TO: David Marshall, Executive Vice Chancellor

FROM: Jeffrey Stopple, Associate Vice Chancellor for Undergraduate Education
       Co-Chair, Ad Hoc Committee on Online Course Evaluation

       George H. Michaels, Executive Director, Instructional Development
       Co-Chair, Ad Hoc Committee on Online Course Evaluation

       **Ad Hoc Committee on Online Course Evaluation Members**
       Matt Hall   Associate Vice Chancellor, Information Technology, CIO
       Glenn Beltz Associate Dean, Undergraduate Studies, Engineering
       John Gilbert Undergraduate Vice Chair, Computer Science
       Madeleine Sorapure Director, Writing Program
       David Morrison Professor, Mathematics and Physics
       Jennifer Sorkin Associate Professor, History of Art & Architecture
       Cierra Sorin  GSA Executive Committee Representative
       Brooke Kopel  AS Executive Committee Representative
       Cindy Doherty Director, Academic Personnel
       Lisa Berry   Senior Instructional Consultant, Instructional Development
       Mindy Colin  Instructional Consultant, Instructional Development
       Toby Lazarowitz Executive Assistant to the Executive Vice Chancellor
       Bret Brinkman Director, Instructional Technology Services, LSIT

RE: Recommendations on the Adoption of Online Course Evaluations
Introduction

Nine years ago, Executive Vice Chancellor Gene Lucas convened faculty, administrators, students, and staff to form an ad hoc committee to consider moving the ESCI course evaluations online. Representatives from key Academic Senate committees and other faculty joined with administrators and staff from Instructional Development and Letters and Science Instructional Technology to assess the possibility and desirability of delivering evaluations online.

Although there was consensus to launch a pilot project, the Academic Senate expressed reservations (in 2011, and after further review, in 2014) about moving to online evaluations. In 2014, the Senate recommended that a more extensive pilot be conducted. Since then, some additional departments have elected to join the pilot, which has become a sort of longitudinal study, with the continued requirement that any department must participate consistently for specific categories of instructor. Senate concerns focused on the possible decline in student participation and indications that scores might worsen and therefore have an adverse effect on Academic Personnel merit reviews.

In April 2018, almost a decade after the campus began these considerations, Executive Vice Chancellor David Marshall reconvened another ad hoc committee to take a fresh look at the possibility of moving forward with ESCI online. The committee was formed in consultation with the Academic Senate, with representatives from Undergraduate Council, Graduate Council, the Committee on Academic Personnel, the Committee on Research and Instructional Resources, Instructional Development, Academic Personnel, Associated Students, and the Graduate Student Association. To emphasize, the charge of this committee is only to investigate questions around moving the current paper based ESCI to the online format.

In May of 2018, Undergraduate Council proposed the formation of a different ad hoc committee to make an assessment of whether Academic Senate should continue to support ESCIs, revise them, or possibly replace them with other modes of teaching assessment. Their concerns were for both the quality of instruction and the promotion/merit-review process at UCSB. Specifically, the statistical limitations of the ESCIs; misunderstandings about what they measure; and the body of empirical evidence demonstrating that student evaluations like the ESCI display biases against faculty of color, female faculty, and faculty from other underrepresented groups. That committee’s work is still in progress.

Areas of Analysis

The Committee chose to focus on four areas of new analysis. This approach was agreed upon at the first meeting of the Committee in July 2018, and was based on the members’ review of the most recent previous report on the ESCI Online Pilot program. That report was filed with the Office of the EVC in December 2015 and is included for reference as Appendix D of this document. The four areas of analysis were:
Ad Hoc Committee on Online Course Evaluation - Pilot Program Recommendations 2018-2019

- Updated literature review on the subject
- Analysis of the data collected as part of the annual survey of student opinion on online course evaluations collected every Winter quarter.
- Analysis of response rate data for the expanded pilot program data set.
- Analysis of faculty/course ratings for the expanded pilot program data set.

The results of these analyses are included as appendices to this document. The updated literature review and annotated bibliography is presented in Appendix A. The analysis of student feedback on online course evaluation is presented in Appendix B. The analyses of response rate patterns and the analysis of faculty/course ratings are included in Appendix C. Brief summary findings of those analyses are presented below.

The literature review, per the Committee’s suggestion, focused on the four questions below; the summary of findings is presented below each question.

- How to improve response rates?
  Suggestions for improving response rates from the literature included: longer survey windows; communicate use/utility; solicit constructive feedback (and train students to give); keep surveys short; central survey administration to decrease survey fatigue; ensure/communicate anonymity; allow time in class; offer incentives. Of these UCSB already does all, both for online and on paper, with the exception of offering incentives to students. Current technology does not support incentivizing individual students (for example, requiring an individual to complete or decline to complete the online ESCI before seeing their grades on GOLD). However, the students collectively may be incentivized if participation reaches some set level.

  Following Nutly’s (2008) recommendation, we recommend suggesting that faculty consider offering an incentive if the class as a whole achieves, or exceeds, an 80% response rate on the course evaluation (Appendix C, slide on p. 16). Incentives will vary; examples are extra office hours or a review session, the option to take a 3 by 5 notecard to the final, or some small number of additional participation points to student’s scores for the quarter. To help facilitate this strategy for faculty who like to employ it, the ESCI Online system already allows faculty to log in during the survey window to check their real time response rates. In addition, beginning in Spring 2019, ESCI Online will automatically send instructors an email when the survey window has closed for the quarter with an embedded summary table of response rates for all of their courses that were surveyed.

- Are online course evaluations better conducted in class or on a student's own time?
  No research could be found to address the question of how ratings or comments are influenced by students completing the evaluations in class or on a student’s own time. It is well-documented that response rates are improved when class time is allocated to complete the online evaluations.

- Are there meaningful effects on ratings if online response rates are lower than for paper surveys?
  The literature related to measuring the effects on ratings of generally lower online response rates yielded only slight differences between paper and online course evaluation surveys. There was
also about an even split between studies that showed slightly better ratings with online instruments, and those that showed the opposite. These results are consistent with our own comparative analysis presented in Appendix C.

● Is there any specific bias in online course evaluations?
The literature on bias in course evaluations is more generally focussed on issues of inherent bias potentially inherent in all forms of course evaluation, and not specifically confined to online evaluations. The whole issue of how best to deal with issues of bias, while examined in the literature review, is not part of the purview of this committee, and we have deferred issues related to that topic to the Senate Ad Hoc Committee on Course Evaluation.

The analysis of student feedback on online course evaluations was based on a survey that has been conducted every Winter quarter since 2012. The sample set covered the six year period between 2012 and Winter 2018. In aggregate over that time period the strong majority of respondents (78%) either agree or strongly agree that online course evaluations surveys are an improvement over paper surveys. Similarly, 55% of respondents agreed with the statement that online surveys allowed them to provide more thoughtful feedback. Students also feel confident that their anonymity is protected using online evaluations. Finally, there has been a steady increase in student ownership of mobile computing devices over that time period, that would allow them to participate in an online course evaluation in class. This was a significant concern when the pilot was started ten years ago, but the combination of less expensive devices and much improved wireless network access on campus has largely eliminated these issues as a concern.

The analysis of response rate patterns and effects on ratings that was conducted on the much expanded data set from the ESCI Online pilot proved to be consistent with the same analysis that was done for the 2015 report. In short, response rates for online surveys overall are lower than they have been on paper (mean = 73.8% on paper; mean = 54.7% online). In both cases, however, there is a very high degree of variability, much of which can be attributed to differences between departments. The expanded data did show a promising increase in response rates after the major expansion of the number of participating departments in Fall 2017. Mean response rates on paper for 2017 and 2018 were 69% and 75.8% respectively, while mean response rates online for 2017 and 2018 were 60.1% and 57.2% respectively. The online response rate means in those two most recent years were 13-19% higher than in the preceding two years (2015 and 2016).

The analysis of the effect of evaluation instrument (online or paper) on instructor ratings was also consistent with the patterns identified in the 2015 study. There is no strong indication of a bimodal distribution of responses, no indication that only the exceptionally pleased and the aggrieved students participate, and no consistent pattern of meaningfully better or worse ratings between online and paper surveys, even with generally lower response rates for online surveys. However, the analysis of student ratings by class size revealed that lower response rates may be problematic in smaller classes, where each student rating carries more weight, and therefore ratings using the online evaluation instrument were more variable and less reliable. The impact
of low response rates in small classes is an issue for both online and paper based ESCI. As noted above, online offers more options than paper to incentivize classes to increase response rates.

**Outside the scope**

Instructional Development was asked to determine the number pages and hours of staff time consumed by paper ESCIs each year. Their estimate is over 600,000 pages, and over 6,000 hours of staff time – more than 3 FTE. Although clearly a consideration for the campus as a whole, the ad hoc committee did not take this into consideration.

**Recommendations**

Based on the new analyses conducted on a much larger data set of results from the ESCI Online Pilot, the Committee strongly recommends:

- Mandate that all departments switch from paper to online course surveys, phased in over a five year period, in order to accommodate the work involved in on-boarding departments and the lead time that it requires.

We concur with all of the recommendations from the original Ad Hoc Committee on Online Course Evaluations with one reversal and one addition.

- We reverse the recommendation not to try to conduct the online surveys in class. That original recommendation was driven by the less than reliable state of wireless networking in the general assignment classrooms, and concerns about fairness for students who may not have had a suitable device to bring to class in order to participate. In the intervening six years both of those situations have dramatically improved, and this Committee feels that they are no longer a concern. Thus we recommend faculty members conduct their ESCI Online surveys in class in an effort to boost response rates. This does not preclude participation later by students not in class or without a device.
- As described in detail earlier, we recommend that faculty offer an incentive if the class as a whole achieves, or exceeds, an 80% response rate on the course evaluation.
APPENDIX A

Updated Literature Review and Presentation on Takeaways

This article investigates the population of respondents vs. non-respondents in online student evaluations of teaching. The authors find that non-respondents are different population than respondents. Response correlates with female, white, higher grade earners, salient courses. The authors do not investigate how these response/non-response patterns may influence bias in evaluation of instructors based on gender, age, etc.

Abstract: Technological advances have enabled institutions of higher education to administer course evaluations online, forgoing the traditional paper-and-pencil methods. Consequently, many of these institutions suffer from low response rates, but little research is available on this topic. To increase understanding about course evaluation participation in the online environment, this study examined over 22,000 undergraduates to whom the university administered about 135,000 evaluations. Multilevel models were constructed to analyze the data, and several variables emerged as significant predictors of participation. The results were mostly consistent with previous research and aligned with theories of survey nonresponse. However, the integration of uncommon variables provided new perspectives about course evaluations in particular. Implications for research and practical applications for institutions are also addressed, including ways to combat survey fatigue, increase the salience of the survey, and increase participation in online course evaluations. (PsycINFO Database Record (c) 2016 APA, all rights reserved)


The literature review revealed several studies that found no statistically significant differences between delivery modes. Two also noted that students provided more comments in the online forms. Response rates varied widely. The University of Kentucky College of Pharmacy, driven by the faculty’s desire for more timely return of results (3 - 4 months typically), launched a pilot study of online evaluations in 3 courses. The response rates for the 3 courses were 85%, 89%, and 75%. The 9 courses using the paper forms averaged an 80% response rate (consistent with the 2 previous years also about 80%). The comments on the online forms were more frequent and longer than the paper forms. Students liked the online form better than the paper form and thought they could provide more effective and constructive feedback online.

Synopsis from Innovate: “Many administrators are moving toward using online student evaluations to assess courses and instructors, but critics of the practice fear that the online format will only result in lower levels of student participation. Joan Anderson, Gary Brown, and Stephen Spaeth claim that such a concern often fails to acknowledge how the evaluation process already suffers from substantial lack of engagement on the part of students as well as instructors; the online format, they assert, merely inherits the fundamental problem of perceived irrelevance in the process itself. After addressing the reasons behind this problem and discussing how well-designed online evaluations can still make a positive difference, the authors describe the development and implementation of a comprehensive, college-wide online evaluation survey at Washington State University’s College of Agricultural, Human, and Natural Resources. In reviewing the survey results, they found that class size, academic discipline, and distribution method played a negligible role in student response rates. However, they found that variances in response rate were significantly influenced by the relative level of participation among faculty members and department heads in the original development of the survey. The authors maintain that online surveys can make the process more relevant and meaningful to students, but they conclude that eliciting greater response rates will still require sustained support, involvement, and advocacy by faculty members and administrators.”


This paper provides a summary of the current research in online vs. paper evaluations as well as results from a student to compare the feedback results. The same form was given to 46 section pairings – one paper and one online. The online response rate was 31% (392 out of 1276 possible responses) and the paper was 69% (972 out of 1415). No significant difference was found in the quantitative ratings between the two methods. They examined the differences on an “overall effectiveness” question in rating for faculty who were above the college average and then for faculty who were below the college average. Faculty who were above the average were scored slightly lower online and the faculty who were below the college average were scored higher online. There was no significant difference in the number of students giving open-ended feedback online, however, there was a significant increase in the length of open-ended feedback online.


The Department of Policy Analysis and Management a Cornell University did a study of course evaluation data from 1998 - 2001. Using the same form, data was analyzed from

29 courses (20 using the paper version, 9 using the online version). The study examined response rates and mean scores between the methods. While specific response rates varied, online was typically lower than the paper form. For example, in fall 2000 paper was 69% compared with 47% online. Using a 5-point scale on their 13 questions, 4 questions had a significant difference in mean scores between methods. This was a greater than 0.10 difference with the web having the higher mean score. The other 9 questions had a less than 0.10 difference in mean scores again with web having the higher means.


Abstract: One way many universities have approached the process of better understanding and meeting the needs of their students is through student evaluations. The evaluation data provide not only diagnostic feedback but also useful information in terms of the quality of learning and teaching experiences. In an effective quality cycle, the data gathered are analysed and used to make improvements. This is often referred to as ‘closing the loop’. However, for any evaluation data to be of value an important prerequisite for ‘closing the loop’ is that response rates should be sufficiently high to be representative of the student cohort. This paper describes how a faculty within Monash University utilising only web based surveys developed a successful communication strategy to elicit staff and student participation in the unit (subject) evaluation process which achieved a response rate as high as 83.2%.


Abstract: Students' contact with university is becoming increasingly computer-mediated. For instance, students can enrol, check exam marks, even attend some lectures online. Many universities now conduct at least some of their student evaluations of teaching or units online instead of using traditional paper questionnaires. Given the move to more flexible learning environments, with no requirement for students to attend all or any lectures, institutions are beginning to consider whether it is possible to move to online only systems of collecting student feedback. Whilst lower response rates appear to be a major concern with online surveys, commentators have pointed to several important advantages over paper forms: reduced paper use, reduced impact on class teaching time, less time spent scanning paper forms etc. It has also been suggested that the responses to open-ended questions provided online are more frequent, lengthy and thoughtful. This paper looks in more detail at the issue of student comments and reports the findings of several comparisons between paper and online approaches, looking at students enrolled internally and externally, and also recent graduates. The place of student comments as a valuable source of data is considered, along with the implications for an improvement in qualitative data with respect to the more quantitative aspects of student feedback.
Cummings, R. and Ballatyne, C. (1999). “Student feedback on teaching: Online! On target?” Paper presented at the Australasian Society Annual Conference, October, 1999. Murdoch University School of Engineering ran a pilot in 1999 of online course evaluations using the same form online as on paper. Students found the online form easier, faster, and felt it offered greater anonymity. The school has a 50% mandate for response rate in course evaluations. Typically paper evaluations had a 65% response rate. The online pilot averaged 31% with 4 of the 18 courses over the 50% mandate. The response rate range was a wide 3% to 100%. Because the pilot was inadequately promoted, some faculty didn’t know they were using online forms and didn’t adequately prepare students. Students noted that they felt no pressure to fill out the online evaluations. The investigators concluded that the quality of responses was the same because they received the same amount of comments online which is what is used most from the evaluation form.


Abstract: This study compares student evaluations of faculty teaching that were completed in-class with those collected online. The two methods of evaluation were compared on response rates and on evaluation scores. In addition, this study investigates whether treatments or incentives can affect the response to online evaluations. It was found that the response rate to the online survey was generally lower than that to the in-class survey. When a grade incentive was used to encourage response to the online survey, a response rate was achieved that was comparable with that to the in-class survey. Additionally, the study found that online evaluations do not produce significantly different mean evaluation scores than traditional in-class evaluations, even when different incentives are offered to students who are asked to complete online evaluations.


The College of Business And Economics at California State University, Northridge did a study with 16 professors to see how the method of evaluation affects response rate and if online treatments (incentives) affect the response rate. Each professor taught 2 sections of the same undergraduate business course. The same form was used in both methods. Instructors were randomly assigned into 1 of 4 groups using different incentives: 0.25% grade incentive for completion of an online evaluation (4 courses), in-class demonstration on how to do the online evaluation (2 courses), if 2/3 of the class submitted online evaluations students would receive their final grades early (2 courses), or a control group (8 courses). The online evaluations averaged a 43% response rate and the paper evaluations averaged 75%. Looking at just the control group, their average

response rate was 29%. In the individual cases the incentives had the effect of increasing response rate (grade incentive 87% response rate, demonstration 53%, and early final grade 51%).


Abstract: Substantial efforts have been made recently to compare the effectiveness of traditional course formats to alternative formats (most often, online delivery compared to traditional on-site delivery). This study examines, not the delivery format but rather the evaluation format. It compares traditional paper and pencil methods for course evaluation with electronic methods. Eleven instructors took part in the study. Each instructor taught two sections of the same course; at the end, one course received an online course evaluation, the other a traditional pencil and paper evaluation. Enrollment in these 22 sections was 519 students. Researchers analyzed open-ended comments as well as quantitative rankings for the course evaluations. Researchers found no significant differences in numerical rankings between the two evaluation formats. However, differences were found in number and length of comments, the ratio of positive to negative comments, and the ratio of formative to summative comments. Students completing faculty evaluations online wrote more comments, and the comments were more often formative (defined as a comment that gave specific reasons for judgment so that the instructor knew what the student was suggesting be kept or changed) in nature.

*Emery, L., Head, T., Zeckoski, A., and Yu Borkowski, E. (2008) “Deploying an Open Source, Online Evaluation System: Multiple Experiences.” Presentation at Educause 2008, October 31, Orlando, FL. Four institutions, University of Michigan Ann Arbor, Virginia Tech, University of Cambridge and University of Maryland, collaborated on an open source online evaluation system within Sakai. Response rates in the various pilots ranged from 32% to 79%. They found the key benefits of online evaluations to be security, validity, efficiency, cost-savings, rapid results turnaround and higher quality student comments.

*Ernst, D. (2006) “Student Evaluations: A Comparison of Online vs. Paper Data Collection.” Presentation at Educause, 2006, October 10, Dallas, TX. The College of Education and Human Development at the University of Minnesota did a study on 314 class pairs (14,154 student evaluations) from fall 2002 to fall 2004. The goals were to see if there is a difference in response rate, a difference in response distributions, a difference in average ratings between the two methods and what are the common perceptions of each method. In the study group the online form averaged a 56% response rate whereas the paper version averaged 77%. Slightly more students responded on the high and low ends of the 7-point scale than did in the middle. There was no significant difference in the mean rating on 4 required questions.


**Abstract:** Evaluation of teaching effectiveness is considered a critical element in determining whether or not faculty members are retained at higher education institutions; academic milestones such as tenure and promotion often require documentation of the quality of faculty teaching. As methods of assessing teaching effectiveness evolve, concerns about the equity of alternative methods arise. This study compared two methods of student evaluations of faculty, online versus the traditional paper format. Although the student response rate for online evaluations was lower, evaluation scoring patterns were similar for both methods. The findings suggest that conducting faculty evaluations online may be a suitable alternative to the traditional, paper-based approach.

Fraze, S., Hardin, K., Brashears, T., Smith, J., Lockaby, J. (2002) “The Effects Of Delivery Mode Upon Survey Response Rate And Perceived Attitudes Of Texas Agri - Science Teachers.” Paper presented at the National Agricultural Education Research Conference, December 11 - 13, Las Vegas, NV , Texas Tech University studied 3 modes of surveying a random group of Texas Agri Science teachers. The 3 modes were e-mail, web, and paper. No significant difference in the reliability of the responses was found. However the response rates were 60%, 43% and 27% for paper, web and e-mail respectively.


**Abstract:** The purpose of this study is to compare the results of paper and online evaluations. The following analysis examines data from six departments of the School of Business Administration during a programmed switch from paper to online evaluations. The courses that participated in this study were divided and compared in the following manner: advanced and core classes, large and small sections, and courses taught by full-time and part-time faculty. The data was collected over a one-year period and contrasts the Spring 2008 and 2009 semesters, during which a total of 4,424 evaluations were reviewed. In addition, data on the years from 2005 to 2008 are provided as a comparison benchmark of typical responses collected when paper evaluations were used. The

conclusions of this study show that while a drop in response rate did occur when the switch was made, no significant change in instructor and course ratings was observed. Furthermore, the students who did complete online evaluations provided lengthier and more numerous comments.


The Master of Administrative Science program at Fairleigh Dickinson University performed a study on courses taught by adjunct faculty. The online evaluations received a 61% response rate and the in-class evaluations received a 82.1% response rate. They found that the online evaluations received twice as many comments (counting total words) as the in-class evaluations. On the question about “materials being clearly presented” (focused on the faculty member) the variation in mean scores in online and in-class was 0.33 on a 5-point scale with online having a less-positive rating. This is a statistically significant difference. Administrators noted that both means were better than the “agree” and were not considered poor ratings.

