, which results should be included in dossiers, and a number of procedural issues surrounding the selection of Section B questions. One recommendation about written student comments, which units collect simultaneously with SRTEs, was enacted through changes in the Promotion and Tenure Frequently Asked Questions, Student Comments (see Question 28 in http://www.psu.edu/dept/vprov/pdfs/p_and_t_faq.pdf).

The current copy of the Statement of Practices is maintained by the Vice Provost for Academic Affairs (VPAA) and made available at http://www.psu.edu/dept/vprov/pdfs/srte_statement.pdf. The current Statement of Practices is also included as Appendix A of the Guidelines for Promotion and Tenure (published annually by the VPAA) and available at http://www.psu.edu/dept/vprov/pdfs/p_and_t_%20guidelines.pdf.

SRTE Format

The basic structure of the SRTEs has not changed since the original legislation was passed by the Senate in 1985. That legislation specifies that the student ratings will include three sections.

Section A: University (4 required questions, including two global questions rating the overall quality of the course and the overall quality of the instructor)
Section B: Academic unit (5-15 questions selected by the faculty of the academic unit from the pool of 177 questions; http://www.srte.psu.edu/SRTE_Items)
Section C: Instructor (up to 5 additional questions selected by individual faculty members from list of 177 questions)

Faculty in the academic units are, and will continue to be, responsible for selecting which questions are to be used in Section B. Historically, Section B has been referred to as the "department section," and many SRTE documents refer to departments. However, course abbreviations have always been used as a basis for creation of the Section B questions.[5] Most departments have a single course abbreviation, but in academic units that offer courses under more than one course abbreviation, the Section B questions for each may be different.

In 1997, concurrent with a reorganization of the Commonwealth Campuses, each campus selected Section B questions to correspond with the courses offered at their campus, with course abbreviations still serving as a proxy for a department. The restriction to three forms per department (course abbreviation) has been relaxed over the years, but most course abbreviations are still associated with only 2-3 forms. Historically, academic units have rarely made changes to the Section B questions. The

[5] Memo on file at the Schreyer Institute for Teaching Excellence dated 7-10-1987, from Ed Rosenstock, University Testing Services (UTS) to SRTE Area Representatives.

current Section B question selections for all locations are available at http://www.srte.psu.edu/AvailableForms/Reports_Forms_Available.aspx.

The September 16, 2003 revision of the Statement of Practices confirmed that the current campus academic units are responsible for Section B question selection (Recommendation 5, http://www.psu.edu/ufs/agenda/sep16-03agn/sep16-03agn.html#AppendixB) and that all units may review and change Section B questions, and have different Section B questions for different courses (Recommendation 6).

Frequency of SRTE Reviews

The frequency with which faculty and courses are evaluated using the SRTEs has been, and continues to be, determined by the college in consultation with the faculty, with some exceptions (Statement of Practices for the Evaluation of Teaching for Promotion and Tenure, Section I.A.11, http://www.psu.edu/dept/vprov/pdfs/srte_statement.pdf). Provisional (pre-tenure) faculty are expected to have SRTEs administered for all sections of all courses. Under special circumstances an administrator may grant an exemption (see Section I.A.11.a.2 in the Statement of Practices referenced above). This frequency recommendation has been in place since September 16, 2003 (Recommendation 1).

Background and Chronology of the Online SRTE Project

From 2002 to 2004, two University Park departments—Biology and Integrative Arts—participated in testing whether SRTEs could be administered online. Using commercially available testing software (Test Pilot), SRTEs were successfully administered in courses in these two departments. In October 2004, the Vice Provost for Academic Affairs commissioned an Online SRTE Committee to investigate the feasibility of moving the administration of all SRTEs from paper to an online environment. The committee included faculty, students, administrators, and technical and measurement professionals (see Appendix 1). Additional members were added to two subcommittees that explored administering and reporting issues. The committee investigated a variety of issues, including commercial software, the prevalence of online administration at other universities, and implications for Penn State faculty and students.

By April 2005, the Online SRTE Committee created specifications for an Online SRTE tool.[6] In June 2005, Penn State contracted with ANGEL Learning Corporation to develop a prototype tool using the SRTE Committee's specifications. The intent was to test the prototype in a small sample of Penn State course-sections and, if successful, to continue to develop the tool. The ANGEL Learning prototype, which included many, but not all of the specifications requested, was delivered and in use by fall 2005. The prototype continued to be used by pilot project participants between fall 2005 and spring 2008 (the tool was never used in summer semester). During this 3-year pilot, a single person managed all setup and administration for all course-sections. While numerous faculty and academic units requested participation, the pilot could not be expanded beyond the cumulative total of 190 course-sections (http://www.srte.psu.edu/OnlineReports/). Not until more of the functions originally requested by the Online SRTE Committee were added and a number of technical fixes implemented could the project be expanded.

[6] Online SRTE Committee (April 2005) Online SRTEs: Report 1, report on file at the Schreyer Institute for Teaching Excellence.

In early 2007, the University Faculty Senate requested an update on the Online SRTE Pilot project. An Informational Report was provided at the March 20, 2007, senate meeting (http://senate.psu.edu/agenda/2006-2007/mar20-07agn/apph.pdf). Of particular concern to faculty senators was the potential for decreased response rates and an increased frequency of negative responses. While the pilot projects produced response rates consistently at or above 60%, the sample size was small and faculty participated on a voluntary basis and the available data were insufficient to address faculty concerns.

The Senate update on the Online SRTE project generated sufficient interest in obtaining a larger sample of Online SRTE results that it prompted a concerted effort by the Schreyer Institute for Teaching Excellence to collaborate with the Administrative Information Systems (AIS) unit of Information Technology Services on an improved tool. From spring 2007 through spring 2008, the Schreyer Institute and AIS programmers worked to develop a Memorandum of Understanding (MOU) that identified the programming and technical problems to be resolved before the pilot project could be expanded (MOU signed August 2008). The Vice Provost for Academic Affairs provided funding for the AIS programmers and new secure servers for the SRTE system and data storage, the Schreyer Institute continues to provide staff to administer and manage the Online SRTE program. Schreyer Institute staff and AIS ANGEL programmers have been meeting every two weeks since fall 2008 and the team has made significant progress on the most pressing issues.

By December 2008, the Online SRTE tool had been moved off of the ANGEL servers to a separate secure SRTE server. This transition was seamless for students who continue to access the Online SRTE tool through their ANGEL accounts. The early contract with ANGEL Learning, the look of the tool, and the involvement of AIS ANGEL programmers, leads some individuals to mistakenly refer to the Online SRTEs as an ANGEL tool. This common mistake causes both faculty and students to express concerns about confidentiality. The systems are separate and the links within students' ANGEL accounts take them to a separate server to complete the SRTEs. The only information accessible to faculty is through a permanent frame in ANGEL that shows which of the faculty member's courses are being evaluated online and a feed indicating the current response rate.

By early December 2008, the most urgent technical challenges had been addressed, which allowed release of an upgraded version to the SRTE Representatives in all of the academic units that had been participating in the 3-year pilot project. Some of University Park units chose to administer the Online SRTEs in all of their courses, others expanded to a more limited degree, and still others remained at pilot project levels (Table 1). The fall 2008 expansion served as a load-test of the new servers and the upgraded program. This first upgrade and expansion was extremely successful.

Table 1. Overview of participation in the Online SRTE project.

Academic Unit	Unit Code	2005 Fall	2006 Spring	2006 Fall	2007 Spring	2007 Fall	2008 Spring	2008 Fall	2009 Spring	2009 Fall	2010 Spring
Penn State DuBois	DS								**all**	**all**	**all**
Penn State Erie, The Behrend College	BD										**all**
Penn State Great Valley	KP			pilot		pilot					**all**
Penn State York	YK	pilot	pilot	pilot							
College of Agricultural Sciences	AG										**all**
College of Arts and Architecture	AA					pilot	pilot	pilot	1 dept	1 dept	1 dept
College of Communications	CM		pilot	pilot	pilot	pilot	pilot	many	most	most	most
College of Earth and Mineral Sciences	EM		pilot	pilot	pilot	pilot	pilot	pilot	pilot	pilot	**all**
College of Engineering	EN	pilot	pilot	pilot	pilot	pilot	pilot	pilot	pilot	pilot	pilot
College of Health and Human Development	HH		pilot			pilot	pilot		**all**	**all**	**all**
School of Nursing	NR										**all**
Eberly College of Science	SC	pilot	pilot	pilot	pilot	pilot	pilot	1 dept	2 depts	2 depts	all
World Campus	WD			pilot	pilot	pilot		pilot		pilot	pilot

Note: Blank cells indicate use of paper SRTEs; "pilot" indicates a small number of course-sections administered the SRTEs online; "dept" indicates that all course-sections in one or more departments administered the SRTEs online; "many" indicates participation of a large number of course-sections; "most" indicates participation of nearly all course-sections; "all" indicates full participation by the college or campus.