*Layne B.H., DeCristofor J.R., McGinty D (1999). “Electronic versus traditional student ratings of instruction.” Res Higher Educ. 1999; 40:221 - 32. At a southeastern university 66 courses made up of 2453 students did a comparison of response effects between paper-and-pencil and online using the same form. Half did online and half did paper-and-pencil forms. The online response rate was 47% and the traditional group was 60%. Also, 76% of the online forms provided comments compared to 50% of the traditional forms. No significant difference was found in methods.


Georgia State University College of Business ran a voluntary pilot from 2002 to 2003 using an identical online version of their paper course evaluation form in the Department of Computer Information Systems. Faculty feared an online form would yield lower scores and lower response rates. In particular, the fear was that few students would submit online evaluations, poor students would “take revenge” on the faculty and good students wouldn’t bother. The paper form had a 67% response rate and the online form had an 82% response rate. This likely due to the fact that the CIS department had easy access to computer labs for students to take the evaluations online. Using a question on teacher effectiveness, the study found no significant difference between the methods. Good students participated in the same numbers and weaker students did fewer online evaluations.

Summer School on Survey Methodology. The paper presents a short literature review comparing online evaluations with paper. The Economics department at University of Belgrade, Serbia conducted a small pilot in a course of 800 students in May of 2006. Half the students received paper evaluations in class and half were directed to complete an identical online evaluation. The paper evaluation received a 92.5% response rate and the online received a 52% response rate after an incentive was introduced. They found that nearly twice as many students filled out the open-ended question online when compared to the paper group. On the instructor-related questions they found a variation of 0.09 to 0.22 on a 10-point scale. No statistical analysis was done for significance.

*Matz, C. (1999). “Administration of web versus paper surveys: Mode effects and response rates.” Masters Research Paper, University of North Carolina at Chapel Hill. (ERIC document ED439694). In a survey of academic reference librarians in North Carolina, Matz found no significant difference in response contents between the methods used. The online form had a 33% response rate and the paper form had a 43% response rate.

*Monsen, S., Woo, W., Mahan, C. Miller, G. & W (2005). “Online Course Evaluations: Lessons Learned.” Presentation at The C ALI Conference for Law School Computing 2005. Yale Law started online course evaluations in 2001 with a less than 20% response rate. The current 8-question form is run by student representatives and has a 90% response rate. Students cannot see their grades until they fill out the evaluation. Northwestern University School of Law started online course evaluations in 2004. So far they have a 68% response rate which compares to a 70-80% paper response rate. Northwestern is against using any penalties (withholding information from a student until they fill out an evaluation). The University of Denver Sturm College started online course evaluations in 2002 with a pilot of 10 courses. The pilot had an 83% response rate. Continuing into 2003 the pilot expanded to 80 courses (with an 81% response rate) and then expanded to all of their offerings (with a 64% response rate). Currently they maintain a response rate around 70%. Duke Law started online course evaluations in 2003 when their scantron machine broke and the expense of replacing was too great. They proposed a goal of 70% response rate and used the same form online. The first term averaged a 66% response rate (with 29% of the 82 courses reaching the 70% goal). In spring 2004 the average was 60% (with 30% of the 119 courses reaching the 70% goal). In fall 2004 the average was 52% (with 8% of the 93 courses reaching the 70% goal). In spring 2005, after dropping non-law students from the pool, the average was 67% (with 41% of the 117 courses reaching the 70% goal). The school is considering several penalties for failure to fill out an evaluation – withholding registration, withholding grades, or withholding free printing.

*Norris, J., Conn, C. (2005). “Investigating strategies for increasing student response rates to online-delivered course evaluations.” Quarterly Review of Distance Education 2005; 6 (1) p13 - 32 (ProQuest document ID 975834871). This paper reports the findings of 2 studies done at Northern Arizona State University. The first study looked at historic data from 2000 - 2002 to examine student responses to online course evaluations in 1108
course sections. This group had an average response rate of 31%. A follow-up questionnaire was sent to 50 faculty in the group to explore what strategies improved response rate. These results informed the second study on 39 online course sections and 21 sections of a required freshman face-to-face course. The second study used some basic strategies (no penalty strategies) in the implementation of the online course evaluations: 2 weeks before the end of the course the URL to evaluation was posted in the course management system, an announcement containing a statement of course evaluation value and due date was sent in a method appropriate to the class (email, online syllabus or discussion board), and a reminder email was sent 1 week before the class ended containing the URL and due date. The 39 online course sections averaged a 74% response rate and the 21 face-to-face courses averaged a 67% response rate. In addition, 11 sections of the face-to-face course used paper evaluations and received a 83% response rate. These suggestions are very similar to the emerging findings from the TLT Group’s BeTA project.


Quantifies response rate needed for different class sizes to be generalizable using statistical analyses and different confidence intervals. Also provides recommendations for improving response rates.

Abstract: This article is about differences between, and the adequacy of, response rates to online and paper-based course and teaching evaluation surveys. Its aim is to provide practical guidance on these matters. The first part of the article gives an overview of online surveying in general, a review of data relating to survey response rates and practical advice to help boost response rates. The second part of the article discusses when a response rate may be considered large enough for the survey data to provide adequate evidence for accountability and improvement purposes. The article ends with suggestions for improving the effectiveness of evaluation strategy. These suggestions are: to seek to obtain the highest response rates possible to all surveys; to take account of probable effects of survey design and methods on the feedback obtained when interpreting that feedback; and to enhance this action by making use of data derived from multiple methods of gathering feedback.

http://www.marquette.edu/oira/ceval/ Marquette University moved from a copyrighted instrument, IAS, to their own instrument, MOCES. Because of the copyright concerns the new instrument has re-worded items that maintain the intent of the IAS items. In spring semester of 2008 a pilot was conducted in 124 course sections with 3837 students. They evaluated the effectiveness of an online approach versus paper and pencil and the software used to deliver the evaluations. Response rates online were lower in 3 of the 5

pilot departments, comparable in 1 and higher in 1 when compared to 3 semester averages of paper and pencil forms. A “power analysis” of the response rates revealed the rates were high enough of 95% confidence in the results. There was no significant difference in the mean ratings for the 4 core questions between the old IAS form and the MOCES online form.

*Sax, L., Gilmartin, S., Keup, J., Bryant, A., and Plecha, M. (2002). Findings from the 2001 pilot administration of Your First College Year (YFCY): National norms. Higher Education Research Institute, University of California. The YFCY distributed its survey that assesses student development during the first year in college using 3 methods: online, online or paper, and paper. In a pool of 57 schools, 16 used the alternative methods of distribution. The study found no significant difference in responses between the methods. The response rate overall was 21%. The online only method response rate was 17% and the online or paper group had a 24% response rate.

*Schawitch, M. (2005) “Online Course Evaluations: One Institute’s Success in Transitioning from a Paper Process to a Completely Electronic Process!” Presentation at the Association for Institutional Research Forum, June 2005. The Rose - Hulman Institute of Technology piloted an online course evaluation in 2002 with a small group of faculty. Over the academic year the pilot had a 70% response rate. 77% of students preferred the online mode and faculty reacted positively to the pilot. In 2003 the entire campus adopted the online form. Over the 3 terms, the online evaluations had response rates of 86%, 78% and 67%. In 2004 the 3 terms had 75%, 71% and 67%. Historically paper evaluations had an 85 - 87% response rate. They are investigating various incentive possibilities.

*Thorpe, S. (2002) “Online Student Evaluation of Instruction: An Investigation of Non - Response Bias.” Paper presented at the 42nd Annual Forum of the Association of Institutional Research, June 2002. Drexel University studied whether significant differences exist in student responses to course evaluations given on paper and online in 3 courses. Response rates in the 3 classes for paper and online (respectively) were 37% and 45%, 44% and 50%, 70% and 37%. In comparing students who responded to the evaluations across the 3 courses the study found that women were more likely than men to respond, students who earned higher grades were more likely to respond, and students with a higher overall GPA were more likely to respond. For two courses the online evaluations had a slightly higher average item rating. For the other course 2 significant differences were found: students doing the online evaluation were less likely to participate actively and contribute thoughtfully during class and to attend class when compared to the paper evaluation group. But the responses overall were not significantly different.


This article summarizes the scholarly research on student course evaluations in three main sections: The validity and reliability of course evaluations, Online vs. paper course evaluations, Student perceptions of course evaluations, Effects of allowing students access to course evaluation data and Recommendations for improving response rates. The three recommendations for improving response rates are to 1) make evaluation a part of the course (most effective), 2) send reminder notices and 3) offer a small incentive.
ESCI ONLINE

UPDATES TO LITERATURE REVIEW
AD HOC COMMITTEE MEETING, 12/18/18
FOUR QUESTIONS

• How to improve response rates?
• Online SETs: in-class or at home?
• Effect of low response rates on scores?
• Bias in online SETs?
HOW TO IMPROVE RESPONSE RATES (NUTLY, 2008)

• Longer survey window
• Communicate use/utility
• Solicit constructive feedback (and train students to give)
• Keep surveys short
• Central survey administration to decrease survey fatigue
• Ensure/communicate anonymity
• Allow time in class
• Offer incentive
EFFECT OF DIFFERENT APPROACHES TO IMPROVE RR (GOODMAN ET AL, 2015)

<table>
<thead>
<tr>
<th>Tactic</th>
<th>N (%) who used tactic</th>
<th>Response rate for users (%)</th>
<th>Response rate for non-users (%)</th>
<th>Statistical results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Provided incentives to complete the evaluation</td>
<td>251 (22.3)</td>
<td>79</td>
<td>57</td>
<td>$t = 17.16, p &lt; 0.0001$</td>
</tr>
<tr>
<td>Reminded students during class to take evaluation</td>
<td>688 (60.9)</td>
<td>65</td>
<td>57</td>
<td>$t = 6.29, p &lt; 0.0001$</td>
</tr>
<tr>
<td>Took class time to have the students complete the evaluation</td>
<td>54 (4.8)</td>
<td>70</td>
<td>61</td>
<td>$t = 3.24, p = 0.0012$</td>
</tr>
<tr>
<td>Explained how the evaluations would be used to improve</td>
<td>645 (57.0)</td>
<td>65</td>
<td>57</td>
<td>$t = 6.58, p &lt; 0.0001$</td>
</tr>
<tr>
<td>Sent personal emails to students asking them to complete their evaluations</td>
<td>375 (33.1)</td>
<td>66</td>
<td>60</td>
<td>$t = 4.77, p &lt; 0.0001$</td>
</tr>
<tr>
<td>Posted reminder or assignment on Blackboard®</td>
<td>360 (31.7)</td>
<td>64</td>
<td>61</td>
<td>$t = 2.61, p = 0.0093$</td>
</tr>
<tr>
<td>Used formal mid-semester assessment process to get feedback</td>
<td>27 (2.4)</td>
<td>65</td>
<td>62</td>
<td>$t = 0.79, p = 0.4282$</td>
</tr>
<tr>
<td>Provided other type of mid-semester check-in with students</td>
<td>116 (10.2)</td>
<td>65</td>
<td>61</td>
<td>$t = 2.13, p = 0.0331$</td>
</tr>
<tr>
<td>Used classroom assessment techniques during semester to get feedback</td>
<td>64 (5.6)</td>
<td>61</td>
<td>62</td>
<td>$t = 0.43, p = 0.6688$</td>
</tr>
</tbody>
</table>

+ more tactics used = higher response rates
MORE ON INCENTIVES (GOODMAN ET AL, 2015)

<table>
<thead>
<tr>
<th>Threshold response</th>
<th>% of faculty with incentives who used threshold (and number)</th>
<th>Average % response</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>11 (8)</td>
<td>89</td>
</tr>
<tr>
<td>90% +</td>
<td>14 (10)</td>
<td>87</td>
</tr>
<tr>
<td>80% +</td>
<td>31 (22)</td>
<td>86</td>
</tr>
<tr>
<td>Sliding scale</td>
<td>11 (8)</td>
<td>73</td>
</tr>
<tr>
<td>70% +</td>
<td>17 (12)</td>
<td>68</td>
</tr>
<tr>
<td>60% +</td>
<td>4 (3)</td>
<td>62</td>
</tr>
<tr>
<td>Unknown</td>
<td>13 (9)</td>
<td>74</td>
</tr>
</tbody>
</table>

Moral: For class-wide threshold incentives, set threshold to at least 80% (e.g. if 80% of students complete SET, will add 0.25% to everyone’s grade)
### CRITICAL RR THRESHOLDS BY CLASS SIZE (NUTLY, 2008)

Table 3. Required response rates by class size.

<table>
<thead>
<tr>
<th>Total no. of students on the course</th>
<th>'Liberal conditions'</th>
<th>'Stringent conditions'</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Required no. of</td>
<td>Required no. of</td>
</tr>
<tr>
<td></td>
<td>respondents</td>
<td>respondents</td>
</tr>
<tr>
<td></td>
<td>Response rate</td>
<td>Response rate</td>
</tr>
<tr>
<td></td>
<td>required (%)</td>
<td>required (%)</td>
</tr>
<tr>
<td>10</td>
<td>7 75%</td>
<td>10 100%</td>
</tr>
<tr>
<td>20</td>
<td>12 58%</td>
<td>19 97%</td>
</tr>
<tr>
<td>30</td>
<td>14 48%</td>
<td>29 96%</td>
</tr>
<tr>
<td>40</td>
<td>16 40%</td>
<td>38 95%</td>
</tr>
<tr>
<td>50</td>
<td>17 35%</td>
<td>47 93%</td>
</tr>
<tr>
<td>60</td>
<td>18 31%</td>
<td>55 92%</td>
</tr>
<tr>
<td>70</td>
<td>19 28%</td>
<td>64 91%</td>
</tr>
<tr>
<td>80</td>
<td>20 25%</td>
<td>72 90%</td>
</tr>
<tr>
<td>90</td>
<td>21 23%</td>
<td>80 88%</td>
</tr>
<tr>
<td>100</td>
<td>21 21%</td>
<td>87 87%</td>
</tr>
<tr>
<td>150</td>
<td>23 15%</td>
<td>123 82%</td>
</tr>
<tr>
<td>200</td>
<td>23 12%</td>
<td>155 77%</td>
</tr>
<tr>
<td>250</td>
<td>24 10%</td>
<td>183 73%</td>
</tr>
<tr>
<td>300</td>
<td>24 8%</td>
<td>209 70%</td>
</tr>
<tr>
<td>500</td>
<td>25 5%</td>
<td>289 58%</td>
</tr>
<tr>
<td>750</td>
<td>25 3%</td>
<td>358 48%</td>
</tr>
<tr>
<td>1000</td>
<td>26 3%</td>
<td>406 41%</td>
</tr>
</tbody>
</table>
FOUR QUESTIONS

• How to improve response rates?
• Online SETs: in-class or at home?
• Effect of low response rates on scores?
• Bias in online SETs?
<table>
<thead>
<tr>
<th>Study</th>
<th>Research design</th>
<th>Difference by response rate</th>
<th>Difference by mean ratings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Avery et al. (<em>2006</em>)</td>
<td>Static group comparison design</td>
<td>33.3–77.8% online, 50.2–100% in-class</td>
<td>Item level: Slight difference in favour of online method. Four out of 13 yielded significant difference</td>
</tr>
<tr>
<td>Barkhi and Williams (<em>2010</em>)</td>
<td>Static group comparison design</td>
<td>~</td>
<td>Item level: Significant difference in favour of in-class method. Significantly more negative ratings and less positive ratings appear in online</td>
</tr>
<tr>
<td>Dommeyer, Baum, and Hanna (<em>2002</em>)</td>
<td>Two-group randomised experimental design</td>
<td>60% online, 92% in-class, significant difference</td>
<td>~</td>
</tr>
<tr>
<td>Dommeyer et al. (<em>2004</em>)</td>
<td>Two-group randomised experimental design</td>
<td>29% online, 70% in-class, significant difference</td>
<td>Overall: No significant difference</td>
</tr>
<tr>
<td>Donovan, Mader, and Shinsky (<em>2010</em>)</td>
<td>Two-group randomised experimental design</td>
<td>48% online, 52% in-class</td>
<td>Overall: Slight difference in favour of online method, not significant</td>
</tr>
<tr>
<td>Fike, Doyle, and Connelly (<em>2010</em>)</td>
<td>Two-group randomised experimental design</td>
<td>54.14% online, 68.29% in-class, significant difference</td>
<td>Overall: Significant difference, in favour of in-class method. One out of 18 yielded significant difference</td>
</tr>
<tr>
<td>Study</td>
<td>Type of design</td>
<td>Response Rates (Mean)</td>
<td>Conclusion</td>
</tr>
<tr>
<td>----------------------------</td>
<td>---------------------------------------</td>
<td>------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Gamliel and Davidovitz</td>
<td>Mixed-subject experimental design</td>
<td></td>
<td>Overall: No significant difference in mean ratings; however, more variation was found in online method</td>
</tr>
<tr>
<td>(2005)</td>
<td></td>
<td></td>
<td>than in-class method</td>
</tr>
<tr>
<td>Guder and Malliaris</td>
<td>Static group comparison design</td>
<td>Overall 25.99% drop from paper to online method</td>
<td>Overall: Slight difference in favour of in-class method, not significant</td>
</tr>
<tr>
<td>(2010)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Heath, Lawyer, and Rasmussen</td>
<td>Two-group randomised experimental</td>
<td>72.2% online</td>
<td>Overall: Slight difference in favour of in-class method, not significant</td>
</tr>
<tr>
<td>(2007)</td>
<td>design</td>
<td>81.5% in-class</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>no significant difference</td>
<td></td>
</tr>
<tr>
<td>Layne, DeCristoforo, and</td>
<td>Two-group randomised experimental</td>
<td>47.8% online</td>
<td>Dimension level: No significant difference in mean ratings</td>
</tr>
<tr>
<td>McIntyre (1999)</td>
<td>design</td>
<td>60.6% in-class</td>
<td></td>
</tr>
<tr>
<td>Liu (2006)</td>
<td>Two-group experimental design</td>
<td>100% in both. No significant difference was found*</td>
<td>Item level: Slight difference in favour of online method in 13 out of 17 items, not significant</td>
</tr>
<tr>
<td>Morrison (2013)</td>
<td>Two-group randomised experimental</td>
<td></td>
<td>Overall: Mean ratings of in-class were higher than the online method in 74.18% of the runs while only</td>
</tr>
<tr>
<td></td>
<td>design</td>
<td></td>
<td>17.89% were significant. More variation was found in online method than in-class method</td>
</tr>
<tr>
<td>Perrett (2013)</td>
<td>Two-group randomised experimental</td>
<td>60.4–71.1% online</td>
<td>Overall: Slight difference in favour of online method, not significant</td>
</tr>
<tr>
<td></td>
<td>design</td>
<td>61.8–68.8% in-class</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>no significant difference</td>
<td></td>
</tr>
<tr>
<td>Stowell, Addison, and Smith</td>
<td>Static group comparison design</td>
<td>61.4% online</td>
<td>Overall: Slight difference in favour of in-class method, not significant</td>
</tr>
<tr>
<td>(2012)</td>
<td></td>
<td>81.5% in-class</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>significant difference</td>
<td></td>
</tr>
<tr>
<td>Thorpe (2002)</td>
<td>Two-group randomised experimental</td>
<td>37–50% online</td>
<td>Item level: No pattern in findings. Three out of 17 yielded significant difference</td>
</tr>
<tr>
<td></td>
<td>design</td>
<td>37–70% in-class</td>
<td></td>
</tr>
<tr>
<td>Venette, Selinow, and</td>
<td>One-group experimental design</td>
<td></td>
<td>Item level: Slight difference in favour of in-class method in four out of six items, not significant</td>
</tr>
<tr>
<td>McIntyre (2010)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
FOUR QUESTIONS

• How to improve response rates?
• Online SETs: in-class or at home?
• Effect of low response rates on scores?
• Bias in online SETs?
BIAS IN ONLINE SET? (ADAMS-UMBACH, 2012)

- Non-respondents are different population than respondents
- Response correlates with female, white, higher grade earners, salient courses
- Attendance and eval (Capa-Aydin, 2014; Large university in Turkey)
HOLIDAY READING

• “Online and paper evaluations of courses: a literature review and case study,” Morrison, 2013.


• “Student evaluation of instruction: comparison between in-class and online methods,” Capa-Aydin, 2014

• “The adequacy of response rates to online and paper surveys: what can be done?” Nutly, 2008
APPENDIX B

Results from the Annual Winter Survey of Students’ Experience with Online ESCI, 2012-2018
Results from the Annual Winter Survey of Students’ Experience with Online ESCI, 2012-2018

Table of Contents

Results from the Annual Winter Survey of Students’ Experience with Online ESCI, 2012-2018 ... 1
Background .................................................................................................................................................. 3
Q1: The online end of quarter survey is an improvement over the traditional paper-based end of quarter survey. .......................................................................................................................... 6
Q2: I was able to provide more thoughtful feedback in this online format of the ESCI survey as compared to the paper-based survey. .......................................................................................................................... 7
Q3: I am confident that the online ESCI survey maintains my anonymity. ................................................. 10
Q4: Would you have a wireless device (laptop, smartphone, tablet) with you in class, that you could use to complete the ESCI survey in class? ................................................................................................. 12
Q5: What was the primary reason that made you decide to complete the online ESCI survey(s) for your course(s) this quarter? ......................................................................................................................................... 13
Q6: What ideas can you offer for improving response rates as the campus transitions to Online ESCI? ......................................................................................................................................... 14
Q7: Do you have any suggestions or concerns about the online ESCI survey? ........................................ 16
Summary .................................................................................................................................................. 17

Tables and Figures

Table 1 – Yearly breakdown of response rates.............................................................................................. 4
Table 2 – Yearly breakdown of participating departments ................................................................................ 4

Figure 1 – Improvement over paper: combined responses .............................................................................. 6
Figure 2 – Improvement over paper: comparison of response percentage per year ........................................ 6
Figure 3 – More thoughtful feedback: combined responses ............................................................................ 7
Figure 4 – More thoughtful feedback: comparison of response percentage per year ....................................... 7
Figure 5 – “Write better questions” aligned with responses to Q2 (percentages) ........................................... 8
Figure 6 – Frequency of codes explaining how to write better questions ......................................................... 9
Figure 7 – Confident that ESCI maintains my anonymity: combined responses ........................................... 10
Figure 8 – Confident that ESCI maintains my anonymity: comparison of response percentage per year ........................................................................................................................................... 10
Figure 9 – “Anonymity concerns” per year (normalized) ................................................................................ 11
Figure 10 – “Anonymity concerns” aligned with requests for better information and responses to Q3 ........................................................................................................................................... 11
Figure 11 – Device ownership....................................................................................................................... 12
Figure 12 – "Do during class" compared to concerns about "Access to mobile device" (normalized) ................................................................................................................................. 12
Figure 13 – Reason for completing Online ESCI: combined responses ................................. 13
Figure 14 – Reason for completing Online ESCI: comparison of response percentage per year . 13
Figure 15 – Most frequent ideas for improving response rates 2014-2018 (normalized) ........... 14
Figure 16 – All excerpts for "Do during class" organized by year and Q6 or Q7 (normalized)..... 15
Figure 17 – Most frequently mentioned suggestions and concerns, 2014-2018 (normalized) ... 16
Figure 18 – Frequency of major codes 2015-2018 (normalized) ............................................. 18
Background

The Feedback About ESCI Online survey (referred to in this report as the Annual Winter Survey) was first distributed in Fall quarter 2012, and has been distributed annually at the end of Winter quarter since 2013 to UCSB students who received an Online ESCI questionnaire in that same quarter. The purpose of the survey is to gauge student perceptions regarding the benefits and challenges of doing course evaluations online, as well as solicit student feedback and ideas about it.

The survey initially consisted of seven items: six scalar items and one text item. Two scalar items were removed in 2014 because they were technical questions that got resolved: one question asked about ideal length of time an online survey should be available to students, and another about whether students should be able to save a draft of their responses. Two new items replaced those that were removed: one scalar item that asks students to indicate the primary reason that they completed their Online ESCI, and one text item that asks for their ideas for improving response rates for Online ESCI.

Between 2012 and 2018 there were 28,038 Annual Winter Surveys distributed. The breakdown of respondents by year is represented in Table 1. Each student who received at least one Online ESCI survey in Winter quarter received the Annual Winter Survey as well. There were 12,993 total responses to the Annual Winter Survey, with half of those occurring in 2018 because of the influx of new departments to Online ESCI that year (see Table 2).

The response rate has been between 28%-51%, with an average of 46%, which we consider successful. The percentage of respondents each year who wrote text responses to the open-ended questions on the survey fluctuated greatly. We suspect that 2015 has the highest number of comments because it is the first year that Psychology joined Online ESCI, and the Psychology students made up 68% of the population being surveyed. We also suspect that 2018 had the lowest percentage of comments because it was the first year that more than half of the population being surveyed was from STEM departments, who are typically less loquacious than their humanities and social sciences peers.

1 All responses to open-ended questions that consisted solely of text such as n/a, no, or nothing were excluded from analyses conducted for this report.
Table 1 – Yearly breakdown of response rates

<table>
<thead>
<tr>
<th>Year</th>
<th>Number of surveys distributed</th>
<th>Number of scalar respondents</th>
<th>Response Rate</th>
<th>Number of comments from scalar respondents</th>
<th>Percentage of scalar respondents who wrote comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>unknown</td>
<td>196</td>
<td>n/a</td>
<td>11</td>
<td>5.6%</td>
</tr>
<tr>
<td>2013</td>
<td>2132</td>
<td>597</td>
<td>28%</td>
<td>75</td>
<td>12.5%</td>
</tr>
<tr>
<td>2014</td>
<td>1426</td>
<td>617</td>
<td>43%</td>
<td>291</td>
<td>47%</td>
</tr>
<tr>
<td>2015</td>
<td>3948</td>
<td>1665</td>
<td>42%</td>
<td>567</td>
<td>34%</td>
</tr>
<tr>
<td>2016</td>
<td>4164</td>
<td>1558</td>
<td>37%</td>
<td>413</td>
<td>26.5%</td>
</tr>
<tr>
<td>2017</td>
<td>4024</td>
<td>2054</td>
<td>51%</td>
<td>421</td>
<td>20.5%</td>
</tr>
<tr>
<td>2018</td>
<td>12344</td>
<td>6306</td>
<td>51%</td>
<td>482</td>
<td>7.6%</td>
</tr>
<tr>
<td>Total</td>
<td>28,038</td>
<td>12,993</td>
<td>46% (avg)</td>
<td>2,260</td>
<td>17.4% (avg)</td>
</tr>
</tbody>
</table>

Table 2 shows a yearly breakdown of the departments that participated in Online ESCI. Note that until 2018 the vast majority of participating departments are from the Humanities and Social Sciences. Between 2015-2017, 64-68% of surveys were distributed to students in Psychology courses. Fourteen departments joined the Pilot Online ESCI in Fall 2017, six of which are large STEM departments. Surveys from these STEM departments make up about 57% of the total number of surveys distributed in 2018, and almost 40% of those are from Math. While there is no way for us to know for certain that this same percentage of Psychology or STEM students participated in the Annual Winter Survey, it is safe to assume that the results from 2015 and 2018 may be different from other years because a large new population took the survey for the first time in each of those years.

Table 2 – Yearly breakdown of participating departments

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>ART</td>
<td>FHSP</td>
<td>FLMST</td>
<td>ART (new Q)</td>
<td>ART</td>
<td>ART</td>
<td>ART (new Q)</td>
<td>GEOG (new Q: TA, F)</td>
</tr>
<tr>
<td>FHSP</td>
<td>FHSP</td>
<td>MAT</td>
<td>FHSP</td>
<td>FHSP</td>
<td>FHSP</td>
<td>FHSP (new Q)</td>
<td>CHEM</td>
</tr>
<tr>
<td>FLMST STATS</td>
<td>PSY</td>
<td>LAIS</td>
<td>PSY</td>
<td>LAIS</td>
<td>PSY</td>
<td>LAIS</td>
<td>EEMB (new Q: TAs, F: paper)</td>
</tr>
<tr>
<td>ART</td>
<td>CNCSP</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ESM</td>
</tr>
<tr>
<td>FHSP</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>MATH (new Q: TAs)</td>
</tr>
<tr>
<td>FLMST</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>MCDB (new Q: TAs, F: paper)</td>
</tr>
<tr>
<td>ART</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>LING (new Q: TA, F)</td>
</tr>
<tr>
<td>FHSP</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>MUSIC (new Q: TA, F)</td>
</tr>
<tr>
<td>FLMST</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>LIB</td>
</tr>
<tr>
<td>ART</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>GRAD DIV</td>
</tr>
<tr>
<td>FHSP</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>SP&amp;PT (new Q: F)</td>
</tr>
<tr>
<td>FLMST</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>SPCH</td>
</tr>
<tr>
<td>ART</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>TEP</td>
</tr>
<tr>
<td>FHSP</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>ARTHI</td>
<td>THTR</td>
</tr>
</tbody>
</table>

Another notable item in Table 2 is the number of departments that created new ESCI questionnaires in 2018. The ESCI Group at Instructional Development has been revamping the ESCI item pool for the past few years to make a set of items that are more specific and
meaningful to instructors, TAs, and departments. We have been using the updated item pool to help departments create new questionnaires as they transition to Online ESCI.

The remainder of this report provides summary data and brief explanations regarding student responses to survey items that have remained in use since 2012 and/or 2014. Some of the findings presented in this report show data only as from 2014 or 2015 because the populations and proportions of comments from 2012-2014 was so variable compared to later years that it skewed the results, even when normalized. Results from 2018 are also sometimes disproportionate, but are more explainable through comparisons with historical data and the influx of STEM respondents.

Both quantitative and qualitative data analysis were used to analyze the responses to scalar questions against written comments from the students. All text excerpts were categorized and coded, and then compared to related scalar responses to provide a more holistic picture of student perceptions, ideas and suggestions. Some recommendations are provided in this report, and are based on discussions with the Instructional Development ESCI Group regarding student responses and technical possibilities of the ESCI Online system.
Q1: The online end of quarter survey is an improvement over the traditional paper-based end of quarter survey.

Figure 1 indicates that 78% of respondents overall agreed or strongly agreed with this statement, while, notably, only 4% of respondents disagreed or strongly disagreed. Figure 2 represents the percentage of responses to Q1 every year of the survey and shows that increases in agreement noticeably occurred in 2013 and in 2018 (by 14% and 11% respectively).

We conducted analyses comparing comments coded as “Positive feedback from students” with yearly responses to Q1, but no relationships were evident. Thus, the 2018 uptick in strongly agree responses might be attributed to the large influx of STEM students taking the survey.
Q2: I was able to provide more thoughtful feedback in this online format of the ESCI survey as compared to the paper-based survey.

Figure 3 indicates that 55% of respondents strongly agreed or agreed that the online ESCI survey allowed them to provide more thoughtful feedback. 37% of respondents answered neither agree nor disagree to this statement, while only 8% disagreed or strongly disagreed.

Figure 3 – More thoughtful feedback: combined responses

Figure 4 illustrates the pattern of responses to Q2 from 2012 to 2018. Note that the response pattern is especially consistent from 2015-2017, but in 2018 the percentage of strongly agree responses surpasses both agree and neither. Two codes for student comments that we analyzed as possible influences for the 2018 uptick in strongly agree responses are: 1) percentage of positive comments about the flexibility of taking the surveys anytime/anywhere, such as no limits on how long a student can spend responding to a single survey, and 2) a decrease in student pleas to write better questions that might indicate an improvement in question wording within the surveys themselves (since many of the new departments had written new ESCI questionnaires).

Figure 4 – More thoughtful feedback: comparison of response percentage per year
In regards to flexibility, analysis shows that all but one of the 43 excerpts coded for flexibility were from students who answered strongly agree, agree, or neither to Q2. This indicates a positive relationship between having flexibility and giving thoughtful responses. However, only 5 of those excerpts were from 2018 compared to 12 from 2017, suggesting that flexibility of time in taking the survey did not influence the 2018 uptick.

In regards to student pleas to write better questions, we tested to see if the 2018 uptick in positive responses to Q2 were accompanied by a decrease in requests to write better survey questions. However, a comparison of scalar responses to Q2 from respondents who wrote a comment coded as write better questions, reveals that there was a 35% increase in the number of these comments.

Figure 5 shows that the percentage of respondents who wrote a comment suggesting that better questions be written rose especially sharply in 2015 and again in 2018 for those who disagreed and strongly disagreed with Q2. In 2018 even students who agreed with Q2 thought that better written questions would improve the Online ESCIs.

Figure 5 – “Write better questions” aligned with responses to Q2 (percentages)

Figure 6 illustrates what kinds of questions respondents suggested would improve the questionnaires. Note that in 2015 and 2018 respondents wrote that “More specific” questions would be helpful, while in 2015 respondents asked for “More questions” as well. The 2015 codes reflect directly on the ESCI questionnaire for Psychology faculty, which asks only Items A&B. The 2018 spike in “More specific” may be a reflection on the ESCI questionnaire for Math faculty, which only asks Items A&B and provides an open-ended question for general comments (Math makes up almost 40% of all STEM respondents).
Thus, it is unclear if there was any particular influential factor that could account for the 2018 uptick in strongly agree responses to Q2 beyond a dearth of comments in general in 2018 and massive increase in survey respondents from STEM departments who rewrote their questionnaires to make them more specific and meaningful.
Q3: I am confident that the online ESCI survey maintains my anonymity.

Figure 7 shows the breakdown of all responses to Q3 between 2012-2018: I am confident that the online ESCI survey maintains my anonymity. Overall, 72% of the respondents strongly agreed or agreed with the statement. Figure 8 indicates that the response percentages have been maintained since 2015. Notably, in 2018 the percentage of respondents who answered “strongly agree” surpassed the “agree” percentage for the first time.

To investigate the 2018 uptick in “Strongly agree” responses, we analyzed all written responses that were coded as “Anonymity Concerns” (n=189). The most frequent text responses to Q7, “Do you have any suggestions or concerns about the online ESCI survey?” are coded as “Anonymity Concerns” (n=166, see Figure 17 later in this report). Closer analysis reveals that when all excerpts coded as “Anonymity Concerns” are normalized across years, respondents in 2015-2016 wrote fewer comments about this than in the years preceding and following them, with the highest rate of concern in 2018 (see Figure 9). It is possible that the 2013 rise in...
concerns about anonymity reflects the one-time Statistics student population, while the major drop in 2015 could be due to the first group of Psychology students. The gradual rise in concerns might reflect cultural emphases on privacy and anonymity over the past few years. We will address it directly in future communications to students. Thus, the uptick in “Strongly Agree” responses is not due to a decrease in written concerns about anonymity.

Further investigation of “Anonymity Concerns” excerpts reveals interesting patterns amongst excerpts in which students ask for better information about ESCI and express “Anonymity Concerns”. When these excerpts are combined together and organized according to students’ responses to Q3 (Figure 11), it indicates that students who disagreed or strongly disagreed wrote very few comments about why they disagreed, nor offered suggestions for guaranteeing their anonymity. By contrast, students who strongly agreed, agreed, or were neutral indicated that “Providing better information to students” about ESCI Online, “Better written email reminders,” and telling them how ESCI information is used might assuage concerns about anonymity. Although interesting, and definitely actionable (see analysis of Q6), this information does not explain the 2018 uptick in strongly agree responses to Q3. Again, it may simply be from the new STEM population.
Q4: Would you have a wireless device (laptop, smartphone, tablet) with you in class, that you could use to complete the ESCI survey in class?

Figure 1 shows a steady increase in mobile device ownership that surpassed 90% in 2017 and 2018. This is only slightly lower than device ownership statistics amongst American college-aged students, as reported by the PEW Research Center. UCSB’s CIO reports from 2014-2017 indicate that each person with a UCSB wireless account averaged between 2.6 wifi-enabled devices. UCSB Enterprise Technology Services has improved wifi access during this same period.

Since device ownership and access to wifi were stable or rising, we were curious to see if there was a decrease in comments with concerns about “Access to mobile devices” and/or an increase in comments coded as “Do ESCI Online during class”. Concerns about “Access to mobile device” has decreased to almost nothing, while “Do during class” increased 90% between 2013-2015. It hovered at that level through 2017, then decreased 40% in 2018. There were no related patterns in other codes or questions that could explain why it dropped.

---

Q5: What was the primary reason that made you decide to complete the online ESCI survey(s) for your course(s) this quarter?

Question 5 was added to the Annual Winter Survey of Online ESCI in 2014 to track the reasons why students chose to complete the Online ESCI for their course(s). Figure 13 shows the breakdown of all responses over the years, indicating that email has been the primary reason. A yearly comparison of student responses to Q5 (Figure 14) shows that GauchoSpace became more common in 2017, then just surpassed email in 2018. An increasing number of instructors are using GauchoSpace every year, so we suspect that students find the link to Online ESCI quite naturally that way. This may also account for the rise to 51% response rates in 2017-2018.

No particular patterns were apparent between these responses and comments coded as “Send automatic email or GauchoSpace reminders” or “Instructor reminders.” However, a recommendation emerges when considering this data with responses and analysis of Q4: add a new option to Q5 that asks if students completed their Online ESCIs during class time that was expressly set apart for it by the instructor.
Q6: What ideas can you offer for improving response rates as the campus transitions to Online ESCI?

Figure 15 shows the most frequent responses (normalized) to Q6 between 2014-2018. It is obvious that students overwhelming think that incentives will improve response rates, either via raffles or by making it mandatory to complete ESCI surveys online. “Raffle or Giveaways” may be possible campus-wide after ESCI Online is offered by all departments, and also after much careful consideration by administrative bodies. Students’ ideas for enforcing mandatory participation ranged from adding an extra credit point to not releasing students’ grades until all online surveys were complete (which is technically not possible to do on GOLD).

The Online ESCI system currently tracks student completion of all surveys, but no reporting mechanism has been created that would/could be delivered to instructors who wish to give points to individual students for taking the survey. We recommend exploring this idea, with careful consideration of campus culture, anonymity concerns, and technical possibilities. On the other hand, instructors who wanted to incentivize ESCI participation for all their students could do so using the overall class response rate that is reported live on ESCI Online.

Figure 15 – Most frequent ideas for improving response rates 2014-2018 (normalized)

Another theme that is evident in Figure 16 is “better communication to students” through reminders from email, GauchoSpace, and instructors. Once this data came to light in Fall 2018, Instructional Development began creating new email reminders for students to provide information about student anonymity and information about how ESCI information is used. However, there is still work to be done on this including continued editing of reminder emails
for conciseness, providing ideas to instructors about how to encourage student participation, creating better wording for the ESCI landing and login pages, and creating a website for instructors and students with more details.

Since we addressed “Write better questions” in Q2, the last code to discuss here is “Do during class.” In the analysis of Q4 we saw that overall comments coded as “Do during class” decreased by about 40% in 2018. Upon further analysis for Q6 it is evident that this code, when applied to excerpts that are responses to Q6 only, has been decreasing quite dramatically in student responses since 2015 (see Figure 16). Again, the 2018 drop might be due to the influx of STEM departments into the Pilot of Online ESCI, but another explanation could be that students have become more habituated to taking online surveys at all times of the day since polling technologies have become increasingly more common. It is, therefore, possible that the idea of sitting in class to take an online survey is not something current students perceive as a valuable use of class time.

*Figure 16 – All excerpts for "Do during class" organized by year and Q6 or Q7 (normalized)*

![Graph showing the trend of excerpts for "Do during class" organized by year and question, normalized, n=264](image-url)
Q7: Do you have any suggestions or concerns about the online ESCI survey?

It is encouraging to note that “Positive feedback from students” is ranked third on the list of codes applied to excerpts from Q7 (Figure 17). The vast majority of this feedback essentially says, “This is great!” or “Good job – keep it up.” Approximately 25% of the “Positive feedback” excerpts also laud the fact that Online ESCI is more environmentally friendly.

Figure 17 – Most frequently mentioned suggestions and concerns, 2014-2018 (normalized)

Figure 17 also shows that the most frequent suggestions and concerns about Online ESCI between 2014-2018 are “Anonymity concerns” and “Write better questions”. This echoes the data analysis reported above for Q2 and Q3, therefore we will not address them further here.

It is important to note that the codes for making Online ESCI “mandatory”, “Write better questions,” “Better communication to students” and “Do during class” are in the top tier of comments in both Q6 and Q7. Therefore, we recommend that the committee seriously consider these in upcoming discussions and decisions. The information about “Write better questions” and “Better communication with students” may also be interesting to the Academic Senate Committee on Course Evaluations.

On another note, students also gave feedback about technical aspects of Online ESCI that the Instructional Development ESCI Group has begun to consider for future updates to the ESCI system. Their feedback includes requests to update the color/design so it is easier to read on mobile devices and making the class/survey list less confusing.
Summary

Students who responded to the Annual Winter Survey about Online ESCI overwhelming find it an improvement over paper ESCI. There was a noticeable increase in positive responses to all of the survey questions in 2018 when Online ESCI was delivered to approximately 15,000 students from STEM departments for the first time.

Notable findings from this report include:

• There is a positive relationship between the flexibility of completing the Online ESCIs anytime/anywhere and students’ agreement that they can provide more thoughtful feedback to their instructors in Online ESCIs than in paper ESCIs.

• In 2018 students who strongly disagreed that Online ESCI allowed them to provide more thoughtful feedback and wrote comments about it requested that ESCI questions become “more specific.” This finding might be of interest to the Academic Senate Committee on Course Evaluations, and provides some guidance for the Office of Instructional Consultation as they work with departments to revamp their ESCI questionnaires.

• Students’ strong confidence in Online ESCI’s ability to keep their responses anonymous has remained at about 72% since 2015, but the number of comments expressing concern about anonymity has more than doubled since then. Students suggest that better communication about the ESCI process and uses of their responses might assuage their concerns. The ESCI Group at Instructional Development is already working on this.

• Over 90% of UCSB students have a mobile device that they could use in class to complete their Online ESCIs. Wifi access across campus is becoming more robust every year.

• Interest in completing Online ESCIs during class time was high from 2015-2017, but dropped by about 40% in 2018. This might be due to the influx of STEM students, many of which would be completing ESCIs for their TAs during hectic lab time. It could also be due to new cultural norms in which students are so habituated to taking surveys on their phones while they are out and about, that they do not perceive taking ESCIs online during class time as an effective use of that time.

• The most frequent comments made by students between 2015-2018 are (Figure 18):
  o Make Online ESCI “grade-related or mandatory”
  o Do Online ESCIs during class
  o Remind students to take the Online ESCIs through email, GauchoSpace and instructor reminders
  o Write better questions
  o Concerns about anonymity
  o Positive feedback about Online ESCI
  o Requests for better communication about Online ESCI
  o Provide incentives to complete ESCIs via raffles or giveaways
The ESCI Group at Instructional Development will work on the technical requests and communication aspects reported in this document. We will use these findings and other research/literature about online course evaluations to make appropriate alterations to the Annual Winter Survey questions and continue to work with the Ad-hoc Online ESCI committee regarding the group’s decisions and recommendations.