In spring 2009, some important changes to the Online SRTE system were introduced. The Online SRTE tool was upgraded to allow two critical functions: a) the SRTEs could be administered separately for each instructor in a multi-instructor course and b) the SRTEs could be administered at different times for each instructor or for non-traditional semesters. Additionally, this semester marked the first formal expansion of the Online SRTE tool to include all of the course-sections from one campus, Penn State DuBois, and one college, the College of Health and Human Development.

Implementation Plan

With the successful technological implementation of the Online SRTEs by the College of Health and Human Development and the DuBois campus, a draft implementation plan was developed by the Schreyer Institute for Teaching Excellence (Table 2). A number of factors were considered in developing the plan, including system capacity/load, improvements in the Online SRTE software program, the complexity of courses offered by academic units, availability of training materials, and availability of centralized staff support.

The Implementation Plan spans six semesters, from fall 2008 through spring 2011, excluding all summer semesters. Separating the Online SRTEs from ANGEL servers in fall 2008 to dedicated and secure SRTE servers permitted the initial increase in capacity and the doubling of expected submissions in each subsequent semester (excluding summer). Successful administration of SRTEs for all course-sections in the College of Health and Human Development and at Penn State DuBois in spring 2009 and fall 2009 reinforced that the system could easily handle a doubling of the load each semester.

The Provost's designee for administration of the SRTEs, Vice Provost Blannie Bowen, distributed the draft plan at the October 8, 2009 meeting of the Academic Leadership Council (http://www.srte.psu.edu/pdf/SRTE_ImplementationPlan_10-8-09.pdf). The plan indicated that Online SRTEs were fully implemented for all World Campus courses when only a small sample of course-sections were evaluated using the Online SRTE tool. On November 30, 2009, the Vice President and Dean for Undergraduate Education and the Vice President for Outreach requested that the SRTEs be administered for all World Campus courses. The implementation plan was shared with the University

Senate Committee on Faculty Affairs at its December 8, 2009 meeting and the plan and preliminary data were discussed at the January 26, 2010 meeting.

The implementation plan was also shared with the Administrative Council for Undergraduate Education (ACUE) on January 7, 2010. At the request of the ACUE student representative, a similar presentation was made to the Academic Affairs Committee of the University Park Undergraduate Association (UPUA) on January 21, 2010. Over the past year, the Schreyer Institute has visited numerous campus, college, division, and department meetings to discuss the Online SRTEs with faculty and administrators.

Table 2. Online SRTE Implementation Plan.

2008-2009

Fall 2008
Departments
- Biology
- Engineering Design, Tech, & Professional Programs
- Engineering Science & Mechanics
- Integrative Arts
- Physics
- Individual faculty

Colleges
- Communications

Campuses
- World Campus (29 sections)

Spring 2009
New Colleges
- Health & Human Development

New Campuses
- DuBois

Summer 2009
No new units

2009-2010

Fall 2009
No new units

Spring 2010
New Colleges
- Agricultural Sciences
- Earth & Mineral Sciences
- Nursing
- Science (remaining)

New Campuses
- Erie–Behrend
- Great Valley

Summer 2010
No new units

2010-2011

Fall 2010
New Colleges
- Arts & Architecture
- Education
- Information Sciences & Technology
- Law

New Campuses
- Abington
- Altoona
- Beaver
- Berks
- Brandywine
- Fayette
- Greater Allegheny
- Harrisburg
- Hazleton
- Lehigh Valley
- Mont Alto
- New Kensington
- Schuylkill
- Shenango
- Wilkes-Barre
- Worthington Scranton
- York

Interdisciplinary Units
- Continuing & Distance Education
- Military Studies (ROTC)

Spring 2011
New Colleges
- Business
- Engineering
- Liberal Arts

Training and Expansion

In spring 2010, online training materials became available, which eased the expansion to four additional colleges and two additional campuses. Training materials are available in a web-based and portable document format from the SRTE website (http://srte.psu.edu).

The spring 2010 expansion exposed a number of complexities in the way that courses and instructors are listed in the course schedule section of ISIS (the centralized student records system) from which course, section, instructor, and student information are drawn. Additional challenges included resolving how to set up cross-listed courses and courses developed at one campus location, but offered at another. All of these are new issues because the paper SRTE process did not rely on ISIS course

schedule data. Keeping the number of new units relatively small has allowed Schreyer Institute staff to provide personalized attention and problem solving support to support staff in the academic units and to make process improvement recommendations to the Vice Provost for Academic Affairs.

In fall 2010, four additional colleges and the remaining Commonwealth Campuses are scheduled to begin administering SRTEs online. While this appears to be a significant increase, as with previous expansions, it simply doubles the expected submission load as planned. In spring 2011, the final three colleges at University Park (Business, Engineering, and Liberal Arts) will be added.

Changes Associated with Online Administration of SRTEs

Three changes are associated with the move to online administration of SRTEs: the delivery mechanism, the offering period, and additional questions for written comments.

Delivery Mechanism. The SRTEs are administered online through a secure system separate from, but accessible to students through the ANGEL Course Management System. The SRTE link in ANGEL takes the student to a separate location (https://cms.psu.edu/Portal/Nuggets/MySRTE/jump.asp). Faculty do not have access to the Online SRTE system, but are provided with information about their course section response rates through ANGEL during the offering period.

Offering Period. As with paper SRTEs, the Offering Period is the last two weeks of regular classes during the semester for traditional full-semester courses. SRTEs are not offered during finals week. The two-week period is specified in the Statement of Practices for the Evaluation of Teaching for Promotion and Tenure, Section I.A.10.a.7 (see link above). The difference between paper and online administration is in the way that this two-week offering period is used. For paper SRTEs, academic units typically distribute SRTE forms to those students present in class during one day of the offering period. For Online SRTEs, students have the entire offering period to complete their ratings because they must do so outside of class, on their own time, during a busy time of the semester. For courses that are shorter than a traditional semester, students are allocated one business day for each week of class, up to a maximum of ten business days. For example, a six-week course will have a six-day offering period.

Students are notified by email when the Online SRTE offering period opens if they are enrolled in a participating course-section. Students receive a maximum of two email reminders if they have not completed their SRTEs; no reminders are sent to students who have completed their SRTEs.

Faculty members are not notified when the SRTEs are available for their students. However, if SRTEs have been setup for a faculty member, the specific course-sections are listed in the permanent SRTE frame visible on the faculty member's My Profile page in ANGEL (the default homepage). This frame allows faculty to monitor the response rates for their course-section(s). In focus group interviews, students responded positively to faculty discussions about the value of student feedback and how the feedback has led to course improvements.

Only students currently enrolled in a course may evaluate the course. Student enrollment information is updated daily using registration data drawn from the centralized student records system (ISIS). If a student withdraws from the course during the offering period, Online SRTE access for that course is revoked.

Additional Questions. The Online SRTEs have two kinds of additional questions. First, two open-ended questions were added to the system in fall 2008 to allow collection of written student comments. The two open-ended questions are:

Open 1: What helped you learn in this course?
Open 2: What changes would improve your learning?

The phrasing of these questions is a result of years of refinement in colleges and universities across the U.S. They are deliberately focused on students' learning and designed to elicit usable information rather than focused on student likes and dislikes (which are difficult to use as the basis for instructional improvements).

In fall 2009, programming was completed for an "Additional Questions" section that allowed academic units to administer a survey of up to 15 questions concomitant with the SRTEs. This section is not officially part of the SRTEs, but recognizes that many academic units gather written student feedback through a survey administered at the same time as the SRTEs. These surveys are one of the most common methods to gather student feedback additional to that obtained from the SRTEs, which is required for promotion dossiers (see Section C.1.a, Special Guidelines for the Criterion of The Scholarship of Teaching and Learning, in the Administrative Guidelines for HR-23: Promotion and Tenure Procedures and Regulations, http://www.psu.edu/dept/vprov/pdfs/p_and_t_%20guidelines.pdf). The Additional Questions section of the Online SRTEs assures that such processes can continue in an online environment. Some academic units are also using this section to collect student perceptions about unit-specific strategic goals and/or learning outcomes assessment.

Benefits of Online Administration

Penn State faculty, students, and academic units accrue numerous benefits by moving the SRTEs online. Among the most important benefits for faculty are that student feedback is available much sooner, which allows student feedback to be used for improvement before the next semester begins. Faculty no longer need to cede class time for administration of the SRTEs. Students appear to be more likely to complete the written feedback surveys that many departments administer simultaneous with the SRTEs. Faculty report that the written feedback obtained online from students is more thorough and meaningful than the handwritten responses obtained in-class.

The administrative advantages are also significant. Academic units will no longer require significant amounts of support staff time at the end of the semester to order and prepare paper forms, to administer the SRTEs in all course-sections, nor to check forms for errors after students complete them. Many units also devote many staff hours to transcribing students' written feedback in order to preserve student anonymity.