The ESCI Group plans to distribute the Annual Winter Survey again in 2019, with some minor alterations as recommended here and in the updated literature review. We expect a larger pool of students to take the survey because Anthropology and Statistics both (re)joined the Pilot Online ESCI in Fall 2018 and a few other departments are waiting for final approval to join.
APPENDIX C

Updated Analysis of Response Rates and Ratings
Appendix C
Updated Analysis of Response Rates and Ratings

Response Rates

We conducted a new comparative analysis of response rates between paper and online surveys using data from Winter 2015 through Fall 2018. For all paper-based faculty surveys conducted between 2015 and 2018 there is a mean response rate of 73.82% (SD = 23.00%, Median: 79%, N=12,926 course surveys. For the full set of ESCI Online surveys, the average response rate was 54.73% (SD=21.73%, Median: 51.50%, N=2452 course surveys). These statistics for both classes of surveys are consistent with, and almost identical to, the previous analysis, which was based on paper survey data from Winter 2000 through Fall 2014. Figure 1 provides a comparison box plot of response rates for the ESCI Online Pilot compared to the same years of paper data. For paper surveys, note that there are a large number of outliers below the first quartile. Note also, that for both sets of data analyzed, the data were confined to faculty surveys only as those are the only surveys that contain the two consistent questions, Items A and B. This analysis is based on responses to Item A. There is no way to do a similar analysis across campus for TA surveys because there are no campus-wide questions for TAs. The take away here, which is consistent with the literature on online course surveys, is that response rates for online surveys tend to be lower than for paper surveys, but with a very high degree of variability.

In an effort to assess the degree of variability represented in survey results from each instrument (online and paper), we made a series of box plots, first segregating by year, then by quarter, and finally by department. Figure 2 present the analysis of both online and paper surveys by year from Winter 2015 through Fall 2018. Of particular note here is that for 2017 and 2018, while the mean and median response rates for online surveys still were lower than those for paper surveys in the same period, there was a noticeable increase in response rates from the previous two years. The median response rates in both those years for online surveys exceeded the third quartile bound on the same class of surveys from the previous two years. In addition, the distribution of online response rates overlapped the response rate distributions for paper surveys.

Additional analysis of response rates for both classes surveys indicates that while the mean and median response rates annually on paper for the whole campus are quite stable, there is considerable variability from quarter to quarter, with Spring quarter consistently having lower overall response rates. That pattern is illustrated in Figure 3 below. Note that the quarter number designation in the X axis legend maps to 1 = Winter, 2 = Spring, 4 = Fall. Given the much larger number of surveys represented on paper, the pattern is quite evident, and is consistent with our previous analysis. The overall smaller number of online surveys, especially for the quarter prior to Fall 2017, before we expanded the pilot, result in more variability. However, the quarterly variability is definitely present when looking at the box plots from Fall 2017 through Fall 2018. An additional revealing pattern in this analysis by quarter shows the steep drop in response rates in Fall 2017 for paper surveys. This was due to the disruptions caused by the Thomas Fire. Note that response rates for the online surveys for Fall 2017 were substantially higher, and with a median response rate that was higher than the median for paper surveys. We attribute this to the fact that students were still able to complete the online course evaluations regardless whether they, or their faculty members, could be on campus for the end of the quarter. The online system also allowed us to adapt to the extraordinary conditions caused by both the Thomas Fire
and the tragic mudslides afterward, by allowing us to keep the survey system open to collect responses up until the rescheduled finals schedule triggered closing the system.

A much higher degree of variability in both paper and online response rate data is revealed when the data are analyzed by department over time. A box plot of that analysis is presented in Figures 4 and 5 below. When viewed in this way it becomes clear that **there are some departments that have historically maintained very high response rates, and others that have had consistently low response rates using paper surveys.** In both cases there are examples of departments whose entire response rate distribution, or a significant portion of it, falls below the campus mean. We suspect that this disparity between departments, for both survey methods, reflects differing departmental emphases on the importance of the course evaluations, and likely also reflects widely differing faculty engagement with the process, regardless of the survey instrument. As the literature review, as well as long experience with course surveys here at UCSB show, an important predictor of student engagement with the evaluation process is faculty engagement with the process. This is particularly true when the faculty let students know how important this form of feedback is for further improving courses, and the value that the faculty place on that feedback. This pattern of variability by department is also consistent with the same measure done for all paper surveys from 2000 through 2014.

The final response rate analysis that was conducted compares online versus paper response rates for departments that joined the pilot during the 2015-2018 analysis window. These results are reported in Figure 6. Two things to note about this illustration. First, this does not represent all departments who have participated in the ESCI Online pilot. This is due to the fact that some of the departments in the pilot started well before Winter 2015, and so there is no comparable paper survey response data available for those departments and they were left out of the analysis. Second, a number of the departments represented in the figure only show a single horizontal bar for the aggregated response rate data for online. That is an artifact of the data selection method, and indicates departments that had only one or two fully online courses offered during the window, and since they were online courses, had to have online course surveys. As a result they are included in the analysis, but the number of surveys in this period would amount to at most four online courses/surveys, while all the rest of their regular course surveys were conducted on paper. It is encouraging to note that for those departments that do have full box distributions of both online and paper response rates, most of their distributions overlap quite well, even though in general the overall pattern is that of lower response rate online. There is also one exception, which is a department that actually has substantially better response rates online than they did on paper during the sample time window.

In short, our analysis indicates that in general response rates for course surveys online are lower than course surveys conducted on paper. This is consistent with the literature and other institutions’ experience. However, the data show that **online response rates can be as good, and occasionally better, that response rates on paper, and most of the variability in response rates tracks by department, and possible differences in faculty engagement with the process. Both of those factors can be effectively addressed to help increase response rates, regardless the survey instrument being deployed.**
Fig. 1. Comparison of Overall Response Rates
Fig. 2. Comparison of Response Rates By Year
Fig. 3. Comparison of Response Rates By Year and Quarter
Figure 4. Variability in Paper ESCI Response Rates by Department, 2015-2018
Figure 5. Variability in Online ESCI Response Rates by Department, 2015-2018
Figure 6. Variability in Online and Paper ESCI Response Rates by Matched Department
Analyses of ESCI Online Pilot Ratings

Purpose of Analyses

The overarching question guiding the analyses in this report is how the type of survey instrument, ESCI Online Pilot or paper-based ESCI, affects students’ evaluations of UCSB instructors and courses. Additionally, we investigated the effect of course size on student evaluations when shifting from paper-based ESCI to the ESCI Online Pilot.

Summary of Findings

The metrics used to measure the impact of switching from paper-based ESCI to the ESCI Online Pilot revealed that the changes in student evaluations of teaching associated with the different test instruments are consistently small. There is no clear pattern associated with the observed shifts in scores. While some faculty have hypothesized that conducting student evaluations online, rather than on paper, will produce bimodal distributions (where only students who either loved or hated the class will respond), our analyses do not show a bimodal response pattern. In investigating how ratings shift with the transition from paper to online evaluations by class size, there is again no clear pattern in terms of the response categories, but there is clear evidence that small classes have more variability associated with the student responses. This suggests that higher response rates are particularly important for small classes, in order to have confidence that the ESCI results accurately represent students’ experience. The decline in response rates for small classes from paper-based ESCI (76%) to ESCI Online (55%) deserves careful consideration.

Data Included in Analysis

This analysis includes ESCI data from Winter 2015-Winter 2018.

Student ratings for the two campus-wide survey questions, ESCI Item A, “Please rate the overall quality of the instructor’s teaching” and ESCI Item B, “Please rate the overall quality of the course, including its material or content, independent of the instructor’s teaching” are analyzed. In order to make the most direct comparisons between the two survey instruments, courses included in the analysis were restricted to: undergraduate level only, taught by Faculty rank instructor, with 10 or more students. Additionally, to understand how student ratings change when moving from paper-based ESCI to online evaluations, the data were restricted to courses that were evaluated for the same instructor/course combination. There were 174 pairs of data, e.g. the online vs. paper evaluation of “Instructor Jones” teaching course “X,” from ten departments, as summarized in Table 1.
Table 1. Courses Included in Analyses

<table>
<thead>
<tr>
<th>Department</th>
<th>Number of Courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARTHI</td>
<td>1</td>
</tr>
<tr>
<td>CHEM</td>
<td>19</td>
</tr>
<tr>
<td>CNCSP</td>
<td>4</td>
</tr>
<tr>
<td>GEOG</td>
<td>10</td>
</tr>
<tr>
<td>LAIS</td>
<td>1</td>
</tr>
<tr>
<td>LING</td>
<td>20</td>
</tr>
<tr>
<td>MATH</td>
<td>39</td>
</tr>
<tr>
<td>MUSIC</td>
<td>27</td>
</tr>
<tr>
<td>PSY</td>
<td>52</td>
</tr>
<tr>
<td>TEP</td>
<td>1</td>
</tr>
<tr>
<td><strong>Grand Total</strong></td>
<td><strong>174</strong></td>
</tr>
</tbody>
</table>

Only one offering by a unique instructor/course pair were included, to avoid any particular instructor (who may teach the same course many times) or course from influencing the data. The single instance of a unique instructor/course pair was selected based on the first time that the instructor/course was evaluated using ESCI Online, paired with the most recent time that instructor/course was evaluated on paper.

Guidelines for Interpretation of ESCI Scores

Our interpretation of the analysis of the changes in distributions is informed by the same advice on interpreting ESCI data that Instructional Development has provided to faculty for many years. The full explanation is available online at: [http://oic.id.ucsb.edu/interpreting-data](http://oic.id.ucsb.edu/interpreting-data). In particular, **the guidance on interpreting the percentage differences in ratings between the current course and the department norms is helpful in this context.** In general, we suggest that by themselves these results (Items A and B) are accurate enough only to place instructors and courses into three broad categories: the truly outstanding, those with problems, and the vast majority in the middle.

Results are always reported in terms of percentage distributions, so it is easy to look at the results and see that, for example, "46% of students rated the teaching as Excellent or Very Good." Looking at the entire distribution allows for determining if responses are clustered in a couple of categories ("They loved it!") or spread across the categories ("Some loved it, some thought it was OK, and some hated it"), or perhaps even bimodal ("Some loved it and some hated it, but nobody was indifferent").

In using the comparison norms, individual results will differ to some extent from those of the department during the current quarter, the department over time, and the campus over time. It is important to understand when those differences are educationally meaningful and when they are not. It is possible to use statistical procedures to compare two distributions, but it is not uncommon for the number of student respondents to be so large that any difference is statistically significant, even though it may not be educationally meaningful. As a rough rule of thumb to decide when the differences between individual results and the "norms" may be meaningful, Instructional Development suggests the following:

- If the percentage of students who rate your course in any particular category differs by about **10% from the percentage in a norm group**, then there MAY be something educationally meaningful going on, and it's worth examining.
For example, suppose your department this quarter has 54% of students rating "the overall quality of the instructor's teaching" as "Excellent," and 64% of your students rate your teaching as "Excellent." Then you may be doing a better job in this course than the departmental average, and you could look at responses to other questions on the survey and to other sources to understand the differences, and to decide for yourself whether they are meaningful.

• If for any response category the **percentage differs by about 20%**, then there's **probably something meaningful going on**, and it's definitely worth seeking further understanding.

• If the **percentage differs by about 30%**, then there **IS something meaningful going on**.

**Data Metrics**

We attempt to quantify the change in student evaluations of campus-wide Items A and B from paper-based ESCI to ESCI Online using two metrics: Deviation of Scores from Departmental Norms and Meaningful Deviations from Departmental Norms.

Where appropriate, we have used paired T-tests to test whether the differences in metrics for paper-based ESCI to ESCI Online are statistically significant. As with any statistical test, the p-value for the T-tests measure the probability that the paired data comes from the same distribution (in laymen's terms, that scores for the two evaluation instruments are the same). In cases where the p-value is less than the significance level, we reject the null hypothesis that the scores are from the same distribution and conclude that there is statistically significant difference between scores from paper-based ESCI and ESCI Online. Importantly, this approach to hypothesis testing assumes that the only difference between the pairs of data is the type of evaluation instrument (paper-based ESCI or ESCI Online); the differences between metrics for the paired data is attributed solely to the type of evaluation instrument. In reality, differences in the paired data could be due to other factors, including, for example: time of day course is taught, differences in student evaluators, teaching approach changed from year to year, differences in student preparation for course, etc. Thus, it is useful to keep in mind that in cases where the results are "statistically significant," the differences between the paper-based ESCI metric and the ESCI Online metric may be due, in part, to factors beyond the test instruments themselves.

**Deviation of Raw Scores from Departmental Norms**

**Deviation of Rating Distributions from Departmental Norms**

As is indicated in the [Guidelines for Interpretation of ESCI Scores](#), the most meaningful feedback in ESCI results comes from comparing the distributions of ratings for the current course to the department norms so see if there are important deviations either above or below the norms. To investigate the pattern in deviations from department norms between online and paper surveys, we conducted paired T-test for each response category, using the formula, "% Deviation from Departmental Norms, ESCI Online Pilot - % Deviation from Departmental Norms, paper-based ESCI.” The results for Item A are reported in Table 1, below. When using this metric, the changes in response ratings from paper-based ESCI to ESCI Online are small.

Table 1 reveals that there is **no significant difference** in the mean deviations from department norms at the .05 significance level for the “Excellent”, Very Good”, “Good” and “Fair,” ratings for Item A. There is a statistically significant difference in the “Poor” rating categories, with a mean percent change of -1%, indicating that instructors’ scores, compared to the departmental norm, decreased by an average of 1%.
in the “Poor” category, when evaluations were conducted Online instead of paper. While statistically significant, the percentage change of 1% is small.

Table 1. Analysis of Changes in Deviations of Distributions from Departmental Norms, Item A

<table>
<thead>
<tr>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Change</td>
<td>0%</td>
<td>1%</td>
<td>0%</td>
<td>0%</td>
<td>-1%</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>20%</td>
<td>16%</td>
<td>11%</td>
<td>8%</td>
<td>7%</td>
</tr>
<tr>
<td>N</td>
<td>174</td>
<td>174</td>
<td>174</td>
<td>174</td>
<td>174</td>
</tr>
<tr>
<td>T Score</td>
<td>0.204</td>
<td>0.629</td>
<td>0.056</td>
<td>0.028</td>
<td>2.446</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.8392</td>
<td>0.5315</td>
<td>0.9552</td>
<td>0.9777</td>
<td>0.0157</td>
</tr>
</tbody>
</table>

The results for Item B are reported in Table 2, below. These results indicate no significant difference in the mean deviations from department norms at the .05 significance level for the “Excellent”, Very Good”, and “Fair,” ratings for Item B. There are statistically significant differences in the “Good” and “Poor” rating categories, with an average increase of 2% in the “Good” category and average decrease of -1% in the “Poor” category when the Online evaluation method is used instead of the paper-based ESCI evaluation.

Table 2. Analysis of Changes in Deviation of Distributions from Department Norms, Item B

<table>
<thead>
<tr>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Change</td>
<td>-1%</td>
<td>1%</td>
<td>2%</td>
<td>0%</td>
<td>-1%</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>19%</td>
<td>16%</td>
<td>10%</td>
<td>9%</td>
<td>4%</td>
</tr>
<tr>
<td>N</td>
<td>174</td>
<td>174</td>
<td>174</td>
<td>174</td>
<td>174</td>
</tr>
<tr>
<td>T Score</td>
<td>0.697</td>
<td>0.447</td>
<td>2.775</td>
<td>0.500</td>
<td>2.344</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.4880</td>
<td>0.6566</td>
<td>0.0063</td>
<td>0.6184</td>
<td>0.0206</td>
</tr>
</tbody>
</table>

Deviation of Mean from Department Norms

An analysis of the deviation of individual rating means from department mean rating was also conducted. These results are presented for Items A and B in Table 3, below.
Table 3. Analysis of Changes in Deviation of Means from Dept. Norms for Items A and B

<table>
<thead>
<tr>
<th>Measure</th>
<th>Item A</th>
<th>Item B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Difference in Mean Scores</td>
<td>0.02</td>
<td>-0.01</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.50</td>
<td>0.39</td>
</tr>
<tr>
<td>N</td>
<td>174</td>
<td>174</td>
</tr>
<tr>
<td>T Score</td>
<td>0.640</td>
<td>0.251</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.5242</td>
<td>0.8027</td>
</tr>
</tbody>
</table>

Again using the paired T-test, there is no significant difference in the mean deviations of the means between paper and online surveys from the department mean for Item A or Item B.

Changes in “Meaningful” Deviations from Department Norms

Next, we examined how the number of “Meaningful” differences between instructor/course scores and their associated departmental norms changed from paper-based ESCI to ESCI Online. Specifically, based on the Guidelines for Interpretation of ESCI Scores, we defined a “Meaningful” difference between an instructor’s score and the departmental norm as greater than, or less than, a 20% difference. So, if an instructor had 70% of students rate Item A as Excellent, and the corresponding departmental norm for Item A, Excellent, was 50%, that difference is counted as “Meaningful,” whereas an instructor who had 40% of student responses fall in the Excellent category, and the same departmental norm of 50%, would not be counted as a “Meaningful” difference (since 40% - 50% is only -10%, and does not meet the threshold difference of >20% or < -20%).

Table 4 summarizes the percentage of instructor/course combinations that had “Meaningful” differences from the departmental norms for Item A using ESCI Online, while Table 5 summarizes the percent that had “Meaningful” differences for Item A using paper-based ESCI. Table 6 calculates the “Change in Meaningful Differences” by subtracting each cell value in Table 4 from Table 5.

The cells in Table 6 can be interpreted based on the following example: In Table 4, 12% of instructors received student ratings for Item A that were “Meaningfully lower” than the departmental norms for the Excellent category, when evaluations were conducted using ESCI Online. In Table 5, 17% of instructors received student ratings for Item A that were “Meaningfully lower” than the departmental norms for the Excellent category, when evaluations were conducted using paper-based ESCI. Therefore, there were 5% fewer instances of instructors receiving student ratings that were “Meaningfully lower” than the departmental norms in the Excellent response category when shifting from paper-based ESCI to ESCI Online. Note: certain cells in Table 6 may appear to off by 1-2% due to rounding.
**Table 4. Meaningful Differences, Item A: (ESCI Online Scores - Norms)**

<table>
<thead>
<tr>
<th></th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Meaningfully Higher (&gt;=20%)</td>
<td>22%</td>
<td>4%</td>
<td>2%</td>
<td>3%</td>
<td>2%</td>
</tr>
<tr>
<td>% No Meaningful Difference</td>
<td>66%</td>
<td>92%</td>
<td>93%</td>
<td>97%</td>
<td>98%</td>
</tr>
<tr>
<td>% Meaningfully Lower (&lt;=-20%)</td>
<td>12%</td>
<td>4%</td>
<td>5%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Table 5. Meaningful Differences, Item A: (Paper-based ESCI Scores - Norms)**

<table>
<thead>
<tr>
<th></th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Meaningfully Higher (&gt;=20%)</td>
<td>24%</td>
<td>2%</td>
<td>5%</td>
<td>5%</td>
<td>2%</td>
</tr>
<tr>
<td>% No Meaningful Difference</td>
<td>59%</td>
<td>90%</td>
<td>90%</td>
<td>95%</td>
<td>98%</td>
</tr>
<tr>
<td>% Meaningfully Lower (&lt;=-20%)</td>
<td>17%</td>
<td>8%</td>
<td>5%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

**Table 6. Change in Meaningful Differences, Item A: (ESCI Online - Paper-based ESCI)**

<table>
<thead>
<tr>
<th></th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>-1%</td>
<td>2%</td>
<td>-3%</td>
<td>-2%</td>
<td>0%</td>
</tr>
<tr>
<td>Δ % No Meaningful Difference</td>
<td>6%</td>
<td>2%</td>
<td>3%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>-5%</td>
<td>-4%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Looking at each of the response categories in Table 6, we see that when moving from paper to online evaluations:

- For the Excellent category, there is a slight (-1%) decrease in “Meaningfully Higher” instructor ratings, a 6% increase in “No Meaningful Difference” and a 5% decrease in “Meaningfully Lower” ratings.
- For the Very Good category, there is a 2% increase in “Meaningfully Higher” instructor ratings, a 2% increase in “No Meaningful Difference” and a 4% decrease in “Meaningfully Lower” ratings.
- For the Good category, there is a 3% decrease in “Meaningfully Higher” instructor ratings, a 3% increase in “No Meaningful Difference” and no change in “Meaningfully Lower” ratings.
For the Fair category, there is a 2% decrease in “Meaningfully Higher” instructor ratings, a 2% increase in “No Meaningful Difference” and no change in “Meaningfully Lower” ratings.

For the Poor category, there no change in the percentage of “Meaningfully Higher” “No Meaningful Difference” and “Meaningfully Lower” ratings.

Table 7 summarizes the change in “Meaningful differences” for Item B, students’ overall evaluation of the course. Note: the item B analogs of Table 4 and Table 5 are not presented for brevity; the results in Table 7 are analogous to the results for Item A presented in Table 6.

Table 7. Change in Meaningful Differences, Item B: (ESCI Online - Paper-based ESCI)

<table>
<thead>
<tr>
<th></th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>-1%</td>
<td>-2%</td>
<td>1%</td>
<td>1%</td>
<td>1%</td>
</tr>
<tr>
<td>Δ % No Meaningful Change</td>
<td>2%</td>
<td>5%</td>
<td>4%</td>
<td>-1%</td>
<td>-1%</td>
</tr>
<tr>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>-1%</td>
<td>-3%</td>
<td>-5%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Looking at each of the response categories in Table 7, we see that when moving from paper to online evaluations:

- For the Excellent category, there is a slight (-1%) decrease in “Meaningfully Higher” instructor ratings, a 2% increase in “No Meaningful Difference” and a 1% decrease in “Meaningfully Lower” ratings.

- For the Very Good category, there is a 2% decrease in “Meaningfully Higher” instructor ratings, a 5% increase in “No Meaningful Difference” and a 3% decrease in “Meaningfully Lower” ratings.

- For the Good category, there is a 1% increase in “Meaningfully Higher” instructor ratings, a 4% increase in “No Meaningful Difference” and 5% decrease in “Meaningfully Lower” ratings.