Over one million paper SRTE forms typically require weeks to process. During this processing period, Institute staff must work with academic units to correct student coding errors, mismatches between enrollments and numbers of submitted forms, missing faculty information, and incorrect course numbers. Errors and processing time are significantly reduced because all course, faculty, and enrollment data are drawn directly from the student records data system (ISIS).

Faculty Concerns

As the Online SRTE project has become more widely known, a number of faculty have expressed concerns about the transition to an online delivery mechanism. The University Faculty Senate Committee on Faculty Affairs requested that this report focus on two issues, average scores for the two overall questions (A3 and A4) and response rates. The results reported here represent only a fraction of the analyses that have been conducted.

The most important concern expressed by faculty is that average scores will decrease because the SRTEs are no longer restricted only to those students present in class on the day that the paper SRTEs are administered, but instead open to all enrolled students. Despite that score changes and unrepresentative samples could clearly have the greatest impact on faculty, the most commonly raised concern focuses on response rates. In part, this concern appears to be driven by anecdotes from colleagues at other institutions and stories in the educational press (not the research literature) that very low response rates will result in non-representative samples of students' perceptions.

Before proceeding to the analyses and results, these concerns are described in greater detail. The descriptions below are based on comments gathered from faculty and administrators across the university in meetings, individual conversations, and email exchanges.

Overall Course and Instructor Ratings

Some faculty are concerned that average scores might decrease because all enrolled students will have an opportunity to submit SRTEs, not just those in class on a select day. Underlying this concern is an assumption that students who do not regularly attend class are more likely to rate instructors or courses negatively. Accompanying the concern is an expectation that dissatisfied or disengaged students will not only be more likely to submit ratings online, but that those students will have more opportunities to evaluate the course and instructor than they do when the only opportunity to evaluate using paper SRTEs depends on their presence in class. If negative ratings increase, score distributions might be bi- or multi-modal.

Less commonly expressed are concerns about negative- or non-response-bias. Non-response bias results when the non-respondents differ systematically from respondents, which results in a non-representative sample. While decreased response rates for SRTEs administered online might result in non-response bias, the in-class administration of SRTEs is also subject to non-response bias because it excludes a particular kind of student from the survey—students not in attendance on the single day that paper SRTE forms are distributed. In fact, that Online SRTEs are available for the full two weeks might result in a sample of student responses that is more representative of students' perceptions rather than less representative.

The best method for testing for response bias is to compare results for a split-sample of students in the same course (e.g. 50% using the online system and 50% using the paper system). In the two years before the pilot project began, the split-sample methodology was used and no response bias was observed. Unfortunately, the Online SRTE tool was not designed for split-sample research—if one student in a course section has access to Online SRTEs, all students have access.

While research exploring non-response bias in paper and online delivery systems is not common, Thorpe (2002) conducted a limited study in three course sections at Drexel University. He found "that some

students are more likely to respond to an online course evaluation survey than others," including women and students expecting higher grades and that students expecting poor grades were less likely to respond to both web and paper based ratings surveys. Even with some minor differences, Thorp found no patterns that indicate web-based student evaluations would be consistently different from paper-based results. The largest and most comprehensive research on this topic to date is that of Webster, Benton, and Gross (2010) who presented an analysis of results from more than 700,000 courses using the IDEA Center's online student ratings system. The IDEA Center provides one of the most widely used student ratings systems in the US and its researchers are respected in the student ratings research community. In a detailed analysis of numerous items that ask students to rate teaching methods, instructor characteristics, course characteristics, and overall ratings of the course and instructor, the study found no meaningful differences between the delivery mechanism and the students' ratings.

Response Rates

Some faculty are concerned that response rates will decrease dramatically when SRTEs are administered online, possibly because many of the first online delivery systems, implemented in the mid-1990s, had very low response rates (see Sorenson and Johnson 2003). One common explanation of both small and large decreases response rates is that students do not want to make the effort to complete the SRTEs on their own time outside of class. Without the peer pressure of other students completing the SRTEs in class, faculty are concerned that it is easier for students to opt out of the SRTE process. Anderson and her colleagues at Washington State University suggest that the issue is less about the delivery mechanism than about student and faculty engagement in student evaluations (Anderson, et al. 2006).

The low response rates of early online systems may also be explained by students' lack of access to technology outside of class. In the intervening decade, not only has access to technology substantially increased, but online rating systems have become ubiquitous in students' lives through commercial, government, and non-profit websites. Over the six-semesters of the Online SRTE pilot project (Fa05-Sp08, excluding summer), the response rate ranged from 58-66%, which may indicate that technology access is not an issue for Penn State students.

While response rates for the pilot project were comparable to paper response rates (see http://www.psu.edu/president/pia/planning_research/reports/fall_02srte.html), response rates for the pilot project might also have been artificially high because the faculty participated voluntarily and might have created a high awareness among their students and inflated the response rates for Online SRTEs. Subsequent expansions of the Online SRTEs included all course-sections and all instructors within the academic units.

Even with the expanding participation, Penn State has some of the highest reported online response rates (see Appendix 2 and Brigham Young University's OnSET, http://onset.byu.edu). The relatively high response rates for Penn State have been achieved without any incentive programs or campaigns to encourage participation. Penn State has not implemented any processes or a publicity campaign to encourage student participation because only some colleges and campuses are participating. Once the entire university is participating, we plan to focus on this issue.

SRTE Comparative Analyses

Many factors other than the delivery method can influence SRTE response rates and scores, most importantly the instructor and the course. The Schreyer Institute has conducted a variety of detailed analyses to explore these issues in greater detail. At the request of the University Faculty Senate Committee on Faculty Affairs, this report includes information about average response rates and the average scores for the two global SRTE questions that inquire about the quality of the course and the instructor (A3 and A4). Also at the request of the Committee, the results below are reported separately for University Park and Commonwealth Campuses. Comparable data were extracted from the paper and online SRTE databases for the fully participating colleges and campuses.

Paper and Online Sampling Procedure

All datasets used in these analyses include only course-sections with five or more student responses. Paper SRTEs from course-sections with fewer than five responses are not processed, a restriction that was first enacted in 1987 and has been reaffirmed by every Vice Provost for Academic Affairs since. While the Online SRTE program does collect all responses, like paper SRTEs, summary results are not produced for courses with less than five student responses. The Online SRTE datasets used here exclude results from course-sections with fewer than five responses.

Analyses are conducted using SRTE datasets from matching semesters; i.e. fall semester data are not compared to spring semester data. In many academic units, the courses offered in fall semester are quite different from those offered in spring semester. Separating fall and spring semesters reduces the likelihood that differences in SRTE results reflect variations in course offerings by semester. Additionally, students' perceptions of the SRTE process may vary as they progress through the academic year. The Online SRTE data used in these analyses were collected in spring 2009, fall 2009, and spring 2010. Paper SRTE samples used in the analyses are drawn from the most recent corresponding semester.

The data used in these analyses include results only for colleges and campuses in which the Online SRTEs were begun at the same time (see shaded cells in Table 1). In spring 2009, the Online SRTE project was expanded to include all course-sections offered at one campus (DuBois) and in one college (Health and Human Development). In spring 2010, two new campuses (Erie-Behrend and Great Valley), two colleges (Agricultural Sciences and Earth and Mineral Sciences), and the School of Nursing began participating in Online SRTEs. Table 2 specifies the comparisons used in this report.

Table 3. Paper vs. Online SRTE Comparisons. Six paper samples and eight online samples are used in 14 comparisons. The paper sample from spring 2008 is compared to both spring 2009 and spring 2010.

Health and Human Development
- Spring 2008 paper vs. Spring 2009 online
- Spring 2008 paper vs. Spring 2010 online
- Fall 2008 paper vs. Fall 2009 online

DuBois
- Spring 2008 paper vs. Spring 2009 online
- Spring 2008 paper vs. Spring 2010 online
- Fall 2008 paper vs. Fall 2009 online

Agricultural Sciences, **Earth & Mineral Sciences**, and **Nursing** (includes nursing courses offered at University Park and M.S. Hershey Medical Center)
- Spring 2009 paper vs. Spring 2010 online

Erie-Behrend and **Great Valley**
- Spring 2009 paper vs. Spring 2010 online

Two other colleges (Science and Communications) now administer all SRTEs online. However, these colleges added courses and departments at different times, rather than all during the same semester. In colleges that implemented the Online SRTEs at different times, it is difficult to obtain comparative samples of paper SRTE results (see Table 1). For example, the biology and physics departments fully adopted the Online SRTEs before other departments in the Eberly College of Science. While we do have one semester in which all science departments participated (spring 2010), the most recent spring semester in which all science departments used paper SRTEs is spring 2004; results from comparisons that span a six year gap would be difficult to interpret with any degree of certainty. The College of Communications has periodically expanded its use of Online SRTEs over the course of the pilot project and allowed some faculty to opt out of using the system early on. This again results in a large time span between the most recent fully paper and fully online SRTEs. As such, University Park Science and Communications courses are not included in the analyses reported below.

Paper SRTE Results from 2004-2008

As a reference point for the subsequent comparisons of paper and online SRTE results, Table 4 presents aggregate summary statistics from five years of paper SRTEs for Penn State DuBois, Penn State Erie, the Behrend College, and Penn State Great Valley. Table 5 presents aggregate summary statistics from five years of paper SRTEs for the College of Agricultural Sciences, the College of Earth and Mineral Sciences, the College of Health and Human Development, and the School of Nursing.

Table 4. Paper SRTE Summary Statistics for Penn State DuBois, Penn State Erie, and Penn State Great Valley.

Semester	Response Rates			A3. Overall Course			A4. Overall Instructor		
	Sections	Mean	standard deviation	Responses	Mean	standard deviation	Responses	Mean	standard deviation
Spring 2004	621	84	13.17	12214	5.48	1.26	12265	5.71	1.35
Spring 2005	884	83	13.47	15792	5.55	1.23	15896	5.79	1.29
Spring 2006	895	84	13.42	15755	5.59	1.22	15834	5.83	1.31
Spring 2007	909	83	13.95	16751	5.58	1.25	16809	5.83	1.33
Spring 2008	884	84	12.08	17660	5.59	1.27	17711	5.86	1.32
Fall 2004	912	85	11.66	16936	5.52	1.24	17023	5.78	1.31
Fall 2005	939	84	12.24	16622	5.54	1.21	16741	5.80	1.27
Fall 2006	964	85	12.18	18586	5.55	1.25	18670	5.81	1.30
Fall 2007	997	85	11.94	20071	5.55	1.25	20112	5.84	1.30
Fall 2008	1039	86	11.46	21023	5.56	1.26	21084	5.83	1.32

Table 5. Paper SRTE Summary Statistics for the College of Agricultural Sciences, the College of Earth and Mineral Sciences, the College of Health and Human Development, and the School of Nursing.

Semester	Response Rates			A3. Overall Course			A4. Overall Instructor		
	Sections	Mean	standard deviation	Responses	Mean	standard deviation	Reponses	Mean	standard deviation
Spring 2004	830	81	16.84	19066	5.51	1.25	19148	5.81	1.28
Spring 2005	804	79	16.65	19666	5.42	1.28	19759	5.71	1.33
Spring 2006	827	79	16.24	20061	5.54	1.25	20157	5.82	1.29
Spring 2007	888	80	16.33	22140	5.60	1.21	22275	5.90	1.22
Spring 2008	859	80	16.63	20403	5.56	1.25	20484	5.83	1.31
Fall 2004	850	80	15.71	20284	5.48	1.23	20377	5.79	1.26
Fall 2005	908	79	16.77	21688	5.49	1.27	21785	5.79	1.32
Fall 2006	965	81	16.64	23824	5.57	1.24	23972	5.87	1.26
Fall 2007	984	80	17.01	23748	5.60	1.22	23867	5.90	1.25
Fall 2008	922	80	17.01	23337	5.52	1.28	23433	5.80	1.35

Results

The analyses use aggregate samples that encompass a range of variability in types of courses, faculty, and students in the participating units. As such, the results are unlikely to reflect the experiences of individual faculty or the results from a single course-section. Additional analyses have been conducted by course abbreviation and course-level for select colleges and campuses. Those data have been reported to those units, but are not reported here to ensure faculty privacy.

Interpretation of arithmetic means (averages) should be undertaken with care because means are sensitive to outliers. That is, the mean for a college or campus might be impacted by atypical results for particular course abbreviations. Likewise, the means for a particular course abbreviation might be impacted by uncommon results from a particular course or instructor.

Analysis of Scores for Overall Course (A3) and Overall Instructor (A4)

This section is divided into three parts. In the first are histograms (bar charts) of the samples listed above (Table 3), followed by summary statistics, and then statistical analyses. Results for overall ratings of the course (A3) are presented in Figure 1 and results for overall ratings of the instructor (A4) are shown in Figure 2.

The score distributions for paper and online delivery of SRTEs are quite similar, with the highest frequencies of ratings at the high end of the scale (scores 5-7). The online score distributions continue to have few instructors and courses rated at the low end of the scale (scores 1-3). This might alleviate concerns about a strong negative response bias because the distributions do not appear to support the prediction that students with negative views will be more likely to submit Online SRTEs.

Overall ratings of the course (A3) show relatively large increases at the high end of the scale (score 7), but only small increases at the low end of the scale (scores 1-3). For the participating Commonwealth Campuses, scores of 7 make up an average of 31.0% of the paper distribution and 36.8% of the online distribution. At participating University Park colleges, scores of 7 make up 24.8% of the paper distribution and 29.7% of the online distribution. For the campuses, scores at the low end of the scale (1-3 combined) encompass 6.0% of the paper distribution and 7.5% of the online distribution. At University Park, scores of 1-3 encompass 6.1% of the paper distribution and 7.7% of the Online SRTE distribution.

Overall ratings of the instructor (A4) show small increases at both the high and low ends of the scale. At the campuses, scores of 7 make up an average of 44.4% of the paper distribution and 46.2% of the online distribution. At University Park, scores of 7 make up 38.9% of the paper distribution and 39.8% of the online distribution. At the campuses, scores of 1-3 (combined) encompass 6.1% of the paper distribution and 8.6% of the online distribution. At University Park, comparable percentages are 6.6% of the paper distribution and 8.9% of the Online SRTE distribution.

Figure 1. Score distribution histograms for Overall Course (A3). University Park histograms are in the left column and Commonwealth Campus histograms are in the right column.