- For the Fair category, there is a 1% increase in “Meaningfully Higher” instructor ratings, a 1% decrease in “No Meaningful Difference” and no change in “Meaningfully Lower” ratings.

- For the Poor category, there is a 1% increase in “Meaningfully Higher” instructor ratings, a 1% decrease in “No Meaningful Difference” and no change in “Meaningfully Lower” ratings.

**Effect of Class Size**

The Ad Hoc Senate Committee on ESCI Online Pilot requested that we investigate how changes in scores from paper-based ESCI to ESCI Online vary by class size. Based on natural breaks in the data, the 174 pairs of instructor/course data points were divided into three groups:

- Small classes, with 10-40 students enrolled,
- Medium classes, with 41-120 students enrolled, and
- Large classes, with 121-839 students enrolled
The response rates for ESCI Online and Paper-based ESCI are summarized in Table 8.

### Table 8. Response rates by Class Size and Evaluation Instrument

<table>
<thead>
<tr>
<th>Instrument</th>
<th>Response Rate metric</th>
<th>Small (10-40)</th>
<th>Medium (41-120)</th>
<th>Large (120-839)</th>
<th>ALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>ESCI Online Pilot</td>
<td>Mean</td>
<td>55%</td>
<td>49%</td>
<td>55%</td>
<td>53%</td>
</tr>
<tr>
<td></td>
<td>St. Dev</td>
<td>17%</td>
<td>14%</td>
<td>17%</td>
<td>16%</td>
</tr>
<tr>
<td>Previous Paper-based ESCI</td>
<td>Mean</td>
<td>76%</td>
<td>65%</td>
<td>63%</td>
<td>69%</td>
</tr>
<tr>
<td></td>
<td>St. Dev</td>
<td>20%</td>
<td>16%</td>
<td>17%</td>
<td>19%</td>
</tr>
<tr>
<td></td>
<td>Count (N)</td>
<td>65</td>
<td>58</td>
<td>51</td>
<td>174</td>
</tr>
</tbody>
</table>

We repeated the two analyses, Deviation of Scores from Departmental Norms and Meaningful Deviations from Departmental Norms, for these three groups, for Items A and B.

**Deviation of Raw Scores from Departmental Norms, by Class Size**

Together, Table 9 and Table 10 show no clear pattern in terms of the magnitude of mean changes in deviation from departmental norms for the three class size groups. Most of the changes are small and are not statistically significant. However, a striking and important pattern once again emerges from the class size grouping: **the variation in mean changes (as measured by the standard deviation) increases with decreasing class size.** For example, among Large courses, there is a mean change of -1% in the Excellent category, indicating that instructors’ scores, compared to the departmental norm, decreased by an average of 1%, when evaluations were conducted Online instead of on paper. The variation associated with this 1% decrease is +/-13% (the associated standard deviation). However, when we examine the results for Small classes, we see that there is a mean change of -2% in the Excellent category, but the variation associated with this 2% decrease is now +/- 25%. The pattern of increasing variation by class size is also evident in Table 10, where the standard deviation ranges from 4-11% for Large classes (across response categories), 4-18% for Medium sized classes, and 4-25% for Small classes.
Table 9. Analysis of Changes in Deviations of Distributions from Departmental Norms, Item A, by Class Size

<table>
<thead>
<tr>
<th>Class Size</th>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large (121-839 enrolled)</td>
<td>Mean Change</td>
<td>-1%</td>
<td>2%</td>
<td>0%</td>
<td>-1%</td>
<td>-1%</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>13%</td>
<td>10%</td>
<td>7%</td>
<td>8%</td>
<td>9%</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
</tr>
<tr>
<td></td>
<td>T Score</td>
<td>0.512</td>
<td>1.459</td>
<td>0.177</td>
<td>0.909</td>
<td>0.418</td>
</tr>
<tr>
<td></td>
<td>T-Test (P)</td>
<td>0.6144</td>
<td>0.1548</td>
<td>0.8617</td>
<td>0.3725</td>
<td>0.6805</td>
</tr>
<tr>
<td>Medium (41-120 enrolled)</td>
<td>Mean Change</td>
<td>4%</td>
<td>-2%</td>
<td>-2%</td>
<td>2%</td>
<td>-1%</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>20%</td>
<td>14%</td>
<td>13%</td>
<td>8%</td>
<td>6%</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
<td>58</td>
</tr>
<tr>
<td></td>
<td>T Score</td>
<td>1.398</td>
<td>1.246</td>
<td>1.155</td>
<td>1.689</td>
<td>1.630</td>
</tr>
<tr>
<td></td>
<td>T-Test (P)</td>
<td>0.1710</td>
<td>0.2218</td>
<td>0.2572</td>
<td>0.0995</td>
<td>0.1117</td>
</tr>
<tr>
<td>Small (10-40 enrolled)</td>
<td>Mean Change</td>
<td>-2%</td>
<td>2%</td>
<td>2%</td>
<td>-1%</td>
<td>-2%</td>
</tr>
<tr>
<td></td>
<td>Std. Dev.</td>
<td>25%</td>
<td>20%</td>
<td>11%</td>
<td>8%</td>
<td>5%</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td>T Score</td>
<td>0.569</td>
<td>0.986</td>
<td>1.424</td>
<td>0.942</td>
<td>2.849</td>
</tr>
<tr>
<td></td>
<td>T-Test (P)</td>
<td>0.5741</td>
<td>0.3315</td>
<td>0.1625</td>
<td>0.3534</td>
<td>0.0063</td>
</tr>
</tbody>
</table>
### Table 10. Analysis of Changes in Deviations of Distributions from Departmental Norms, Item B, by Class Size

<table>
<thead>
<tr>
<th>Class Size</th>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large (121-839 enrolled)</td>
<td>Mean Change</td>
<td>-1%</td>
<td>1%</td>
<td>1%</td>
<td>-1%</td>
<td>-1%</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>11%</td>
<td>10%</td>
<td>8%</td>
<td>7%</td>
<td>4%</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td>51</td>
<td></td>
</tr>
<tr>
<td>T Score</td>
<td>0.410</td>
<td>0.581</td>
<td>0.519</td>
<td>0.662</td>
<td>1.305</td>
<td></td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.6868</td>
<td>0.5677</td>
<td>0.6095</td>
<td>0.5152</td>
<td>0.2023</td>
<td></td>
</tr>
</tbody>
</table>

| Medium (41-120 enrolled) | Mean Change | 0% | 0% | 2% | 0% | -1% |
| Std. Dev. | 18% | 13% | 11% | 8% | 4% |
| N | 58 | 58 | 58 | 58 | 58 |
| T Score | 0.059 | 0.263 | 1.627 | 0.258 | 1.000 |
| T-Test (P) | 0.9539 | 0.7952 | 0.1123 | 0.7988 | 0.3256 |

| Small (10-40 enrolled) | Mean Change | -2% | 1% | 3% | -1% | -1% |
| Std. Dev. | 25% | 21% | 10% | 10% | 4% |
| N | 65 | 65 | 65 | 65 | 65 |
| T Score | 0.767 | 0.476 | 2.396 | 0.492 | 1.745 |
| T-Test (P) | 0.4491 | 0.6381 | 0.0204 | 0.6273 | 0.0881 |

Changes in “Meaningful” Deviations from Department Norms, by Class Size

The change in “Meaningful Differences” in instructor/course scores are summarized by class size, for Item A in Table 11 and for Item B in Table 12. There is no discernable difference between the three groups using this metric.
Table 11. Change in Meaningful Differences, Item A: (ESCI Online - Paper-based ESCI), by Class Size

<table>
<thead>
<tr>
<th>Class Size</th>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large (121-839 enrolled)</td>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>-4%</td>
<td>2%</td>
<td>-4%</td>
<td>-8%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % No Meaningful Difference</td>
<td>10%</td>
<td>4%</td>
<td>6%</td>
<td>8%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>-6%</td>
<td>-6%</td>
<td>-2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Medium (41-120 enrolled)</td>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>3%</td>
<td>2%</td>
<td>-5%</td>
<td>3%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % No Meaningful Difference</td>
<td>-2%</td>
<td>-2%</td>
<td>3%</td>
<td>-3%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>-2%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Small (10-40 enrolled)</td>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>-3%</td>
<td>2%</td>
<td>0%</td>
<td>-2%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % No Meaningful Difference</td>
<td>11%</td>
<td>5%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>-8%</td>
<td>-6%</td>
<td>0%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>
Table 12. Change in Meaningful Differences, Item B: (ESCI Online - Paper-based ESCI), by Class Size

<table>
<thead>
<tr>
<th>Class Size</th>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Large (121-839 enrolled)</td>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>0%</td>
<td>-2%</td>
<td>0%</td>
<td>2%</td>
<td>2%</td>
</tr>
<tr>
<td></td>
<td>Δ % No Meaningful Difference</td>
<td>2%</td>
<td>2%</td>
<td>2%</td>
<td>-2%</td>
<td>-2%</td>
</tr>
<tr>
<td></td>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>-2%</td>
<td>0%</td>
<td>-2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Medium (41-120 enrolled)</td>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>0%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % No Meaningful Difference</td>
<td>-3%</td>
<td>3%</td>
<td>5%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>3%</td>
<td>-3%</td>
<td>-7%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Small (10-40 enrolled)</td>
<td>Δ % Meaningfully Higher (&gt;=20%)</td>
<td>-2%</td>
<td>-5%</td>
<td>0%</td>
<td>2%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % No Meaningful Difference</td>
<td>6%</td>
<td>9%</td>
<td>5%</td>
<td>-2%</td>
<td>0%</td>
</tr>
<tr>
<td></td>
<td>Δ % Meaningfully Lower (&lt;=-20%)</td>
<td>-5%</td>
<td>-5%</td>
<td>-5%</td>
<td>0%</td>
<td>0%</td>
</tr>
</tbody>
</table>

Summary of Effect of Class Size

The two analyses, “Raw Scores, by Class Size” and “Deviation of Raw Scores from Departmental Norms, by Class Size,” highlight the high variability associated with changes in scores when moving from Paper-based ESCI to ESCI Online for small classes. This result is not unexpected—in small classes each student response carries more weight—it highlights the fact that we should expect ESCI results to be highly variable and less stable in smaller courses. High response rates are particularly important for small courses as illustrated by the following example: When there are 20 students in a course, each student response counts as 5% (100% divided by 20 students) of the response distribution. If, in a class with enrollment of 20, the response rate is 50%, each of the 10 students’ responses now counts as 10% of responses. Nutly (2008)1 addresses the problem of low response rates in student evaluations of teaching based on statistical analyses of sampling errors by class size. The author concludes that, for a class size of 40 (the upper end of our “Small class” group), the required responses rate is 95% (38 respondents) for a 95% confidence level, and 40% (16 respondents) for an 80% confidence level. The response rates for ESCI Online and paper-based ESCI are summarized in Table 8. For small courses, Table 8 shows that the average response rate is 55% with ESCI Online, and 76% for the previous Paper-based ESCI evaluations of the 65 small-course instructor/course pairs. While

---

neither instrument yields a response rate that satisfies the required 95% threshold associated with a 95% confidence rating, the Paper-based ESCI response rate is much closer than the ESCI Online response rates. Thus, careful consideration should be given to how the ESCI results are used in merit and promotion cases for small courses, especially when response rates are low.
APPENDIX D

Online ESCI Pilot Response to Senate Memos
Online ESCI Pilot
Response to Senate Memos

The ESCI Online Pilot Project was designed as a pilot program to evaluate the potential of shifting from a paper-based end of course evaluation instrument to an online course evaluation instrument. The sole intent was to assess feasibility, along multiple dimensions, of changing the manner in which the end of course survey questionnaires are delivered and responses collected. The pilot was monitored closely by the Ad Hoc Committee on Online Course Evaluation established by EVC Gene Lucas at the end of Fall 2010. The committee consisted of faculty representatives from three Senate Committees, representatives of Associated Students and the Graduate Student Association, as well as Instructional Development staff, and representatives from Academic Personnel, Letters and Science Information Technology, and the Office of the EVC.

The project report was transmitted to then EVC Gene Lucas on December 2, 2013. In November 2014, EVC David Marshall received and forwarded to the Executive Director of UCSB’s Instructional Development a memo, dated November 20, 2014, from the Senate Divisional Chair, Professor Kum-Kum Bhavnani. The Senate memo discussed a range of general and specific concerns related to the move to online evaluations that were discussed and reported back to Academic Senate from ten committees and councils at UCSB*. Those concerns are addressed in this response prepared by the consultants and staff of Instructional Development.

*Council on Planning and Budget (CPB), Undergraduate Council (UgC), Graduate Council (GC), Council on Research and Instructional Resources (CRIR), Council on Faculty Issues and Awards (CFIA), the Committee on Diversity and Equity (D&E), and the Faculty Executive Committees from the College of Engineering (COE FEC), College of Letters and Science (L&S FEC), the College of Creative Studies (CCS FEC), and the BREN School FEC.
Table of Contents/Categories of Concern:

Administrative Issues ........................................................................................................... 3
Who is Involved in Steering the Pilot? .................................................................................. 3
Pilot Participation .................................................................................................................. 4
Best Practices ....................................................................................................................... 6
Timing of ESCI Online Availability ..................................................................................... 9
How ESCI Should be Used by CAP ....................................................................................... 9
Literature Review ................................................................................................................. 10
Cost Savings ......................................................................................................................... 11
IT Resources ......................................................................................................................... 13
Communication .................................................................................................................. 15
Response Rates .................................................................................................................. 16
Clarification ......................................................................................................................... 22
General Issues .................................................................................................................... 28
Appendix A - Winter 2015 ESCI Survey about Online Surveys ........................................... 30
Appendix B - Annotated Bibliography ................................................................................. 34
Appendix C - Tips to ESCI Online Instructors ................................................................. 42

Key to Commenters:
BREN - Bren School of Environmental Science and Management
CLIIR - Committee on Libraries, Information and Instructional Resources
CPB - Committee on Planning and Budget
GC - Graduate Council
HB - Henning Bohn
KKB - Kum-Kum Bhavnani
Administrative Issues

Comments/Concerns:
• Groups questioned why the Pilot Project had started without support from the Academic Senate. KKB
• They also questioned under whose authority the Pilot Program was initiated. KKB
• If another pilot is instituted, we request that Instructional Development discuss these possibilities with Senate. KKB

Response:
The Ad Hoc Committee on Online Course Evaluation, which filed the report on the pilot’s progress through the beginning of Fall 2013, was a joint Senate and Administrative Committee established to oversee the pilot program. The Ad Hoc Committee had Senate representatives from CAP, CLIIR, and Undergrad Council, as well as Associated Students and the Graduate Students Association. The Ad Hoc Committee was established in the Fall of 2010 with Carol Genetti and George Michaels as Co-Chairs under the joint auspices of the EVC Lucas and then Senate Chair Bohn. When Carol Genetti became the Dean of Graduate Division, Jeffrey Stopple (one of the CAP representatives on the Ad Hoc Committee) agreed to take her place as Co-Chair. Due to the long time frame for getting the pilot up and running, the terms of most of the Senate representatives on the Ad Hoc Committee, while they continued their involvement with the Ad Hoc Committee, had expired on their respective Senate Committees. As a result, while in the initial phases of the pilot those Senate committee members were making regular progress reports to their respective Senate committees about the Pilot Program, once their tenure on their committees came to an end, so did the progress reports. The Committees did not appoint new representatives to the Ad Hoc Committee. Unfortunately neither Co-Chair Stopple nor Co-Chair Michaels recognized the need to get new Senate representatives on the Ad Hoc Committee, as the original representatives were happy to continue their participation in the Ad Hoc Committee.

As part of either extending the pilot program over an even longer period, or moving toward voluntary availability for departments to choose to move to a fully online evaluation system; we would request the establishment of a new Ad Hoc Committee on Online Course Evaluation, with representatives from the original Senate committees (CAP, CLIIR, UGC) but with a longer charter. Perhaps a charter on the order of five years would be appropriate. The Ad Hoc Committee could have a rotating Senate membership, consistent with normal terms on Senate committees, with a provision for the appointment of replacement Ad Hoc Committee members from the appropriate Senate permanent committees.

Who is Involved in Steering the Pilot?

Comments/Concerns:
• Involve faculty members from UgC, GC and CAP… lecturers. HB
• A description of the pilot program should be made available for scrutiny to other Senate committees as soon as possible. KKB

Response:
The Pilot program was steered by the Ad Hoc Committee on Online Course Evaluation, as described above. The pilot program was developed with input from three working sub-committees: a Academic Personnel Sub-Committee, a Implementation Sub-Committee, and a Content Sub-Committee. The Academic Personnel Sub-Committee developed recommendations for minimizing potential negative impacts on the academic personnel review process and...
included ladder faculty and lecturers. The Implementation Sub-Committee worked on the actual implementation details for the system, a survey of student impressions of the online system, as well as details related to student and faculty notifications, reminders, and a realtime response rate monitoring system. The Content Sub-Committee reviewed and evaluated the enormous number of existing ESCI questions, with a goal of developing a set a recommended questions that departments could review when updating their questionnaires. The committee developed a taxonomy of topics and recommended new items to be added (including many open-ended questions), and developed criteria to make existing items more suitable for online forms.

Based on the number of Senate committees responding to the Ad Hoc Committee’s report of December 2013, it appears that the report was widely circulated through the Senate during Winter and Spring quarters 2014. In addition, this response contains additional analysis of effects on response rates and on ratings based on a larger data pool than was available in December 2013.

**Pilot Participation**

**Comments/Concerns:**

- *Using the new system in a few large courses whose syllabi and instructors have remained stable for several years, in order to investigate the effect of the change on ESCI results.*  
  
  HB

- *Only tenured faculty should use an online evaluation system, and they recommend that pilot testing be restricted to courses taught by tenured faculty.* KKB

**Response:**

The Academic Personnel Sub-Committee made the following recommendations (pp. 22-23 of the Pilot Program Report, December 2013) guiding selection of departments and courses to be included in the pilot and the final program if approved, which were included in the report. Additional information on these points as they pertain specifically to the pilot follow the recommendations below.

1. **Move entire departments to the online system at one time; do not select individual courses within a department for the pilot.**

   **Rationale:** It will be very important for departments and other reviewing agencies in academic personnel cases to know precisely which courses were evaluated online. The process of record-keeping will be much simpler if all of a given instructor’s courses were evaluated the same way in a given quarter. It will also make it easier to provide a standardized evaluation of the teaching, given the consistency of the assessment tool in a given quarter.

2. **Ensure that all personnel materials clearly indicate which courses were evaluated online, and whether the course was evaluated as part of the pilot project.**

   **Rationale:** Given that we do not know to what extent or how changing the evaluation instrument will impact student evaluation of teaching, it is critical that everyone is aware of which courses were evaluated by which means until the transition is complete. It will be critical that reviewing agencies factor this into the assessment of teaching as part of academic personnel reviews. In addition, the evaluation system that is finally implemented may differ in important ways from the pilot, and this may also have an impact on the evaluation results. For these reasons, we recommend that the following
documents prominently indicate that the courses were evaluated online, and as part of the pilot program:

a. Official ESCI and narrative reports
b. Class Instruction Histories provided by Budget and Planning
c. Bio-bibliography; statement under “Teaching” header

3. **Reviewing agencies should explicitly note the method of evaluation in their letters until the transition period is complete or until campus agencies judge that it is no longer necessary.**

**Rationale:** This safeguard ensures that each reviewing agency is explicitly aware of the method by which the reviews were collected. This practice need not be cumbersome; one might for example indicate that it was reviewed online with a parenthetical comment, e.g. ‘received a score of 1.7 (online pilot evaluation)’.

4. To further mark the difference, change the name of the evaluation system from ESCI. One possibility: ESCI 2.0.

**Response:** The system has been called “ESCI Online Pilot” for the pilot period and will be called “ESCI Online” if it goes to production from pilot.

5. **Departments should consider adopting other additional means of assessing teaching during the transition period.**

**Rationale:** Since there will be a period of transition when results from the new method are being evaluated and new standards for their interpretation are emerging, faculty members might find it reassuring if additional means of assessing teaching supplement the online course evaluations. As noted in Appendix 1 of the APM, these might include: (a) opinions of other faculty members knowledgeable in the candidate’s field, particularly if based on class visitations, on attendance at public lectures or lectures before professional societies given by the candidate, or on the performance of students in courses taught by the candidate that are prerequisite to those of the informant; (b) opinions of students; (c) opinions of graduates who have achieved notable professional success since leaving the University; (d) number and caliber of students guided in research by the candidate and of those attracted to the campus by the candidate’s repute as a teacher; and (e) development of new and effective techniques of instruction.

6. **Inform instructors not to administer online course evaluations in class, even if all students have access to computers. Perhaps include a statement on this in the instructions to students, e.g. ‘Course evaluations should be not be completed during regularly scheduled instructional hours of the course being evaluated’.**

**Rationale:** Since the quality of student evaluations may vary depending on whether or not the evaluation is administered in class or online, regularizing the parameters of their collection promotes equity and equitability across courses.

7. Add a spell-check facility for students typing narrative evaluations.
Explanation: This is already implemented as it is a standard feature of all modern web browsers and required no additional programming.

In adhering to Recommendation 1 above, there were additional constraints on the determination of appropriate departments to approach about volunteering to participate in the Pilot Program. The Implementation Sub-Committee recognized that for the pilot it was not initially possible to include departments with cross-listed or concurrent enrollment courses until we had an effective way to determine correct student enrollments between cross-listed and/or concurrent courses. That eliminated a significant number of departments from consideration as potential volunteer candidates for the pilot. Those constraints have now been addressed in the ESCI Online Pilot, and both cross-listed and concurrent enrollment courses can now be handled by the system. The Academic Personnel Sub-Committee’s Recommendation number 1 was to ensure that all faculty members and TA’s in participating departments had their ESCI surveys collected using the same methodology to help ensure equity. Recommendations 2 and 3 were to ensure that no faculty member or TA would be unfairly judged on the ESCI reports collected using the online survey system.