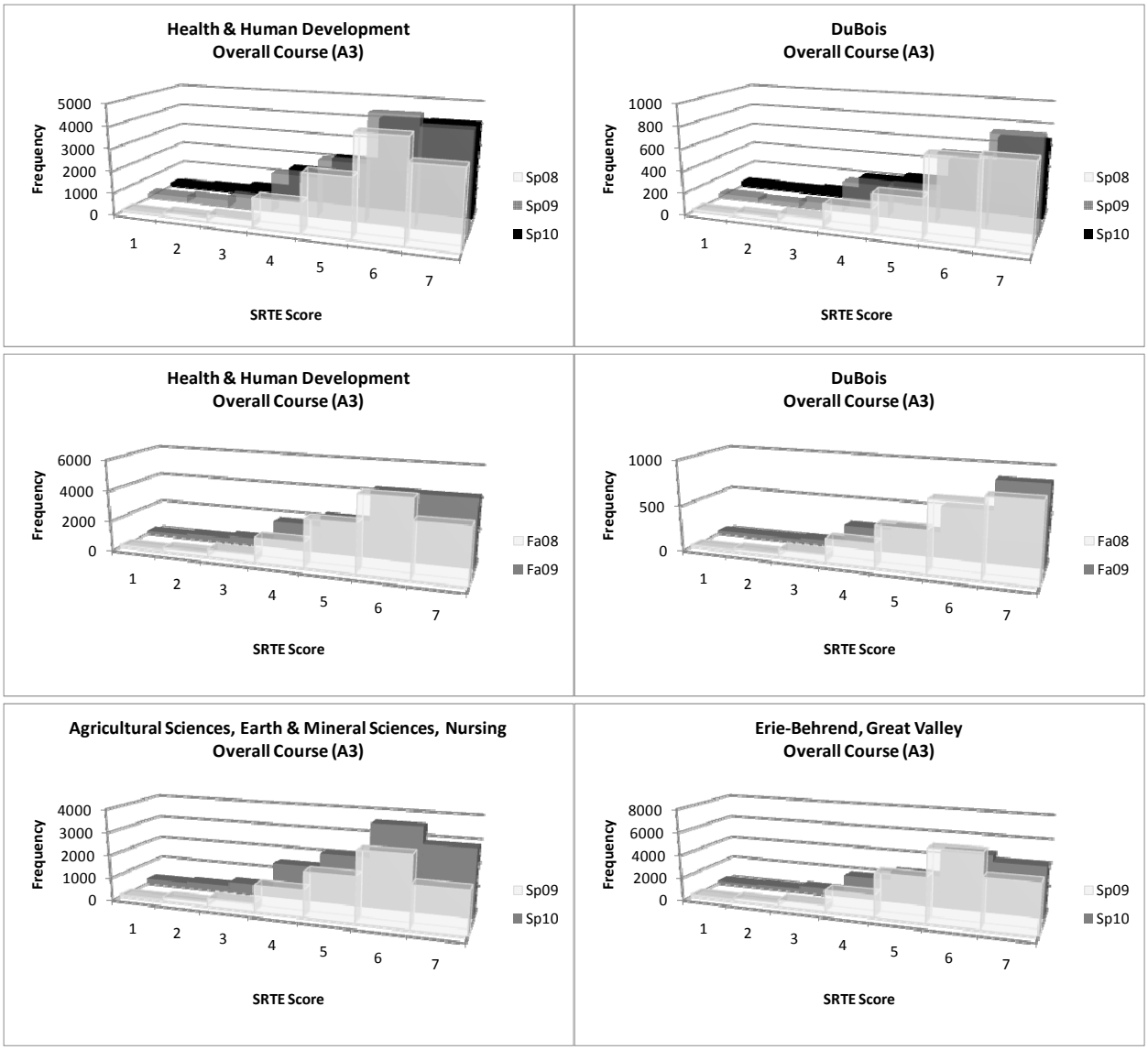
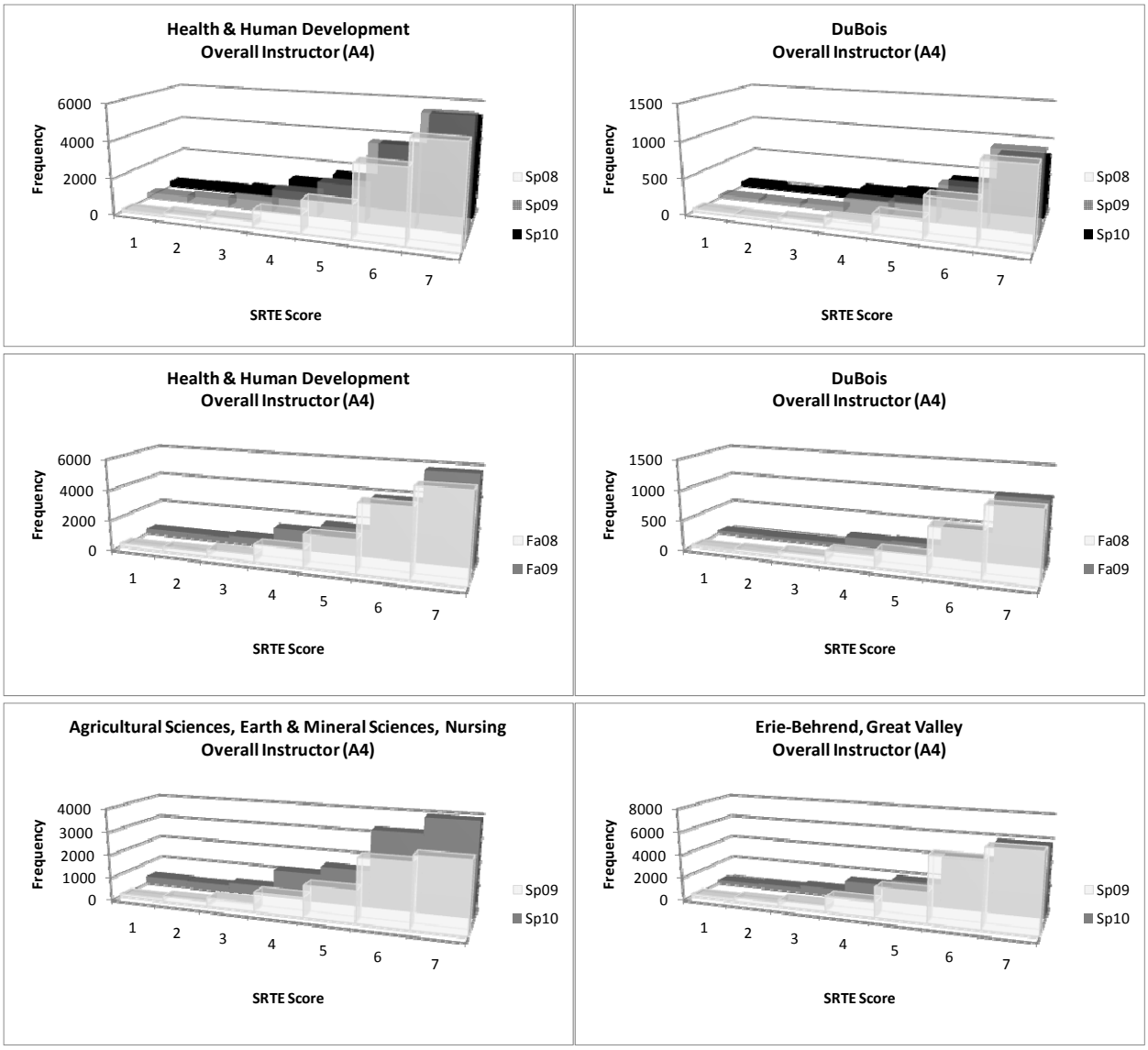


Figure 2. Score distribution histograms for Overall Instructor (A4). University Park histograms are in the left column and Commonwealth Campus histograms are in the right column.



Summary Statistics: Overall Course (A3) and Overall Instructor (A4)

Before presenting the statistical analyses of the average scores we present the summary statistics (Tables 6-11) for the comparisons listed above in Table 3. Commonwealth Campus results are followed by University Park results.

Table 6. Spring semester summary statistics for course-sections offered by Penn State DuBois.

Question	Method (sem/yr)	Submissions	Mean	standard deviation	95% Confidence Interval[7]	
A3. Overall Course	Paper (Sp08)	2124	5.75	1.30	5.70	5.81
A3. Overall Course	Online (Sp09)	2164	5.72	1.42	5.66	5.78
A3. Overall Course	Online (Sp10)	1778	5.66	1.51	5.59	5.73
A4. Overall Instructor	Paper (Sp08)	2133	6.00	1.35	5.95	6.06
A4. Overall Instructor	Online (Sp09)	2158	5.91	1.45	5.85	5.97
A4. Overall Instructor	Online (Sp10)	1770	5.80	1.55	5.73	5.87

Table 7. Fall semester summary statistics for course-sections offered by Penn State DuBois.

Question	Method (sem/yr)	Submissions	Mean	standard deviation	95% Confidence Interval[7]	
A3. Overall Course	Paper (Fa08)	2421	5.75	1.27	5.70	5.80
A3. Overall Course	Online (Fa09)	2100	5.81	1.39	5.75	5.87
A4. Overall Instructor	Paper (Fa08)	2440	5.98	1.29	5.93	6.03
A4. Overall Instructor	Online (Fa09)	2085	5.95	1.46	5.89	6.02

Table 8. Spring semester summary statistics for course-sections offered by Penn State Great Valley and Penn State Erie, the Behrend College.

Question	Method (sem/yr)	Submissions	Mean	standard deviation	95% Confidence Interval[7]	
A3. Overall Course	Paper (Sp09)	16806	5.59	1.24	5.57	5.61
A3. Overall Course	Online (Sp10)	15016	5.56	1.37	5.54	5.59
A4. Overall Instructor	Paper (Sp09)	16827	5.86	1.29	5.84	5.88
A4. Overall Instructor	Online (Sp10)	14972	5.73	1.47	5.71	5.76

Differences in the total number of observations for Health and Human Development course-sections between spring 2008 and spring 2009 reflect increased numbers of course-sections and increased enrollments in the college. While the college continued to expand its enrollments and course-sections, the number of SRTE submissions decreased. Interestingly, the mean scores for the course and instructor are higher in spring 2010, despite the drop in the average response rate.

Table 9. Spring semester summary statistics for course-sections offered by the College of Health and Human Development.

Question	Method (sem/yr)	Submissions	Mean	standard deviation	95% Confidence Interval[7]	
A3. Overall Course	Paper (Sp08)	12293	5.68	1.20	5.66	5.70
A3. Overall Course	Online (Sp09)	14716	5.59	1.37	5.57	5.62
A3. Overall Course	Online (Sp10)	12920	5.67	1.35	5.65	5.70

[7] The Confidence Interval is the range likely to contain the true response rate for the program in 95% of an infinitely large sample of course response rates. In other words, this means that 95% of the time, this interval would encompass the response rate if we had an infinite number of rate samples.

A4. Overall Instructor	Paper (Sp08)	12341	5.95	1.25	5.93	5.97
A4. Overall Instructor	Online (Sp09)	14691	5.77	1.46	5.75	5.79
A4. Overall Instructor	Online (Sp10)	12871	5.81	1.44	5.78	5.83

Note: The datasets used to generate the summary statistics in this table do not include results for Nursing courses. The School of Nursing was administratively included in the College of Health and Human Development in spring 2008, but was a separate administrative unit by fall 2008.

Table 10. Fall semester summary statistics for course-sections offered by Health and Human Development.

Question	Method (sem/yr)	Submissions	Mean	standard deviation	95% Confidence Interval[8]	
A3. Overall Course	Paper (Fa08)	13919	5.58	1.26	5.56	5.60
A3. Overall Course	Online (Fa09)	13978	5.62	1.37	5.60	5.64
A4. Overall Instructor	Paper (Fa08)	13975	5.84	1.34	5.82	5.86
A4. Overall Instructor	Online (Fa09)	13924	5.78	1.47	5.75	5.80

Note: The datasets used to generate the summary statistics in this table do not include results for Nursing courses. The School of Nursing was administratively included in the College of Health and Human Development in spring 2008, but was a separate administrative unit by fall 2008.

Table 11. Spring semester summary statistics for course-sections offered by the College of Agricultural Sciences, the College of Earth and Mineral Sciences, and the School of Nursing.

Question	Method (sem/yr)	Submissions	Mean	standard deviation	95% Confidence Interval[8]	
A3. Overall Course	Paper (Sp09)	8405	5.46	1.27	5.44	5.49
A3. Overall Course	Online (Sp10)	11165	5.47	1.40	5.45	5.50
A4. Overall Instructor	Paper (Sp09)	8423	5.70	1.38	5.67	5.73
A4. Overall Instructor	Online (Sp10)	11058	5.66	1.47	5.63	5.68

Note: The increased number of submissions between spring 2009 and spring 2010 does not reflect an increased response rate or increased enrollments. Most of the difference can be attributed to one Earth and Mineral Sciences faculty member who teaches over 1500 enrolled students. Since the faculty member has participated in the Online SRTE project from spring 2006 onward, his results are not included in the paper data from 2009, but are included in the online data from 2010.

Statistical Analyses: Overall Course (A3) and Overall Instructor (A4)

Statistical analyses of the rating distributions using Pearson's chi-square (χ^2) tests whether the frequencies in the categories of one variable are independent of the frequencies in the second variable. Here we use χ^2 to explore whether there is an association between the delivery mode and overall ratings of course quality and instructor quality. This non-parametric statistical test is appropriate for data that are not normally distributed and for a combination of ordinal and nominal data. Pearson's χ^2 is also sensitive to large sample sizes, which means that relatively small differences in cell values are likely to result in statistically significant χ^2 values.

Tables 12-17 include the results of the test of independence, i.e. whether the frequencies of SRTE scores along the 1-7 scale are independent of the delivery method. In every comparison, the null hypothesis of independence is rejected; each χ^2 value is statistically significant with p-values <.01. This result appears to indicate that there is an association between the delivery method and the frequencies of SRTE scores

[8] The Confidence Interval is the range likely to contain the true response rate for the program in 95% of an infinitely large sample of course response rates. In other words, this means that 95% of the time, this interval would encompass the response rate if we had an infinite number of rate samples.

in each category. However, statistically significant results are not surprising given the large numbers of SRTE submissions, which results in cell counts ranging from 20-6092. In other words, while the association may be statistically significant, it may have little practical significance.

The relationship between delivery method and SRTE scores is further explored using Cramér's V, which examines the strength of the association between a nominal variable (delivery method) and an ordinal variable (SRTE score) and is not sensitive to sample size effects. Cramér's V scores that are close to 1 indicate a strong association. None of the Cramér's V scores reported in Tables 12-17 are close to 1. Instead, they are all quite low, ranging from 0.05 to 0.11. Based on Cramér's V, none of the associations between delivery method and SRTE scores appear to be strong.

Table 12. Chi-square and Cramér's V analyses for spring semester course-sections offered by Penn State DuBois.

Comparison	Question	Pearson χ^2	Probability	Cramér's V
Sp08 paper: Sp09 online	A3. Overall Course	37.35	<.001	.093
Sp08 paper: Sp10 online	A3. Overall Course	49.26	<.001	.112
Sp08 paper: Sp09 online	A4. Overall Instructor	16.94	<.005	.063
Sp08 paper: Sp10 online	A4. Overall Instructor	28.17	<.001	.085

Table 13. Chi-square and Cramér's V analyses for fall semester course-sections offered by Penn State DuBois.

Comparison	Question	Pearson χ^2	Probability	Cramér's V
Fa08 paper: Fa09 online	A3. Overall Course	54.96	<.001	.110
Fa08 paper: Fa09 online	A4. Overall Instructor	46.13	<.001	.101

Table 14. Chi-square and Cramér's V analyses for spring semester course-sections offered by Penn State Great Valley and Penn State Erie, the Behrend College.

Comparison	Question	Pearson χ^2	Probability	Cramér's V
Sp09 paper: Sp10 online	A3. Overall Course	218.30	<.001	.083
Sp09 paper: Sp10 online	A4. Overall Instructor	194.37	<.001	.078

Table 15. Chi-square and Cramér's V analyses for spring semester course-sections offered by Health and Human Development.

Comparison	Question	Pearson χ^2	Probability	Cramér's V
Sp08 paper: Sp09 online	A3. Overall Course	172.416	<.001	.080
Sp08 paper: Sp10 online	A3. Overall Course	188.625	<.001	.086
Sp08 paper: Sp09 online	A4. Overall Instructor	172.579	<.001	.080
Sp08 paper: Sp10 online	A4. Overall Instructor	137.507	<.001	.074

Table 16. Chi-square and Cramér's V analyses for fall semester course-sections offered by Health and Human Development.

Comparison	Question	Pearson χ^2	Probability	Cramér's V
Fa08 paper: Fa09 online	A3. Overall Course	221.273	<.001	.089
Fa08 paper: Fa09 online	A4. Overall Instructor	110.496	<.001	.063

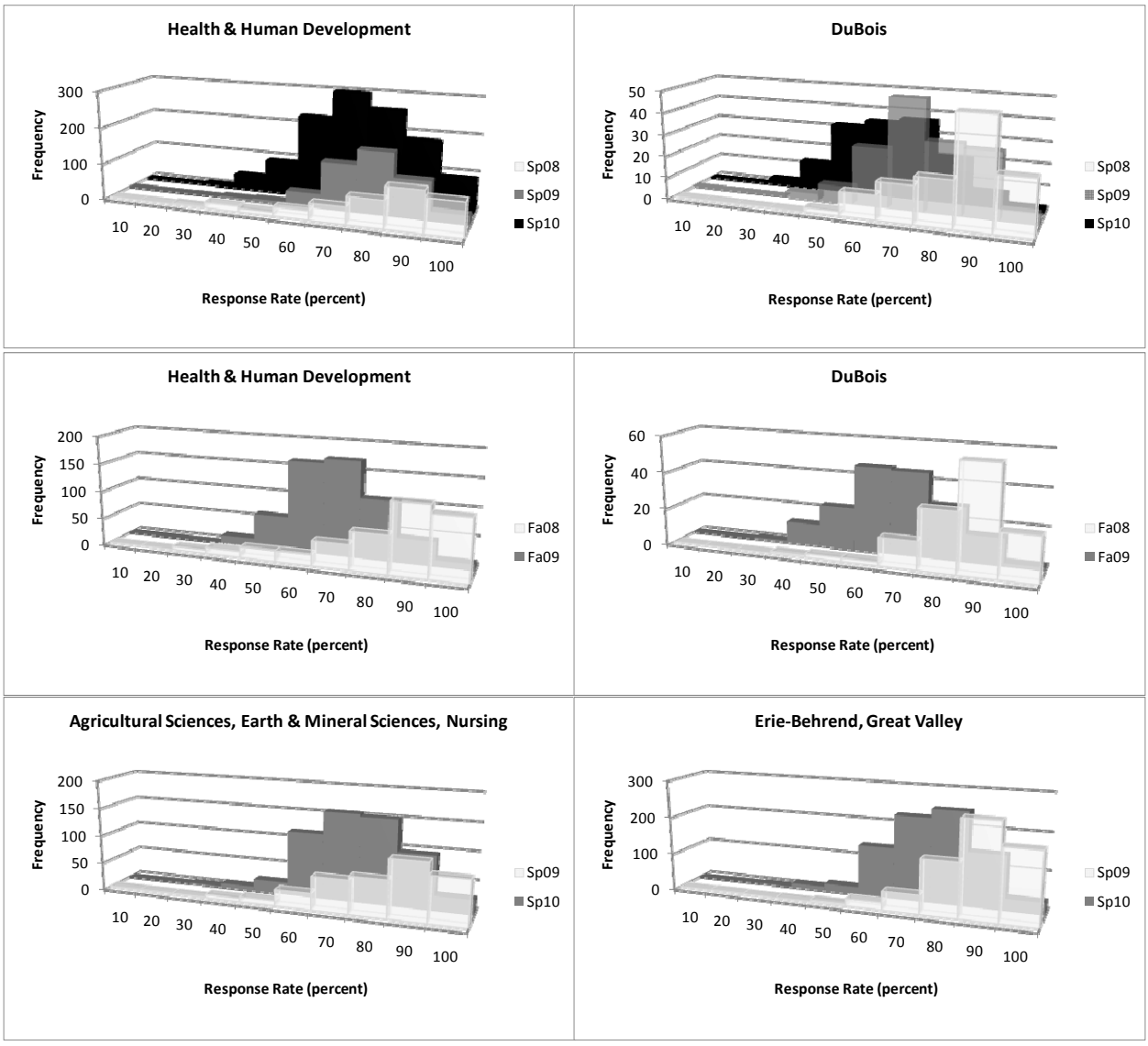
Table 17. Chi-square and Cramér's V analyses for spring semester course-sections offered by the College of Agricultural Sciences, the College of Earth and Mineral Sciences, and the School of Nursing.