There is one additional clarification to make here. The ESCI system has always been configured such that the normative data is based only on responses to items from surveys that were collected in the same manner and at the same time of the quarter. For example, ESCI reports for mid-quarter surveys only use responses from other mid-quarter surveys as the basis for the norms (baselines for comparison): course over time, department over time, division over time, college over time, and campus over time. In the case of the courses included in the ESCI Online Pilot, their norms are also restricted to only the other courses that were participating in the ESCI Online Pilot, and not against all of the other course end of quarter surveys that were collected using the paper Scantron forms. Thus, depending on the survey collection method and time of the quarter, there are multiple response pools used in the analysis, and the pools are never mixed. This approach will help us to identify trends in response rates and distributions that may arise with the shift to online evaluations.

Best Practices

Comment/Concern:
Research in this area also shows that when the person being evaluated is present during the survey process, the outcome is different than if they were not present. HB

Response:
As outlined on our ESCI Recommended Methods and Practices, instructors should not be present while students are evaluating the course and instruction. This recommendation is independent of the mode of survey delivery (paper or online).

Historically, the ESCI survey policy established by CETIS (former Senate Committee on Evaluation, Teaching and Instructional Support) specifically prohibits the instructor being evaluated from being in the room while the students are completing the surveys. The instructor is to hand out the questionnaires and response sheets, ask for a volunteer to collect the response sheets at the end, and return those to the department office; at which point the instructor is to leave the room. The same policy would apply if students were completing the surveys online but in class on laptops, tablets or smart phones; unless CAP or another appropriate Senate committee chose to change that long-standing policy.
Comment/Concern:
Asking students to fill out the survey on their laptops (or smartphones) might work. HB
Rephrased as: Completing Online ESCI evaluations in-class

Response:
Asking students to complete an online survey in class on either a laptop, tablet or smart phone as present may be problematic for two reasons. When asked in winter quarter 2015 only 88% of respondents (N=1,665) self-reported that they have an appropriate device that they could bring to class to engage in the online survey in the classroom (full report in Appendix A). For reasons of equal accessibility, the campus would have to develop an interim plan to provide devices for the ~12% of students who do not have a device they could bring to class for this purpose. In addition, the campus would need to make substantial additional investments in wifi access points serving the classrooms to ensure both adequate coverage in all general assignment classrooms, and a sufficient number of access points to reliably handle the volume of simultaneously connected devices. Finally, when asked, many students indicate that their preference for completing ESCI surveys online rather than in class is to avoid losing precious instructional time during dead week to the survey process.

Comment/Concern:
Students provide comments in part because time is allowed and because they see other students doing the same (argument for time to fill out in class). HB
Rephrased as: Effects of “in-class” or “at home” responses on open-ended questions.

Response:
We agree that it would be useful to investigate how comments are influenced by mode of ESCI delivery. The research literature suggests that students provide more comments when filling out surveys online, as compared to filling them out in-class, perhaps because they have more time, or perceive their typed responses as being more anonymous than handwritten feedback. It could also be interesting to compare online comments filled out on laptops/smartphones in class with online comments completed outside of class; this would shed light on the suggestion that students provide comments because they see other students doing the same.

Ultimately, assuming that the student access and campus wifi robustness issues can be effectively addressed, the determination as to whether or not to allow students to complete ESCI Online surveys in class will have to be addressed by the Senate, participating departments, and the faculty.

Comment/Concern:
CRIIR says “Decline to fill out” must be an option offered in the survey. KKB

Response:
In the pilot, currently, a student can effectively register a “Decline to fill out” response by simply submitting the survey form without having responded to any of the questions. This is equivalent to students submitting a blank response sheet in class. In the online system, they get “credit” for having completed the survey so that they stop getting pestered with reminders, but do not have to provide any feedback. It would certainly be possible to have a more explicit “Decline to Respond” button on the online survey form that would do the same thing.
Comment/Concern:
Communication with Departments and Faculty in Advance of Department Adoption

Response:
In addition to the recommendations made by the Ad Hoc Committee and mentioned in the section “Pilot Participation” above, the Committee also made to following recommendations regarding advance work and communications with faculty and departments prior department’s adopting ESCI Online (pp. 24-25 of the Pilot Program Report).

Proposed Communication Strategy For Moving Beyond the Pilot

Plentiful information should be provided to the campus about this project, especially emphasizing the safeguards that are being put into place to ensure that the process does not adversely affect the evaluation of instructors. The communication should emphasize the following points:

A. This is being done across the country, and that every UC campus is somewhere on the trajectory of shifting to online evaluations;

B. The committee is constructing the system based on research on similar processes from other campuses;

C. This process has been going for a year (Editor’s Note — almost three years now), and is being done carefully and thoughtfully, with input from relevant committees of the Academic Senate, as well as Senate faculty, Unit 18 lecturers, and graduate-student instructors;

D. We are working to determine what factors will best ensure high response rates on this campus;

E. Safeguards are being adopted to ensure that instructors are not unfairly disadvantaged by this move;

F. This change will have positive impacts on the quality of the evaluations, leading to a more useful product that can be used to improve instruction:
   1. The new system will allow students to provide more detailed and helpful comments, thus allowing instructors better feedback that can be used to improve teaching;
   2. Ease of collecting evaluations of TA’s in courses without labs or discussion sessions;
   3. Eventually, faculty members will be able to access evaluations at all times, and there will be a search capability.
   4. Eventually, faculty members will be able to see data in multiple formats, charts, etc.
   5. It will be easier to read narrative evaluations when they are typewritten.
   6. There will be a tremendous cost savings, especially in staff time, but also in paper—the new initiative is in line with campus sustainability goals.

The communications should be conveyed through the following means:

A. Informational sessions at college-level or divisional briefings for Chairs;

B. Visits of committee members to department meetings before a given department goes online;

C. A memo to the campus from the EVC;

D. Visits of committee members to meetings of the Associated Students and Graduate Student Association;
E. A website with clearly and easily accessible information (FAQ’s);
F. Articles in 93106 and the Daily Nexus, with interviews of students and instructors who have been through the process.

Timing of ESCI Online Availability

Comments/Concerns:
• The suggested time period of starting at week six for administering the surveys seems much too early to every group. …a four week window in which to evaluate a course creates inherent bias for a particular course; an evaluation done in week six would not be comparable to an evaluation done in week ten. HB
• Most groups did not feel that a four week window for completing ESCI forms was appropriate. KKB
• Timing needs to be flexible and able to be controlled on course-by-course basis. BREN

Response:
At no time did the pilot ever allow a four week window for the online surveys. In the first two quarters of the pilot we allowed a three week window for students to complete the surveys. The engagement rate in the first week was always very low, so in the next four quarters we shortened the survey window to the last two weeks of the quarter; starting at midnight on the Friday night two weeks before Finals started. That had no substantial negative effect on response rates. For the past two quarters, we started the survey window on the Monday before Dead Week; forgoing the weekend before. That also had no negative effect on response rates. It may be possible to confine the survey window just to one week, during Dead Week, and still get a reasonable response rate.

Timing the availability of the online survey window open and close dates is a very tricky issue. First, we do not have reliable data on which faculty members are teaching which courses until the fourth week of the quarter. That data is used to then have department staff indicate of those courses, which ones should be surveyed, and which survey to use for each. At the same time department staff can correct faculty assignment errors from the Registrar’s data, and have the opportunity to add survey requests for courses that may be missing. The “Survey Request” process typically takes two weeks. We have developed an online Survey Request Form that went into general use in Winter quarter 2015 that could cut that time down to one week. In week eight of the quarter we do a final extract of course and enrollment data from the Registrar to ensure that we have as accurate a list of all of the students enrolled and which courses and sections they are enrolled. For courses that are taught in the full ten week quarter, this data constraint is not an issue. However for courses, such as the Freshman Seminar courses, or Bren courses, that may only be taught over a two week period sometime during the quarter; we simply can not get all of the necessary data to allow for an online survey window prior to the end of the quarter.

How ESCI Should be Used by CAP

Comment/Concern:
At issue is therefore not the mode of delivery, but the discontinuation of ESCI as we know it, and the design of a new and intrinsically non-comparable evaluation system. HB

Response:
This entire effort has only been about replacing printed questionnaires and custom printed paper Scantron sheets with an online survey instrument as an option for those departments that
wish to take advantage of such a system if available. All other aspects of ESCI have, and will continue to, remain the same. Department staff will continue to specify which courses should be surveyed each quarter and which survey questionnaire to apply. The results will continue to be anonymous. Online survey instruments will allow for easier collection of responses to open-ended questions (qualitative responses) as part of the core ESCI system without requiring academic department staff time for transcription. Processing and reporting of ESCI results will continue to use exactly the same methods as in the past.

Comment/Concern:
“the way in which ESCI scores are used for merit is in need of careful review and reconsideration” (their italics). KKB and CRIIR

Response:
Policy regarding the use of ESCI results in academic personnel cases is solely the responsibility of the Senate, Academic Personnel and the Executive Vice Chancellor. Addressing those issue were not part of the charge to the Ad Hoc Committee for the pilot. Instructional Development provides the service to the campus to manage the collection of the survey data, process and distribute the reports, and manage the data. Instructional Development has no role in setting policy as to the use of ESCI data in personnel cases, or in any other ways. There is clearly widespread concern about the use of ESCI results for academic personnel cases, as well as some concern about the use of ESCI results in Teaching Assistant personnel decisions as well. This is an issue that would best be addressed through joint action between the Committee on Academic Personnel, Academic Personnel, the EVC’s office, Graduate Division, and perhaps Undergraduate Council and Graduate Council. Instructional Development would be more than happy to assist in that effort, but it is and should be, completely out of Instructional Development’s responsibilities.

Comment/Concern:
Would like to see ESCI scores, “even as they are currently administered, downplayed in personnel reviews for all our colleagues, faculty or graduate students. …These changes should not be entirely about labor-saving, [and] should produce a better system. We do not see much evidence of that in the proposal as it stands.” KKB and GC

Response:
The proposal is simply to provide another option to departments for managing the ESCI data collection process. There is no intention to change the way in which ESCI data are processed and reported. Because of the valuable baseline data at the department, division, college and campus levels, the ECSI system’s reports are arguably the most helpful of any of the course evaluation systems in use in the UC system.

Literature Review

Comment/Concern:
Data from other institutions in which a change from paper to online surveys resulted in drastically reduced participation rates…urge the Task Force to provide background research to faculty. Although the Report mentions literature on the subject, several believe that surveying institutions who have implemented online surveys could be very useful. HB

Response:
A literature review and background research on the topic of changes in response rates when shifting from paper to online surveys was conducted and studied by the original task force prior to submission of the first proposal. Dr. Lisa Berry, and members of the Content Sub Committee of the Ad Hoc Committee further added to that corpus. The Ad Hoc Committee was also kept
informed of developments in online course survey development UC system-wide through Dr.
Berry’s involvement in the system-wide Evaluation and Assessment Working Group, sponsored
by the Educational Technology Leadership Group (ETLG), on which she served as Chair for two
years. Included in Appendix B is the resulting annotated bibliography developed by the Sub
Committee addressing both changes in response rates and effects on ratings, the more pertinent
point.

If the attached annotated bibliography leaves remaining research questions, we would be happy
to work with the Senate to seek further information from institutions that have switched from
paper to online evaluations. We have already received results of preliminary inquiries to UCLA
who switched to online course surveys two years ago. We would be happy to follow up with
additional inquiries to the other UC campuses who have already made this switch on their
results.

Cost Savings

Comments/Concerns:
• Can the cost savings be quantified? KKB

Response:
In terms of quantifying costs of the paper system to academic departments, we were able to
collect personnel time data and printing costs for fall quarter 2014 from a set of eleven
departments. We generated averages based on that set of data to develop quarterly and annual
costs per department and for the campus as a whole. The results of that analysis and the
estimates are presented in Table 1 below.

| Table 1 - Academic Department Costs for ECSI Logistics |
|-----------------------------------------------|-------------------|-----------------|-----------------|-----------------|
| Category                                | Units  | Amount | Rate   | Cost  | Totals |
|-----------------------------------------------|-------------------|-----------------|-----------------|-----------------|
| Survey Request                           | Hours  | 1.25    | $20.00 | $25.00          |        |
| Printing Questionnaires                  | Hours  | 2.25    | $20.00 | $45.00          |        |
| Packet Collation/Distribution            | Hours  | 9.25    | $20.00 | $185.00         |        |
| Survey Return                           | Hours  | 9.50    | $20.00 | $190.00         |        |
| Comment Processing                      | Hours  | 4.25    | $20.00 | $85.00          | $530.00 |
| Questionnaire Printing Cost             | Page   | 2.625   | $0.05  | $131.25         | $131.25 |
| Total/Quarter:                          | $661.25 |
| Personnel Costs                         | Hours  | 1537.00 | $20.00 | $30,740.00      | $30,740.00 |
| Questionnaire Printing Cost             | Page   | 152,250 | $0.05  | $7,612.50       | $7,612.50 |
| Total/Quarter:                          | $38,352.50 |
| Personnel Costs                         | Hours  | 6,148.00| $20.00 | $122,960.00     | $122,960.00 |
| Questionnaire Printing Cost             | Page   | 609,000 | $0.05  | $30,450.00      | $30,450.00 |
| Total/Year:                             | $153,410.00 |
These estimates are probably quite conservative given that we estimated staff costs at only $20/hour, but in many departments a more expensive staff member actually handles many of these tasks. The bottom line is that the paper ESCI questionnaire system easily costs campus academic departments over $150,000 per year for labor and printing costs. By comparison, annual cost to Instructional Development for managing the central administration of the ESCI system is only a little over $42,000 for equipment, software, maintenance agreements, custom printed Scantron sheets, and student assistance (see Table 2 below). That does not include personnel costs for the ESCI Administrator and ESCI Assistant.

As was pointed out in both the original proposal and the Ad Hoc Committee’s report, any cost savings to the campus that come from converting to collecting ESCI data online will be to the academic departments, and not Instructional Development or the Administration. For academic departments, shifting to online ESCI surveys would eliminate all labor and printing costs other than the labor involved in completing the Survey Request Form each quarter (on average 1.25 hours of time). Another way to look at it is that moving to online surveys will save on the order of 5,858 hours of academic department staff time campus-wide per year.

<table>
<thead>
<tr>
<th>Category</th>
<th>Units</th>
<th>Amount</th>
<th>Rate</th>
<th>Cost</th>
<th>Totals</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Equipment</td>
<td></td>
<td>$2,960.42</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Operating Supplies</td>
<td></td>
<td>$2,780.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Custom Printed Response Forms</td>
<td></td>
<td>$17,635.68</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Student Assistance</td>
<td></td>
<td>$8,580.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Furniture Services</td>
<td></td>
<td>$100.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shredding</td>
<td></td>
<td>$213.36</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transportation</td>
<td></td>
<td>$200.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hardware Maintenance Agreements</td>
<td></td>
<td>$5,165.76</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software</td>
<td></td>
<td>$1,995.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software Maintenance Agreements</td>
<td></td>
<td>$1,443.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Software Tech Support</td>
<td></td>
<td>$1,443.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grand Total</td>
<td></td>
<td>$42,516.22</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Department averages based on data from: Black Studies; Computer Science; Economics; Evolution, Ecology and Marine Biology; Earth Science; Film & Media Studies; Global Studies; Materials; Music; Physics; and Theater and Dance.

Campus estimates based on 58 academic departments.
IT Resources

Comments/Concerns:
• Questions whether the current IT systems have the necessary capacity. They say, “It is imperative to allow for simultaneous and synchronous data collection and we are not convinced that this can be handled well.” HB, CRIIR
• Questions arose about costs of enhanced security systems if necessary. HB-CRIIR
• There was concern that the IT infrastructure does not have the necessary bandwidth and infrastructure to sustain online evaluation, particularly if the whole evaluation time frame is during the last week of classes. KKB
• Example of email switchover problem KKB
• Human error (in conversion of surveys from paper to online) KKB

Response:
In preparation for the ESCI Online Pilot, Instructional Development, at the suggestion of the Implementation Sub-Committee of the Ad Hoc Committee, invested in new, dedicated server hardware on which to run the survey service. The server hardware is housed in the North Hall Data Center, and hence is in a physically secure environment with more than adequate network bandwidth, redundant power supplies using conditioned power, and with more than adequate cooling and ventilation. The survey service is run as part of a highly customized Drupal instance with the intention of being able to add capabilities to the system as need and faculty suggestions allow. During the Pilot Program at no time did the system ever come close to capacity on the number of simultaneous requests/submissions. At the point in time where we get approval to offer online surveys as a production service for departments on request, we can easily increase capacity of the system by adding one to four more web servers with a load balancing module to handle any significantly increased simultaneous demand. Network bandwidth should never be a problem, as the amount of data required for transmission is actually quite small compared to, for example, video streaming (which ID also has experience hosting).

Physical security of the entire ESCI system is taken care of by housing all ESCI related servers, web services and database servers in the North Hall Data Center. Data security of the system has been reviewed with the previous Campus Security Officer, and he had no issues or concerns with the measures that had been put in place. No FERPA data is ever transmitted through the system. All connections to the system are encrypted. Student responses are tagged with a respondent number (in the event a faculty member would like to do an item analysis on the ESCI results), but the respondent number has absolutely no connection to the specific student. All responses are kept strictly anonymous.

In short, however, based on our discussions with the Campus Information Security Officer, the system developed for the ESCI Online pilot met all of his criteria for data security and process integrity. The current paper system is more prone to security problems than the online system has been. This issue of security and other problem associated with the paper ESCI survey system are discussed, and examples given, below.

During the early phases of the Pilot we identified potential problem areas in matching the correctly formatted version of the surveys to the correct courses, and we were able to develop internal, programmatic safeguards to eliminate those problems. In many ways the online survey process is far less prone to logistical and process errors than the traditional paper survey system. Some examples of errors and security problems that occur with paper surveys include:
Assembling and Distributing Survey Packets

- Departments have included a mix of questionnaires in a survey packet. If the mix is faculty and TA questionnaires, departmental normative information could be contaminated for both faculty and TA norms.

- Departments have included a mix of ESCI response forms in a survey packet. If this happens to a faculty survey where the mix is TA and Faculty response forms, the faculty member will not have data for Campus Items A & B for those who responded using the TA response form.

- One department uses two different types of Faculty response forms; one form has a 5-response scale, and the other has a 7-response scale. There have been times where a mix of these two forms occurred. For this type of problem, if the questionnaire was referring to a 5-point response scale, then all of the student responses given on the 7-response scale form must be discarded.

- The survey packets have been incorrectly distributed out to the instructors. For example, faculty has mistakenly been given a TA survey packet for their course.

Survey Completion

- There are limitations with acceptable writing utensils used to bubble in scantron response forms. The mark sensing scanner can only read items bubbled with blue or black ink and pencil. Data goes unread on forms using red ink.

- Students have answered questions with a response that is out of the range of response choices. For example, if the response choices for a question are “a) True” or “b) False”, and a response of “C” is given, that response is out of range and is considered blank.

- Students have submitted responses for more items than were on the questionnaire. These are cases where the department has no questionnaire that matches the number of items responded to by the student. This becomes a survey that cannot be processed.

- Response forms that have multiple marks for a single item will cause the scanner to stop scanning (adding delay to the overall processing time). The sheet must be inspected to determine if it is obvious what the intention of the respondent was. This is corrected by hand and rescanned. This kind of error is frequently encountered during the scanning process. On average, one in every 25 sheets.

Returning Surveys

- Completed evaluations have been misplaced and returned to the ESCI office several quarters later.

- Multiple surveys have been merged together appearing as a single survey. For example, a TA with multiple discussion sections has been combined into one survey.

- Separate faculty surveys have accidentally been bundled together as one survey. This type of error is particularly hard to spot because the resulting survey appears as a single survey. In order to find this type of occurrence, a department must notice that a survey is missing. In order to find the missing survey, someone must peruse the entire department’s submitted surveys, one sheet at a time.

- Blank ESCI response forms are not stored in a secure fashion by some departments. It is possible for an individual to obtain a batch of blank ESCI response forms. For example, there was a case where an instructor was able to pre-fill a set of blank response forms and submit them as part of their survey.
• Blank response forms are not secured in the classroom. When distributing the blank response forms in the classroom, it is possible for a student to grab more than one form.

• A whole quarter’s worth of completed surveys and response sheets were sent to recycling by a new staff member in an academic department who thought that they were old files.

Communication

Comment/Concern:
The current survey does not ask about the level of student engagement in the course (did they attend class, did they do the reading/homework prior to lectures, etc.). HB

Response:
All ESCI surveys, by design are completely customizable by the faculty member and/or their department. Unfortunately, most faculty members and departments are either unaware of this, or are so pressed for time that they do not request changes to their surveys. The only two campus-wide required items on faculty ESCI surveys are Items A and B. There are no campus-wide required items of TA surveys. Many departments have departmentally required items on either faculty surveys or TA surveys or both, but faculty members can, and always have been able to, add survey items, or change survey items from quarter to quarter. The only requirement is that they either use an existing item from the item pool of approximately 5,000 questions, or work with an Instructional Consultant from ID to craft a completely new item to add to the item pool. There are questions available in the item pool that address student engagement, and additional items can be added for those faculty or departments that want to include them in their ESCI surveys. The item pool includes, and can accommodate, both scalar response questions and open-ended questions. That is true regardless whether the survey is administered on paper or online.