Comparison	Question	Pearson χ^2	Probability	Cramér's V
Sp09 paper: Sp10 online	A3. Overall Course	110.816	<.001	.075
Sp09 paper: Sp10 online	A4. Overall Instructor	50.629	<.001	.051

Response Rate Analyses

Figure 3 presents the response rates graphically in histograms (bar charts) that reflect the comparisons listed in above in Table 3. As predicted the response rates for online administration of SRTEs are lower than for paper administration.

Figure 3. Response rate histograms.



Summary Statistics: Response Rate

Before presenting the statistical analyses of the average scores the summary statistics are presented in Tables 18-23 for the comparisons listed above in Table 3. Commonwealth Campus results are followed by University Park results. As expected, the mean response rates for online SRTEs are lower than for paper SRTEs with decreases ranging from 11%-23%. The paper response rates used in the analyses range from 79-86%, while the online samples have response rates ranging from 59-73%.

Table 18. Spring semester response rates for course-sections offered by Penn State DuBois.

Method (sem/yr)	**Mean**	**standard deviation**	**95% Confidence Interval**[9]	**# Sections**
Paper (Sp08)	82%	13.87	79% – 84%	146
Online (Sp09)	69%	14.15	67% – 71%	162
Online (Sp10)	59%	15.88	56% – 61%	155

Table 19. Fall semester response rates for course-sections offered by Penn State DuBois.

Method sem/yr)	**Mean**	**standard deviation**	**95% Confidence Interval**[9]	**# Sections**
Paper (Fa08)	84%	12.59	82% – 86%	169
Online (Fa09)	61%	14.24	58% – 63%	174

Table 20. Spring semester response rates for course-sections offered by Penn State Great Valley and Penn State Erie, the Behrend College.

Method (sem/yr)	**Mean**	**standard deviation**	**95% Confidence Interval**[9]	**# Sections**
Paper (Sp09)	86%	11.98	85% – 86%	814
Online (Sp10)	70%	13.12	70% – 71%	849

Table 21. Spring semester response rates for course-sections offered by the College of Health and Human Development.

Method (sem/yr)	**Mean**	**standard deviation**	**95% Confidence Interval**[9]	**# Sections**
Paper (Sp08)	79%	16.77	77% – 81%	414
Online (Sp09)	73%	12.14	72% – 74%	498
Online (Sp10)	59%	13.03	58% – 60%	529

Note: The datasets used to generate the summary statistics in this table do not include results for Nursing courses. The School of Nursing was administratively included in the College of Health and Human Development in spring 2008, but was a separate administrative unit by fall 2008.

Table 22. Fall semester response rates for course-sections offered by Health and Human Development.

Method (sem/yr)	**Mean**	**standard deviation**	**95% Confidence Interval**[9]	**# Sections**
Paper (Fa08)	80%	17.45	78% – 82%	445
Online (Fa09)	63%	12.54	62% – 64%	576

[9] The Confidence Interval is the range likely to contain the true response rate for the program in 95% of an infinitely large sample of course response rates. In other words, this means that 95% of the time, this interval would encompass the response rate if we had an infinite number of rate samples.

Table 23. Spring semester response rates for course-sections offered by the College of Agricultural Sciences, the College of Earth and Mineral Sciences, and the School of Nursing.

Method (sem/yr)	**Mean**	**standard deviation**	**95% Confidence Interval**[10]	**# Sections**
Paper (Sp09)	81%	16.41	79% – 82%	431
Online (Sp10)	70%	14.60	69% – 71%	615

Statistical Analyses: Response Rates

Statistical analyses of the response rates are presented in Tables 24-29 for each of the comparisons listed above (Table 3). The results for Commonwealth Campuses are followed by University Park results.

Statistical comparisons of the mean response rates for online and paper SRTEs, using t-tests for paired semesters, indicate that the differences are statistically significant. At the participating Commonwealth Campuses, average online response rates range from 59-70%, with decreases ranging from 13-23% lower than (average) paper SRTE response rates. In the participating colleges, online response rates range from 59-73%, with decreases ranging from 6% to 20%.

Table 24. Two-tailed t-test with equal variances analyses for spring semester course-sections offered by Penn State DuBois.

Comparison	**t-statistic**	**degrees of freedom**	**Probability**
Sp08 paper: Sp09 online	7.71	306	<.001
Sp08 paper: Sp10 online	13.31	299	<.001

Table 25. Two-tailed t-test with equal variances analyses for fall semester course-sections offered by Penn State DuBois.

Comparison	**t-statistic**	**degrees of freedom**	**Probability**
Fa08 paper: Fa09 online	16.39	341	<.001

Table 26. Two-tailed t-test with unequal variances analyses for spring semester course-sections offered by Penn State Great Valley and Penn State Erie, the Behrend College.

Comparison	**t-statistic**	**degrees of freedom**	**Probability**
Sp09 paper: Sp10 online	24.63	1657.15	<.001

Table 27. Two-tailed t-test with unequal variances analyses for spring semester course-sections offered by Health and Human Development.

Comparison	**t-statistic**	**degrees of freedom**	**Probability**
Sp08 paper: Sp09 online	5.85	735.18	<.001
Sp08 paper: Sp10 online	20.07	762.28	<.001

[10] The Confidence Interval is the range likely to contain the true response rate for the program in 95% of an infinitely large sample of course response rates. In other words, this means that 95% of the time, this interval would encompass the response rate if we had an infinite number of rate samples.

Table 28. Two-tailed t-test with unequal variances analyses for fall semester course-sections offered by Health and Human Development.

Comparison	t-statistic	degrees of freedom	Probability
Fa08 paper: Fa09 online	17.38	773.87	<.001

Table 29. Two-tailed t-test with unequal variances analyses for spring semester course-sections offered by the College of Agricultural Sciences, the College of Earth and Mineral Sciences, and the School of Nursing.

Comparison	t-statistic	degrees of freedom	Probability
Sp09 paper: Sp10 online	11.17	854.92	<.001

The previous analyses of average SRTE scores indicate that this statistical significance does not necessarily translate into practical significance. One of the most common statistical misconceptions is to equate statistical with practical significance by suggesting that significant results indicate "strong relationships between variables or big differences between comparison groups" (Huck 2009: 227). The similarity in SRTE scores, despite substantial response rate decreases, argues against inferring practical significance. That is, even though the online samples of student responses are smaller than the paper response samples, they do not result in compelling differences in average SRTE scores.

These response rates are much higher than those reported by faculty at other institutions and in the earliest literature reporting on experiments with online student ratings in the mid- to late-1990s (Sorenson and Johnson 2003) and anecdotally in the educational press. Interestingly, the response rates for the samples included in these analyses are comparable to those obtained during Penn State's six-semester pilot study, which ranged from 58-66% (see http://www.srte.psu.edu/OnlineReports/). Penn State's rates are also higher than the average response rate of 53% for ratings administered online by the IDEA Center (Webster et al. 2010).

A portion of the observed decrease in response rates is likely a result of requiring students to complete the SRTEs on their own time rather than during class. Yet, even with lower response rates, the sample of students completing the SRTEs may be more representative than the samples obtained through paper SRTEs because students have more than a single opportunity to complete the forms.

Both Penn State DuBois and the College of Health and Human Development have been using the online tool for three semesters and show a marked decrease in the third semester (spring 2010). These decreases may reflect a 'novelty effect', i.e. the response rates for the first semesters could be higher because students are intrigued by the idea of completing their ratings online. We will continue to monitor trends in response rates for participating units. As noted above, once all colleges and campuses have transitioned to online delivery, Penn State can consider implementing university-wide strategies for increasing response rates (see Appendix 3). One of the most common strategies, and reportedly one of the most successful for increasing response rates is to allow students early access to their grades. Since Penn State students' grades are already available relatively soon after the end of the semester, we may need to consider other strategies.

Conclusions

This report includes analyses of data from the first three semesters in which a number of campuses and colleges fully participated in online administration of the SRTEs. The Committee on Faculty Affairs of the University Senate requested that analyses be conducted on data aggregated into two groups, University Park and Commonwealth Campuses. Each analysis compared Online SRTE results to the most recent results from in-class administration of SRTEs using paper forms. Fall and Spring semesters were not aggregated in order to reduce the likelihood that differences were due to variations in curricula over the two semesters.

The statistical analyses addressed faculty members' primary concerns about decreased response rates and decreased average scores. The analyses were limited to results from two questions rating the overall quality of the course (A3) and the overall quality of the instructor (A4) because these results are used by all academic units in promotion and tenure dossiers and annual reviews.

At the aggregate level, all statistical tests indicate statistically significant differences between paper and online administration of the SRTEs. Summary statistics show small differences in the average scores for questions A3 and A4. Statistical comparisons using Pearson's chi-square test whether the frequencies of SRTE scores along the 1-7 scale are independent of the delivery method (paper vs. online). The test results indicate that there is an association between the distribution of scores and the method of administration. However, because Pearson's chi-square is sensitive to large sample sizes, the differences are predictably statistically significant. As such, we performed additional statistical analyses that test the strength of the relationship between scores and delivery method. The results indicate that none of the associations between delivery method and SRTE scores are strong.