Comment/Concern:
We are also concerned that online evaluation may be used for graduate seminars, which seems antithetical to the pedagogical modality of seminars.” KKB

Response:
Department staff determine which courses are to be surveyed, ideally in consultation with the faculty as part of the quarterly Survey Request Process. In our instructions to the department staff we discourage them from requesting surveys for graduate seminars, small undergraduate seminars, independent studies courses and tutorials. We specifically advise them against requesting a survey for any course with fewer than ten students enrolled. For courses of that small size, there are far more effective and personal ways to for faculty members to get the feedback that they need. None of this process changes with shifting to delivering the surveys online rather than on paper.

Comment/Concern:
Members thought it would be useful to break down the data further for evaluation, particularly by gender and ethnicity.” KKB

Response:
While an analysis of that kind might prove useful, the ESCI system does not capture the pertinent demographic data to enable that kind of cross-classification analysis. The ESCI staff would be happy to work with any department that chooses to undertake such an analysis, providing that they obtain their faculty member’s permission to use their data in such a fashion, and of course the department would have to provide the demographic data on those faculty members.
Comment/Concern:
Wonders how partially-completed evaluations would be managed. CPB

Response:
In the online system, as with the paper system, partially completed surveys are handled the same way: any extant responses to items are counted, and any items for which a student neglected to make a response is simply a blank response. That is why there can be a disparity in the response rate from one item to another on the same ESCI report, and why we always report individual item response counts, blank counts, and response rates, rather than an over all response rate for the survey as a whole.

Response Rates

Comment/Concern:
• (On increasing response rates with early access to grades): Do we even have evidence that faculty are able to submit grades significantly prior to the deadline to make the offer of this incentive meaningful?” HB

• Students would either have to complete the evaluation or “actively” opt out of doing it before having access to their grades. HB and KKB

• The College of Engineering FEC recommends that individual faculty be given discretion about whether it should be mandatory for students to fill out an online evaluation.

Response:
It has been suggested that early access to grades or withholding of grades will/should be used as an incentive to complete surveys. However, the current architecture of the student information system precludes either the option for early grade release and the option for withholding grades. Some feel that faculty should be able to decide whether it should be mandatory for students to fill out an online evaluation. The Ad Hoc Committee on Online Course Evaluation recommended that neither incentives nor disincentives be used to increase response rates. Thus, early access to grades and withholding of grades are not being considered as strategies to increase response rates. Rather the Committee recommended, and the pilot program implemented a strategy that included:

• frequent email reminders to students (55% of students indicate this was most effective)

• an embedded reminder icon in GauchoSpace on log-in (18% of students found this effective)

• best practices tips for faculty and TA’s whose departments were participating in the pilot to actively promote student participation (See Appendix C)

• a real-time response rate report that faculty and TAs could use to monitor response rates for their courses during the survey period

The suggestion that faculty should be given discretion about whether it should be mandatory for students to fill out an online evaluation runs counter to the recommendations of the Ad Hoc Committee, and without the ability to withhold grades, has no enforcement mechanism. A better alternative would allow faculty members to have discretion about whether to award extra point credit for completing an online survey, which is consistent with Wode and Kaiser’s (2011) recommendation 1 (see below). It would be possible to provide faculty members with a report on students who had completed surveys online in time for final grade to be submitted, allowing them to add the extra credit points to the student’s final scores. Research in this area, as well as on-campus experience with minor participation points awarded of participation in class using
student response systems (iClicker) demonstrate that this provision of a small number of extra points has an outsized inducement effect for the tiny overall effect on the student’s final grade.

Comment/Concern:
Evaluations should be filled out before grades are available to students. KKB

Response:
Regardless whether the surveys are conducted on paper or online, the survey period will always be over before grades are available. In addition, the final ESCI reports for courses will never be available to faculty until well after final grades for courses have been submitted.

Comment/Concern:
• How best to improve response rates for online ESCIs, and how to avoid a possible selection bias that could accompany lower response rates? KKB
• Student responses will likely represent the extremes, both positive and negative, thereby losing the middle and providing skewed and/or inconclusive results. HB

Response:
The research literature, as summarized by Wode and Kaiser (2011) suggests that there are three main ways to improve response rates: 1) make evaluation a part of the course (most effective), 2) send reminder notices and 3) offer a small incentive.

While most of the research literature suggests that average student ratings do not vary significantly between paper-based evaluations and online evaluations, we agree that the distribution of scores should be compared for the two modes of delivery. This could be accomplished by analyzing online evaluations for faculty members who have taught large and small lecture for several years and have had stable ESCI distributions for their paper-based evaluations. We have conducted such an analysis and those results are presented in the section below titled “Clarification”.

Comment/Concern:
On response rate benchmark—CAP takes strong issue with the report’s statement that “90% of our students complete evaluations on average.” KKB

Response:
We appreciate CAP raising a red flag about the reported average response rate of 90%. When we investigated this calculation, we found that the data contained response rates of >100%, which elevated the average. In response, we have reviewed the data and reset all response rates of >100% to 100%. We also removed all cases where the response rate was 0. Once these changes were implemented, we recalculated the average response rates for the original set of data used for the report for all courses taught by Faculty between 1991 and 2008. That resulted in an average response rate of 62% (SD=17%). Based on that revised mean for the original data set compared to the reported mean response rate for the ESCI Online Pilot in the original Committee report, the difference between the two approaches was much less than originally reported. In an effort to do a more complete analysis on more current data, we went further and analyzed responses rates for all paper-based faculty surveys conducted between 2000 and 2014 and found a mean response rate of 73.4% (SD = 19.52%, Median: 75%, N=52,723 course surveys and 1,868,784 survey sheets). This updated analysis also indicates that response rates deviated less from the norms for paper surveys than originally reported in the initial ESCI Online Pilot report. For the full set of ESCI Online surveys, the average response rate was 56.31% (SD=24.47%, Median: 52.55%, N=480 course surveys). Figure 1 below provides a comparison box plot of response rates for the ESCI Online Pilot compared to the 14 years of paper data. For
paper surveys, note that there are a large number of outliers below the first quartile. Note also, that for both sets of data analyzed, the data were confined to faculty surveys only as those are the only surveys that contain the two consistent questions, Items A and B. There is no way to do a similar analysis across campus for TA surveys because no TA surveys contain the same items between departments.

![Comparison of Response Rates Online vs. Paper](image)

**Figure 1. Comparison of Paper and Online ESCI Response Rates**

Additional analysis of response rates for paper surveys indicates that while the mean and median response rates annually for the whole campus are quite stable, there is considerable variability from quarter to quarter, with Spring quarter consistently having lower overall response rates. That pattern is illustrated in Figure 2 below.
A much higher degree of variability in the paper response rate data is revealed when the data are analyzed by department over time. A box plot of that analysis is presented in Figure 3 below. When viewed in this way it becomes clear that there are some departments that have historically maintained very high response rates, and others that have had consistently low response rates using paper surveys. This degree of variability completely encompasses the variability in the response rates for the online surveys, as illustrated in Figure 4 below.
Without disclosing the departments, we can report that the variation in response rates for online surveys is consistent with the historical variation in responses rates on paper surveys for those departments. Four of the six departments show substantial overlap in response rates between paper and online surveys. The large difference in department “G” in the graph we attribute to a very small sample of online surveys, as that department has significantly shrunk since the period covered by paper surveys, and so is not a very reliable indicator.
Concern:
Students who do not attend class will be able to complete online surveys. HB and KKB

Response:
Most paper-based ESCI evaluations are administered during the last week of classes (dead week). While we have not researched attendance rates during dead week vs. other weeks of the quarter, anecdotal evidence suggests that more students attend the final week of classes, as review sessions are oftentimes held during this time. Thus, it may be the case that students who do not regularly attend classes are already given the opportunity to complete ESCI evaluations. We appreciate your suggestions for how to address the potential issue of students evaluating courses who did not attend on a regular basis.

Comment/Concern:
Most groups suggested that online ESCIs be filled out during class time, although it was also suggested that need not be the only time that evaluations are available for completion. KKB

Response:
The Ad Hoc Committee recommended that online evaluations NOT be filled out during class time. Asking students to complete an online survey in class on either a laptop, tablet or smartphone as present may be problematic for two reasons. When asked in winter quarter 2015 only 88% of respondents self-reported that they have an appropriate device that they could bring to class to engage in the online survey in the classroom. In addition, the campus would need to
make substantial additional investments in wifi access points serving the classrooms to ensure both adequate coverage in all general assignment classrooms, and a sufficient number of access points to reliably handle the volume of simultaneously connected devices. Importantly, when asked, many students indicate that their preference for completing ESCI surveys online rather than in class is to avoid losing precious instructional time during dead week to the survey process. This is a recommendation that could be revised once the issue of equitability and reliable network coverage have been addressed, or for courses that are already regularly taught in a computer equipped space; as we recognize that allowing class time for online evaluations could greatly improve response rates.

**Comment/Concern:**
*Reviewing groups suggest that loaning laptops or similar to students during class time would likely enhance participation rates. Multiple Committees/Individuals*

**Response:**
It may be possible, through Collaborate, to develop a laptop/tablet loan program for both in-class online ESCI and in-class online examinations. That would require a significant investment given that 12% (~2400 students) of our undergraduate students would need to participate in such a loaner program. At the present time however, even if such a loan program could be funded and effectively managed, the problem of insufficient wifi access points to reliably cover the general assignment classroom inventory with service is still a primary constraint on this approach.

**Clarification**

**Comment/Concern:**
*Stated that while they appreciated seeing the data in the report, “the conclusions drawn do not follow from that data and the statistical methods used were [inadequate]. KKB-CRIR*

**Response:**
Based on dissatisfaction with the initial analysis of the data, a more thorough analysis was conducted, including additional data not originally available at the time the original report was prepared. The new pilot dataset includes comparative response data from a total of 270 courses; 135 included in the pilot, matched to the ESCI responses from the last time the same course with the same instructor was surveyed using paper forms. For the new analysis, in addition to basic descriptive statistics we also applied paired T-tests to means and response distributions, whereas the original analysis only focused on descriptive statistics applied to differences in the means. As one faculty member reviewer correctly pointed out, the means are not particularly informative indicators, whereas the distributions of responses are far more informative. In addition, we conducted an analysis to determine if there was an important change in the deviation of response percentages among the choices “Excellent”, “Very Good” “Good”, “Fair” and “Poor” for Items A and B. The analysis was confined to Items A and B, as those are the only two items common to all faculty ESCI surveys. In all of these analyses, the last paper-based survey for the course and instructor are treated as the baseline, all measures are treated as deviations above or below that baseline in the surveys collected online. The full data set, with course, instructor and department identifiers removed, is available on request as an Excel spreadsheet with all department, course and instructor information redacted. It is simply too large a dataset to present in printed form.

**Guidelines for Interpretation**

Our interpretation of the analysis of the changes in distributions is informed by the same advice on interpreting ESCI data that we have provided to faculty for many years. The full explanation is available online at: [http://oic.id.ucsb.edu/interpreting-data](http://oic.id.ucsb.edu/interpreting-data). In particular the guidance on
interpreting the percentage differences in ratings between the current course and the department norms is helpful in this context. In general, we suggest that by themselves these results (Items A and B) are accurate enough only to place instructors and courses into three broad categories: the truly outstanding, those with problems, and the vast majority in the middle.

Results are always reported in terms of percentage distributions, so it is easy to look at the results and see that, for example, “46% of students rated the teaching as Excellent or Very Good.” Looking at the entire distribution allows for determining if responses are clustered in a couple of categories (“They loved it!”), or spread across the categories (“Some loved it, some thought it was OK, and some hated it”), or perhaps even bimodal (“Some loved it and some hated it, but nobody was indifferent”).

In using the comparison norms, individual results will differ to some extent from those of the department during the current quarter, the department over time, and the campus over time. It is important to understand when those differences are educationally meaningful and when they are not. It is possible to use statistical procedures to compare two distributions, but it is not uncommon for the number of student respondents to be so large that any difference is statistically significant, even though it may not be educationally meaningful. As a rough rule of thumb to decide when the differences between individual results and the “norms” may be meaningful, we suggest the following:

- If the percentage of students who rate your course in any particular category differs by about 10% from the percentage in a norm group, then there MAY be something educationally meaningful going on, and it’s worth examining.

  For example, suppose your department this quarter has 54% of students rating “the overall quality of the instructor’s teaching” as “Excellent,” and 64% of your students rate your teaching as “Excellent.” Then you may be doing a better job in this course than the departmental average, and you could look at responses to other questions on the survey and to other sources to understand the differences, and to decide for yourself whether they are meaningful.

- If for any response category the percentage differs by about 20%, then there’s probably something meaningful going on, and it’s definitely worth seeking further understanding.

- If the percentage differs by about 30%, then there IS something meaningful going on.

**Analysis of Means**

In examining the effect of conducting the surveys online compared to paper, we conducted a paired T-test to determine if there was a significant difference in the mean scores for Items A and B. Table 3 presents those results.

<table>
<thead>
<tr>
<th></th>
<th>Item A</th>
<th>Item B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Difference in Mean Scores</td>
<td>0.16</td>
<td>0.13</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.52</td>
<td>0.53</td>
</tr>
<tr>
<td>N</td>
<td>135</td>
<td>135</td>
</tr>
<tr>
<td>T Score</td>
<td>3.6405</td>
<td>2.7821</td>
</tr>
<tr>
<td>T-test (P)</td>
<td>0.0004</td>
<td>0.0064</td>
</tr>
</tbody>
</table>
As the table indicates, the average difference in mean scores for Item A in the online surveys are only slightly worse (higher scores by .16) than for the paper surveys. The effect is even less for the results for Item B (a difference of .13). Both have very similar standard deviations. The results of the paired T-test shows that there is a statistically significant difference in means for both Items A and B between online surveys and paper surveys with a significance level of 0.05. That is, the value of P for the paired T-test in both cases is less than 0.05, causing us to reject the null hypothesis that the difference of means for Items A and B are the same between the two survey methods. While the differences are statistically significant, the average differences are very small with relatively large standard deviations.

**Analysis of Response Distributions**

We also conducted an analysis of differences in the distribution of responses on the five point scale for both Items A and B between online and paper survey results. Results for Item A are presented in Table 4 below. Item A is the rating of the instructor.

**Table 4 - Analysis of Changes in Distributions Item A**

<table>
<thead>
<tr>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Change</td>
<td>-10%</td>
<td>4%</td>
<td>4%</td>
<td>0%</td>
<td>0%</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>31.0%</td>
<td>25.3%</td>
<td>16.6%</td>
<td>7.9%</td>
<td>4.5%</td>
</tr>
<tr>
<td>N</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
</tr>
<tr>
<td>T Score</td>
<td>3.621</td>
<td>2.046</td>
<td>2.958</td>
<td>0.448</td>
<td>0.561</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.0004</td>
<td>0.0435</td>
<td>0.0038</td>
<td>0.6559</td>
<td>0.5773</td>
</tr>
</tbody>
</table>

For Item A, the analysis indicates no significant difference in the distributions of responses for the two lowest rating categories (4 and 5) as the value of P is greater than our significance level of .05. There is a significant difference in the mean distributions for the three highest rating categories (1, 2 and 3) indicating that there is a subtle, but statistically significant, shift in ratings between these three highest categories. There is an average shift from “Excellent” to “Very Good” and “Good”. That is also indicated by the 0% mean change for responses 4 and 5 and the very small standard deviations for responses 4 and 5 compared to the other three. Using the 10%, 20% and 30% cut off guides from above however, the changes in distributions are so small as to not be meaningful. More importantly, there is no indication of a bimodal distribution of responses, as some faculty members feared, with only very low ratings and very high ratings.
Results for Item B are summarized in Table 5 below. Item B is the rating of the course overall.

### Table 5 - Analysis of Changes in Distributions Item B

<table>
<thead>
<tr>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Change</td>
<td>-7%</td>
<td>2%</td>
<td>5%</td>
<td>-0%</td>
<td>0%</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>28%</td>
<td>24%</td>
<td>18%</td>
<td>11%</td>
<td>4%</td>
</tr>
<tr>
<td>N</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
</tr>
<tr>
<td>T Score</td>
<td>3.009</td>
<td>1.073</td>
<td>3.294</td>
<td>0.373</td>
<td>1.117</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.0032</td>
<td>0.2869</td>
<td>0.0013</td>
<td>0.7104</td>
<td>0.2677</td>
</tr>
</tbody>
</table>

In this analysis there is no statistically significant difference in the ratings in the middle (2) and the low end (4 and 5) of the scale. There is a significant difference in the ratings for rating categories 1 and 3. The percentages of mean change in distributions indicates that students rating the course online tended to shift from “Excellent” to “Very Good” and “Good” when rating the course overall. Once again, while the changes are statistically significant, examining the magnitude of the changes in the distributions indicates that these are not meaningful changes applying the 10%, 20% and 30% guidelines.

### Deviation of Mean from Department Norms

An analysis of the pattern of deviation of individual rating means from department rating means was also conducted. These results are presented in Table 6 below.

### Table 6 - Analysis of Changes in Deviation of Means from Dept. Norms for Items A and B

<table>
<thead>
<tr>
<th>Measure</th>
<th>Item A</th>
<th>Item B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average Difference in Mean Scores</td>
<td>-0.06</td>
<td>-0.21</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0.53</td>
<td>0.28</td>
</tr>
<tr>
<td>N</td>
<td>135</td>
<td>135</td>
</tr>
<tr>
<td>T Score</td>
<td>1.425</td>
<td>8.947</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.1581</td>
<td>0.0000</td>
</tr>
</tbody>
</table>

Again using the paired T-test, there is no significant difference in the mean deviations of the means between paper and online surveys from the department mean for Item A. There is a statistically significant difference in the mean deviations of the means between paper and online surveys from the department mean for Item B. In this case, for Item B, the pattern of deviation is that over all course rating means were very slightly better as a whole that the various department norms.
Deviation of Rating Distributions from Department Norms

As was indicated in the guidelines to interpretation above, the most meaningful feedback in ESCI results is in comparing the distributions of ratings for the current course to the department norms so see if there are important deviations either above or below the norms. To determine if there was any pattern in deviations from department norms between online and paper surveys, we conducted a comparison and paired T-test for those metrics as well. The results for Item A are reported in Table 7 below. There is no significant difference in the mean deviations from department norms at the .05 significance level for the “Excellent”, “Very Good”, and “Good” ratings. There is a significant difference for the two lowest ratings categories, “Fair” and “Poor”. While statistically significant, the percentage changes of -1% are not large enough to be meaningful. The same can be said for the very small percentage changes in the upper three ratings categories as well.

Table 7 - Analysis of Changes in Deviation of Distributions from Dept. Norm for Item A

<table>
<thead>
<tr>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Change</td>
<td>-1%</td>
<td>2%</td>
<td>1%</td>
<td>-1%</td>
<td>-1%</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>31%</td>
<td>25%</td>
<td>17%</td>
<td>8%</td>
<td>5%</td>
</tr>
<tr>
<td>N</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
</tr>
<tr>
<td>T Score</td>
<td>0.257</td>
<td>0.780</td>
<td>0.462</td>
<td>2.166</td>
<td>2.849</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.9609</td>
<td>0.4384</td>
<td>0.6464</td>
<td>0.0327</td>
<td>0.0052</td>
</tr>
</tbody>
</table>

The results of this analysis for Item B are reported in Table 8 below. There is no significant difference in the mean deviations from departments norms at the .05 significance level for the “Excellent”, “Very Good”, “Good” and “Fair” ratings. There is a significant difference for the lowest ratings category, “Poor”. While statistically significant, the percentage changes of +/- 1% to 2% are not diagnostic or particularly meaningful.

Table 8 - Analysis of Changes in Deviation of Distributions from Dept. Norm for Item B

<table>
<thead>
<tr>
<th>Measure</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean Change</td>
<td>2%</td>
<td>2%</td>
<td>1%</td>
<td>-2%</td>
<td>-1%</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>29%</td>
<td>25%</td>
<td>18%</td>
<td>11%</td>
<td>4%</td>
</tr>
<tr>
<td>N</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
<td>135</td>
</tr>
<tr>
<td>T Score</td>
<td>0.956</td>
<td>1.091</td>
<td>0.547</td>
<td>1.647</td>
<td>3.275</td>
</tr>
<tr>
<td>T-Test (P)</td>
<td>0.3426</td>
<td>0.2792</td>
<td>0.5865</td>
<td>0.1032</td>
<td>0.0014</td>
</tr>
</tbody>
</table>
Analysis of Overall Impact on Ratings

A final analysis simply looked at the counts of cases where the ratings in each category for Items A and B were either meaningfully better for online versus paper surveys, meaningfully worse, or effectively no change applying the interpretation guidelines and using the 20% difference metric as the discriminator. The results for Item A are presented in Table 9, and the results for Item B are presented in Table 10.