The Schreyer Institute is currently conducting similar analyses on smaller samples from each college and campus by course level and by course abbreviation. Subsequent reports will be available to the Senate Committee on Faculty Affairs and on the SRTE website (http://srte.psu.edu); publically available information will only include aggregate information from which individual courses cannot be identified.

The response rates also indicate a statistically significant decrease in response rates. In the aggregate, the mean response rates for online SRTEs are 11%-23% lower than for paper SRTEs. While this is disappointing, it is expected given that students complete their SRTEs outside of class on their own time, with University Park showing a smaller range than the Commonwealth Campuses.

While the decrease is a concern and will be closely monitored, the outlook is positive. First, the analyses of scores for questions A3 and A4 indicate that despite substantial response rate decreases, SRTE scores do not show a comparable substantial decrease and are remarkably similar to those for paper. Second, once Online SRTEs are administered for all course-sections, Penn State will be able to consider implementing strategies for increasing response rates that have been successful at other universities. Since Penn State has some of the highest recorded response rates for online administration of student ratings, universities across the nation will be following our efforts.

In the meantime, the research literature for both paper and Online SRTEs continues to indicate that faculty are the most important determinant of students' participation. Faculty that let students know that their feedback is important and that the feedback is used to improve teaching and learning tend to

continue to receive higher response rates. When students feel that their feedback will not be taken seriously, response rates tend to be lower.

Again, the response rates and scores will continue to be monitored. However, at this time the advantages appear to outweigh the potential changes. The current strategy moving all course-sections to the Online SRTE system at the same time protects faculty under review. With a single transition date, review committees and administrators will be unlikely to attribute changes associated with this new delivery method to a change in quality of the instructor.

References Cited

Anderson, Joan, Brown, Gary, and Spaeth, Stephen (2006) Online Student Evaluations and Response Rates Reconsidered. *Innovate* 2(6), http://www.innovateonline.info/index.php?view=article&id=301.

Huck, Schuyler W. (2009) *Statistical Misconceptions*. New York: Routledge.

Sorenson, D. Lynn, and Johnson, Trav D. (2003) *Online Student ratings of Instruction*. New Directions for Teaching and Learning, No. 96. San Francisco: Jossey-Bass.

Thorpe, Stephen W. (2002) Online Student Evaluation of Instruction: An Investigation of Non-Response Bias. Paper Presented at the 42nd Annual Forum of the Association for Institutional Research, June 2002, Toronto Canada.

Webster, Russell J., Benton, Steve, and Gross, Amy (2010) Online Versus Paper Survey Delivery of College Student Ratings of Instruction, paper presented by Kansas State University and The IDEA Center at the 2010 American Educational Research Associate (AERA) Annual Meeting, Denver, CO.

Appendices

Appendix 1: 2004 Online SRTE Committee

Online SRTE Committee 2004 (alphabetical order, with 2004 affiliations)

- Nicole Belolan (Undergraduate Student, American Studies, Liberal Arts Undergraduate Council)
- Renata Engel (Chair, Schreyer Institute for Teaching Excellence, Executive Director)
- Leonard Berkowitz (Senator, Committee on Faculty Affairs, Professor, Penn State York)
- Will Kerr (Schreyer Institute for Teaching Excellence, Programmer)
- Ralph Locklin (Schreyer Institute for Teaching Excellence, Measurement Specialist)
- Jeff Markowitz (undergraduate student, Biobehavioral Health, Schreyer Honors Scholar, Commonwealth Campuses Student Government, Academic Affairs Director)
- Ron Rash (Academic Information Systems, Information Technology Services)
- Karen Schultz (Academic Information Systems, Information Technology Services)
- Al Turgeon (Professor, Agricultural Sciences)
- Rebecca Young (Office of the Vice Provost for Academic Affairs)

Administering Online SRTE Subcommittee

- Renata Engel
- Nicole Belolan
- Judy Bowman (Penn State York, Campus SRTE Representative)
- Will Kerr
- Debbie Lucas (Department of Biology, SRTE Representative)
- Jeff Markowitz
- Ron Rash
- Karen Schultz
- Cindy Brewer (Committee on Faculty Affairs, University Faculty Senate)
- Peter DeVries (Academic Information Systems, Information Technology Services)

Online SRTE Reporting Subcommittee

Leonard Berkowitz
Cindy Brewer
Renata Engel
Ralph Locklin
Dean Snow (Head, Department of Anthropology)
Alfred Turgeon
Rebecca Young

Appendix 2: Online Student Ratings at Peer Institutions

Penn State is a member of the Committee on Institutional Cooperation (CIC), which includes schools that are members of the "Big 10" college athletic conference, plus the University of Chicago. Twelve universities are included in this group and in 2011 it will also include the University of Nebraska - Lincoln. The CIC includes what many consider to be our closest peer institutions. In July 2010, the Schreyer Institute asked these institutions to share their online student ratings response rates.

Indiana University - Bloomington

Only a few departments have made the transition and no attempt has been made to summarize response rates. "There is now some movement on our campus toward establishing a dedicated online system, and eventually doing all evaluations that way." (J. David Perry, Director, IUB Evaluation Services & Testing (7-20-10).

Michigan State University

No response

Northwestern University

No response

Ohio State University

	# Possible Submissions	# Submissions	Response Rate
Winter 2009 (paper)	189,847	119,933	63%
Winter 2009 (online)	57,698	18,784	33%
Autumn 2009 (online)	251,095	102,881	41%
Winter 2010 (online)	236,143	95,595	41%

Purdue University

	# Possible Submissions*	# Submissions	Response Rate
paper			69%
Fall 2008 (online)	90,546	57.243	63%
Spring 2009 (online)	142,976	86,361	60%
Fall 2009 (online)	198,164	120,509	61%

* Reflects gradual rollout of online student ratings. Scores are identical for paper and online.

University of Illinois—Urbana Champaign

Instructor & Course Evaluation System: ICES Online Pilot

- 5,000 course sections
- 170,000 submissions

	Response Rate	Item #1 (instructor)		Item #2 (course)	
		Mean	s.d.	Mean	s.d.
ONLINE Su07-Sp09	54%	3.9	1.1	3.8	1.1
PAPER Su06-Sp07	~66%	4.1	0.7	3.9	0.8

University of Iowa

Pilot project to begin fall 2010.

University of Michigan

Fall 2009 (online): 59%
Winter 2010 (online): 54.4%

University of Minnesota

"[T]here is quite a difference in response rate for the two types of evaluation delivery."
(Ole R. Gram, Office of the Senior Vice President for Academic Affairs and Provost, University of Minnesota, 7-20-10)

University of Wisconsin-Madison

"The overall response rate for online course evaluations has been about 65% - 70%."
http://testing.wisc.edu/online%20course%20evals.html

Appendix 3: Strategies to Increase Response Rates

The list below was compiled at the request of numerous Penn State faculty who wanted to know what strategies other institutions are using to prompt an increase in online student ratings response rates. The strategies listed should not be considered suggestions or an endorsement by the Schreyer Institute or Penn State.

Strategies used by other colleges and universities or recommended by student ratings administrators:

- Instructors encourage students to complete the ratings.
- Instructors let students know they are interested in students' feedback.
- Instructors discuss how they have used past ratings results to improve their courses.
- University communications assure students that responses are confidential and that results are not released to faculty until after final grades are posted.
- Provide links to the student ratings website on homepages (university, campus, college, course). Student link to Penn State's Online SRTEs http://rateteaching.psu.edu.
- Send e-mail reminders to non-responders (Penn State does this already).
- Implement a university-wide marketing campaign, advertise or publicize the SRTE website and the importance of student feedback in general through posters, flyers, etc.
- Foster a response rate competition by publishing college, campus, or department response rates (but not rates for individual courses); update response rates throughout offering period (Penn State provides this information for faculty on their ANGEL My Profile site).
- Provide budgetary incentives for academic units with the highest or most improved response rates.
- Automatically enroll participants in a raffle for prizes (e.g., iPods, laptops, bookstore certificates, plane tickets).
- Provide participants access to information of interest, e.g.
  - ratings for all courses on core items
  - ratings for select questions from all courses
  - early access to grades
- Offer a registration advantage to participants (e.g. each student gets a certain number of points to bid on elective courses or on multiple-section courses).

Strategies to Avoid

- Making participation a course requirement
- Offering extra credit for participation
- Mentioning the role of student ratings in decisions about faculty salary, promotion, or tenure

Faculty who provide incentives to their students risk being misinterpreted as manipulative or pressuring students for high scores. Many higher education scholars recommend that grades be clearly based on evaluations of students' *academic performance* and likewise, they discourage the practice of assigning points or grades for compliance with requirements unrelated to the course or student learning.