Table 9 - Analysis of Importance of Change in Distributions Item A

<table>
<thead>
<tr>
<th>Category</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningfully Better (&gt;=20% Increase) N</td>
<td>21</td>
<td>31</td>
<td>21</td>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>No Meaningful Change N</td>
<td>70</td>
<td>84</td>
<td>108</td>
<td>127</td>
<td>133</td>
</tr>
<tr>
<td>Meaningfully Worse (&gt;=20% Decrease) N</td>
<td>44</td>
<td>20</td>
<td>6</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>% Better</td>
<td>16%</td>
<td>23%</td>
<td>16%</td>
<td>4%</td>
<td>1%</td>
</tr>
<tr>
<td>% No Change</td>
<td>52%</td>
<td>62%</td>
<td>80%</td>
<td>94%</td>
<td>99%</td>
</tr>
<tr>
<td>% Worse</td>
<td>33%</td>
<td>15%</td>
<td>4%</td>
<td>2%</td>
<td>1%</td>
</tr>
</tbody>
</table>

Table 10 - Analysis of Importance of Change in Distributions Item B

<table>
<thead>
<tr>
<th>Category</th>
<th>1 Excellent</th>
<th>2 Very Good</th>
<th>3 Good</th>
<th>4 Fair</th>
<th>5 Poor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meaningfully Better (&gt;=20% Increase) N</td>
<td>20</td>
<td>28</td>
<td>23</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>No Meaningful Change N</td>
<td>73</td>
<td>88</td>
<td>107</td>
<td>124</td>
<td>134</td>
</tr>
<tr>
<td>Meaningfully Worse (&gt;=20% Decrease) N</td>
<td>42</td>
<td>19</td>
<td>5</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>% Better</td>
<td>15%</td>
<td>21%</td>
<td>17%</td>
<td>4%</td>
<td>0%</td>
</tr>
<tr>
<td>% No Change</td>
<td>54%</td>
<td>65%</td>
<td>79%</td>
<td>92%</td>
<td>99%</td>
</tr>
<tr>
<td>% Worse</td>
<td>31%</td>
<td>14%</td>
<td>4%</td>
<td>4%</td>
<td>1%</td>
</tr>
</tbody>
</table>

For Item A, 33% of ratings in the “Excellent” category in the online surveys deviated by 20% or more below the rating in that category on paper. The fall-off curve of meaningfully worse ratings form the “Very Good” through “Poor” ratings categories indicates that all of the shift that occurred is on the high end to middle of the ratings. Recall Item A is the rating of the instructor. For Item B, the rating of the course, the pattern is quite similar.

Conclusion

Based on the combination of the discussion of the changes in response rates between paper and online surveys, and the preceding analysis of changes in the ratings two things are clear. First, that the response rate goes down in the switch from paper to online surveys. This is consistent with all of our research and the findings of peer UC schools who have already made the switch.
Second, while response rates do go down, there is no dramatic negative or positive effect on instructor or course ratings. This is also consistent with results from UC peer institutions. Is it possible to increase response rates? Yes, in three ways. First instructors who actively monitor their online response rates and engage with the students to encourage them to participate have demonstrably higher response rates than those who do not. Second, it would be possible to report to faculty members those students who did participate in the online survey before grades are due, and faculty members could award some small number of participation points toward the student’s grade as an inducement. Third and finally, as online ESCI surveys become more the norm on campus, more students would be familiar with it and hence may be more likely to engage, particularly if the first two recommendations above also came into play.

**General Issues**

**Comment/Concern:**
- “Course evaluations [are] fundamentally part of the instructional responsibility and privilege of the faculty,” and CDE requests that “the ESCI process be made a more self-reflective process for the students.” CPB
- Non-ESCI evaluations – qualitative evaluations – are often critical for analyzing teaching, yet the Report contains little commentary on how a transition between ESCI scores and other methods of evaluation would be achieved. CPB

**Response:**
This issue was extensively discussed by the Ad Hoc Committee as a whole as well as by the Content Sub Committee. While the current ESCI Item Pool, has some self-reflective questions, there is no reason that more could not be developed and departments and faculty members encouraged to use those in their questionnaires. Similarly, there are already open-ended questions in the ESCI Item Pool. Some departments use them, and many others do not, we suspect primarily because they do not have the staff resources to manage transcription of the responses to maintain student anonymity. There again is no reason that more open-ended questions could not be developed and added to the ESCI Item Pool. Furthermore, for departments who chose to switch to an online ESCI survey system, all of the logistical friction of using open-ended questions is removed as there would be no need for department staff time to be spent on processing those responses. In fact the responses to open ended questions for online surveys can, and have been, easily incorporated into the standard ESCI report for faculty and TAs.

**Comment/Concern:**
**Pedagogical rationale for moving the process online. CPB**

**Response:**
There is no specific pedagogical rational for moving the ESCI survey process online. From a survey process standpoint moving to online surveys allows for the delivery of a potentially more flexible, richer (through the regular inclusion of self-reflective and open-ended questions for qualitative information), and logistically simpler system for faculty, TAs and academic department staff. Because the response pools are different between surveys collected online and those collected on paper, it would also be possible to process the results of the online surveys much more quickly and get the feedback to the faculty weeks earlier than the current process allows. From a pedagogical standpoint, the move to an online system offers the potential for providing faculty members with
more useful qualitative and quantitative feedback more quickly to help inform them of pedagogical changes that they may want to make early enough in the following quarter to have a meaningful difference in student reaction to their courses.
ONLINE ESCI PROCESS SURVEY
ID 1 0100
Survey Number: 299570

1. The online end of quarter survey is an improvement over the traditional paper-based end of quarter survey.

   (a) Strongly Agree (b) Agree (c) Neither Agree Nor Disagree (d) Disagree (e) Strongly Disagree

   Response weighting: 1 2 3 4 5
   (a) (b) (c) (d) (e)

   **This COURSE current quarter**
   42% 32% 20% 5% 2%

   **Blank Response**
   3

   **Total Students**
   1665

   **Courses**
   1

   **Mean Median**
   1.9 2.0

   **Student-weighted Norms (GR students)**
   42% 32% 20% 5% 2%

   **Dept ESCI INSTRUCTORS current qt**
   3

   **Dept ESCI INSTRUCTORS over time**
   11

   **Campus INSTRUCTORS over time**
   11

   **Dept ESCI INSTRUCTORS current qt**
   43% 31% 20% 4% 2%

   **Dept ESCI INSTRUCTORS over time**
   43% 31% 20% 4% 2%

   **Campus INSTRUCTORS over time**
   43% 31% 20% 4% 2%

2. I was able to provide more thoughtful feedback in this online format of the ESCI survey as compared to the paper-based survey.

   (a) Strongly Agree (b) Agree (c) Neither Agree Nor Disagree (d) Disagree (e) Strongly Disagree

   Response weighting: 1 2 3 4 5
   (a) (b) (c) (d) (e)

   **This COURSE current quarter**
   23% 25% 40% 9% 2%

   **Blank Response**
   12

   **Total Students**
   1665

   **Courses**
   1

   **Mean Median**
   2.4 3.0

   **Student-weighted Norms (GR students)**
   23% 25% 40% 9% 2%

   **Dept ESCI INSTRUCTORS current qt**
   12

   **Dept ESCI INSTRUCTORS over time**
   18

   **Campus INSTRUCTORS over time**
   18

Guidelines for “Interpreting ESCI Data” and a description of the “Report Output” can be found at http://oic.id.ucsb.edu/esci.

Department and Campus Norms taken over time span: Spring Quarter 2010 - Winter Quarter 2015

Special processing - Winter Quarter 2015 - ESCI Online Pilot

Appendix A - Winter 2015 ESCI Survey about Online Surveys

Page 1
### ESCI Online Pilot Survey Statistics

**SPECIAL PROCESSING - Winter Quarter 2015 -- ESCI Online Pilot**

**Department and Campus Norms taken over time span: Spring Quarter 2010 - Winter Quarter 2015**

**Guidelines for “Interpreting ESCI Data” and a description of the “Report Output” can be found at http://oic.id.ucsb.edu/esci.**

**NOTICE:** Please examine these evaluations upon receipt and immediately report any suspected errors to: ESCI Coordinator, Instructional Consultation, 1130 Kerr Hall (x4278).

---

#### Question 3.

I am confident that the online ESCI survey maintains my anonymity.

(a) Strongly Agree (b) Agree (c) Neither Agree Nor Disagree (d) Disagree (e) Strongly Disagree

<table>
<thead>
<tr>
<th>Response</th>
<th>Total Courses</th>
<th>Mean</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student-weighted Norms (GR students)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dept ESCI INSTRUCTORS current qt</td>
<td>34% 39% 20% 5% 2%</td>
<td>10</td>
<td>1665</td>
</tr>
<tr>
<td>Dept ESCI INSTRUCTORS over time</td>
<td>32% 38% 22% 6% 2%</td>
<td>20</td>
<td>2879</td>
</tr>
<tr>
<td>Campus INSTRUCTORS over time</td>
<td>32% 38% 22% 6% 2%</td>
<td>20</td>
<td>2879</td>
</tr>
</tbody>
</table>

---

#### Question 4.

Would you have a wireless device (laptop, smartphone, tablet) with you in class, that you could use to complete the ESCI survey in class?

(a) Yes (b) No

<table>
<thead>
<tr>
<th>Response</th>
<th>Total Courses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student-weighted Norms (GR students)</td>
<td></td>
</tr>
<tr>
<td>Dept ESCI INSTRUCTORS current qt</td>
<td>88% 12%</td>
</tr>
<tr>
<td>Dept ESCI INSTRUCTORS over time</td>
<td>84% 16%</td>
</tr>
<tr>
<td>Campus INSTRUCTORS over time</td>
<td>84% 16%</td>
</tr>
</tbody>
</table>

---

**ONLINE ESCI PROCESS SURVEY**

**ID** 1 0100

**Survey Number:** 299570

Page 2
5. What was the primary reason that made you decide to complete the online ESCI survey(s) for your course(s) this quarter?

(a) Instructor encouraged me to participate
(b) Automated e-mail reminder from ESCI
(c) Reminder generated by GauchoSpace
(d) Self-motivated
(e) Other

Response weighting: 0 0 0 0 0

<table>
<thead>
<tr>
<th>This COURSE current quarter</th>
<th>Blank</th>
<th>Total</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>17% 55% 18% 8% 2% (a) (b) (c) (d) (e)</td>
<td>2</td>
<td>1665</td>
<td>1</td>
</tr>
</tbody>
</table>

Student-weighted Norms (GR students)

<table>
<thead>
<tr>
<th>Dept ESCI INSTRUCTORS current qt 17% 55% 18% 8% 2%</th>
<th>2</th>
<th>1665</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dept ESCI INSTRUCTORS over time 18% 54% 18% 7% 2%</td>
<td>5</td>
<td>2282</td>
<td>2</td>
</tr>
<tr>
<td>Campus INSTRUCTORS over time 18% 54% 18% 7% 2%</td>
<td>5</td>
<td>2282</td>
<td>2</td>
</tr>
</tbody>
</table>
Omitted the last 54 pages, for reasons of space, containing the open-ended text responses to the questions:

6. What ideas can you offer for improving response rates as the campus transitions to Online ESCI?

7. Do you have any suggestions or concerns about the online ESCI survey?
Appendix B - Annotated Bibliography

Student Evaluations of Courses and Instruction: Online Response Rates
Compilation of Abstracts and Annotated Bibliographies


The literature review revealed several studies that found no statistically significant differences between delivery modes. Two also noted that students provided more comments in the online forms. Response rates varied widely. The University of Kentucky College of Pharmacy, driven by the faculty’s desire for more timely return of results (3 - 4 months typically), launched a pilot study of online evaluations in 3 courses. The response rates for the 3 courses were 85%, 89%, and 75%. The 9 courses using the paper forms averaged an 80% response rate (consistent with the 2 previous years also about 80%). The comments on the online forms were more frequent and longer than the paper forms. Students liked the online form better than the paper form and thought they could provide more effective and constructive feedback online.


Synopsis from Innovate: “Many administrators are moving toward using online student evaluations to assess courses and instructors, but critics of the practice fear that the online format will only result in lower levels of student participation. Joan Anderson, Gary Brown, and Stephen Spaeth claim that such a concern often fails to acknowledge how the evaluation process already suffers from substantial lack of engagement on the part of students as well as instructors; the online format, they assert, merely inherits the fundamental problem of perceived irrelevance in the process itself. After addressing the reasons behind this problem and discussing how well-designed online evaluations can still make a positive difference, the authors describe the development and implementation of a comprehensive, college-wide online evaluation survey at Washington State University’s College of Agricultural, Human, and Natural Resources. In reviewing the survey results, they found that class size, academic discipline, and distribution method played a negligible role in student response rates. However, they found that variances in response rate were significantly influenced by the relative level of participation among faculty members and department heads in the original development of the survey. The authors maintain that online surveys can make the process more relevant and meaningful to students, but they conclude that eliciting greater response rates will still require sustained support, involvement, and advocacy by faculty members and administrators.”


This paper provides a summary of the current research in online vs. paper evaluations as well as results from a student to compare the feedback results. The same form was given to 46 section pairings – one paper and one online. The online response rate was 31% (392 out of 1276 possible responses) and the paper was 69% (972 out of 1415). No significant difference was found in the quantitative ratings between the two methods. They examined the differences on an
“overall effectiveness” question in rating for faculty who were above the college average and then for faculty who were below the college average. Faculty who were above the average were scored slightly lower online and the faculty who were below the college average were scored higher online. There was no significant difference in the number of students giving open-ended feedback online, however, there was a significant increase in the length of open-ended feedback online.


The Department of Policy Analysis and Management at Cornell University did a study of course evaluation data from 1998 - 2001. Using the same form, data was analyzed from 29 courses (20 using the paper version, 9 using the online version). The study examined response rates and mean scores between the methods. While specific response rates varied, online was typically lower than the paper form. For example, in fall 2000 paper was 69% compared with 47% online. Using a 5 - point scale on their 13 questions, 4 questions had a significant difference in mean scores between methods. This was a greater than 0.10 difference with the web having the higher mean score. The other 9 questions had a less than 0.10 difference in mean scores again with web having the higher means.


Abstract: One way many universities have approached the process of better understanding and meeting the needs of their students is through student evaluations. The evaluation data provide not only diagnostic feedback but also useful information in terms of the quality of learning and teaching experiences. In an effective quality cycle, the data gathered are analyzed and used to make improvements. This is often referred to as ‘closing the loop’. However, for any evaluation data to be of value an important prerequisite for ‘closing the loop’ is that response rates should be sufficiently high to be representative of the student cohort. This paper describes how a faculty within Monash University utilizing only web-based surveys developed a successful communication strategy to elicit staff and student participation in the unit (subject) evaluation process which achieved a response rate as high as 83.2%.


Abstract: Students’ contact with universities is becoming increasingly computer-mediated. For instance, students can enroll, check exam marks, even attend some lectures online. Many universities now conduct at least some of their student evaluations of teaching or units online instead of using traditional paper questionnaires. Given the move to more flexible learning environments, with no requirement for students to attend all or any lectures, institutions are beginning to consider whether it is possible to move to online only systems of collecting student feedback. Whilst lower response rates appear to be a major concern with online surveys, commentators have pointed to several important advantages over paper forms: reduced paper use, reduced impact on class teaching time, less time spent scanning paper forms etc. It has also been suggested that the responses to open-ended questions provided online are more frequent, lengthy and thoughtful. This paper looks in more detail at the issue of student comments and reports the findings of several comparisons between paper and online approaches, looking at students enrolled internally and externally, and also recent graduates. The place of student
comments as a valuable source of data is considered, along with the implications for an improvement in qualitative data with respect to the more quantitative aspects of student feedback.


Murdoch University School of Engineering ran a pilot in 1999 of online course evaluations using the same form online as on paper. Students found the online form easier, faster, and felt it offered greater anonymity. The school has a 50% mandate for response rate in course evaluations. Typically paper evaluations had a 65% response rate. The online pilot averaged 31% with 4 of the 18 courses over the 50% mandate. The response rate range was a wide 3% to 100%. Because the pilot was inadequately promoted, some faculty didn’t know they were using online forms and didn’t adequately prepare students. Students noted that they felt no pressure to fill out the online evaluations. The investigators concluded that the quality of responses was the same because they received the same amount of comments online which is what is used most from the evaluation form.


Abstract: This study compares student evaluations of faculty teaching that were completed in-class with those collected online. The two methods of evaluation were compared on response rates and on evaluation scores. In addition, this study investigates whether treatments or incentives can affect the response to online evaluations. It was found that the response rate to the online survey was generally lower than that to the in-class survey. When a grade incentive was used to encourage response to the online survey, a response rate was achieved that was comparable with that to the in-class survey. Additionally, the study found that online evaluations do not produce significantly different mean evaluation scores than traditional in-class evaluations, even when different incentives are offered to students who are asked to complete online evaluations.


The College of Business And Economics at California State University, Northridge did a study with 16 professors to see how the method of evaluation affects response rate and if online treatments (incentives) affect the response rate. Each professor taught 2 sections of the same undergraduate business course. The same form was used in both methods. Instructors were randomly assigned into 1 of 4 groups using different incentives: 0.25% grade incentive for completion of an online evaluation (4 courses), in-class demonstration on how to do the online evaluation (2 courses), if 2/3 of the class submitted online evaluations students would receive their final grades early (2 courses), or a control group (8 courses). The online evaluations averaged a 43% response rate and the paper evaluations averaged 75%. Looking at just the control group, their average response rate was 29%. In the individual cases the incentives had the effect of increasing response rate (grade incentive 87% response rate, demonstration 53%, and early final grade 51%).
Abstract: Substantial efforts have been made recently to compare the effectiveness of traditional course formats to alternative formats (most often, online delivery compared to traditional on-site delivery). This study examines, not the delivery format but rather the evaluation format. It compares traditional paper and pencil methods for course evaluation with electronic methods. Eleven instructors took part in the study. Each instructor taught two sections of the same course; at the end, one course received an online course evaluation, the other a traditional pencil and paper evaluation. Enrollment in these 22 sections was 519 students. Researchers analyzed open-ended comments as well as quantitative rankings for the course evaluations. Researchers found no significant differences in numerical rankings between the two evaluation formats. However, differences were found in number and length of comments, the ratio of positive to negative comments, and the ratio of formative to summative comments. Students completing faculty evaluations online wrote more comments, and the comments were more often formative (defined as a comment that gave specific reasons for judgment so that the instructor knew what the student was suggesting be kept or changed) in nature.

Four institutions, University of Michigan Ann Arbor, Virginia Tech, University of Cambridge and University of Maryland, collaborated on an open source online evaluation system within Sakai. Response rates in the various pilots ranged from 32% to 79%. They found the key benefits of online evaluations to be security, validity, efficiency, cost-savings, rapid results turnaround and higher quality student comments.

The College of Education and Human Development at the University of Minnesota did a study on 314 class pairs (14,154 student evaluations) from fall 2002 to fall 2004. The goals were to see if there is a difference in response rate, a difference in response distributions, a difference in average ratings between the two methods and what are the common perceptions of each method. In the study group the online form averaged a 56% response rate whereas the paper version averaged 77%. Slightly more students responded on the high and low ends of the 7-point scale than did in the middle. There was no significant difference in the mean rating on 4 required questions.

This white paper outlines 9 best practices for moving to online course evaluations. Key benefits to moving online are listed as well as strategies to build response rates.
Texas Tech University studied 3 modes of surveying a random group of Texas Agri Science teachers. The 3 modes were e-mail, web, and paper. No significant difference in the reliability of the responses was found. However the response rates were 60%, 43% and 27% for paper, web and e-mail respectively.


Abstract: The purpose of this study is to compare the results of paper and online evaluations. The following analysis examines data from six departments of the School of Business Administration during a programmed switch from paper to online evaluations. The courses that participated in this study were divided and compared in the following manner: advanced and core classes, large and small sections, and courses taught by full-time and part-time faculty. The data was collected over a one-year period and contrasts the Spring 2008 and 2009 semesters, during which a total of 4,424 evaluations were reviewed. In addition, data on the years from 2005 to 2008 are provided as a comparison benchmark of typical responses collected when paper evaluations were used. The conclusions of this study show that while a drop in response rate did occur when the switch was made, no significant change in instructor and course ratings was observed. Furthermore, the students who did complete online evaluations provided lengthier and more numerous comments.


The Master of Administrative Science program at Fairleigh Dickinson University performed a study on courses taught by adjunct faculty. The online evaluations received a 61% response rate and the in-class evaluations received a 82.1% response rate. They found that the online evaluations received twice as many comments (counting total words) as the in-class evaluations. On the question about "materials being clearly presented" (focused on the faculty member) the variation in mean scores in online and in-class was 0.33 on a 5-point scale with online having a less-positive rating. This is a statistically significant difference. Administrators noted that both means were better than the “agree” and were not considered poor ratings.


At a southeastern university 66 courses made up of 2453 students did a comparison of response effects between paper-and-pencil and online using the same form. Half did online and half did paper-and-pencil forms. The online response rate was 47% and the traditional group was 60%. Also, 76% of the online forms provided comments compared to 50% of the traditional forms. No significant difference was found in methods.


Georgia State University College of Business ran a voluntary pilot from 2002 to 2003 using an identical online version of their paper course evaluation form in the Department of Computer Information Systems. Faculty feared an online form would yield lower scores and lower response rates. In particular, the fear was that few students would submit online evaluations, poor students
would “take revenge” on the faculty and good students wouldn’t bother. The paper form had a 67% response rate and the online form had an 82% response rate. This likely due to the fact that the CIS department had easy access to computer labs for students to take the evaluations online. Using a question on teacher effectiveness, the study found no significant difference between the methods. Good students participated in the same numbers and weaker students did fewer online evaluations.


The paper presents a short literature review comparing online evaluations with paper. The Economics department at University of Belgrade, Serbia conducted a small pilot in a course of 800 students in May of 2006. Half the students received paper evaluations in class and half were directed to complete an identical online evaluation. The paper evaluation received a 92.5% response rate and the online received a 52% response rate after an incentive was introduced. They found that nearly twice as many students filled out the open-ended question online when compared to the paper group. On the instructor-related questions they found a variation of 0.09 to 0.22 on a 10-point scale. No statistical analysis was done for significance.


In a survey of academic reference librarians in North Carolina, Matz found no significant difference in response contents between the methods used. The online form had a 33% response rate and the paper form had a 43% response rate.


Yale Law started online course evaluations in 2001 with a less than 20% response rate. The current 8-question form is run by student representatives and has a 90% response rate. Students cannot see their grades until they fill out the evaluation. Northwestern University School of Law started online course evaluations in 2004. So far they have a 68% response rate which compares to a 70-80% paper response rate. Northwestern is against using any penalties (withholding information from a student until they fill out an evaluation). The University of Denver Sturm College started online course evaluations in 2002 with a pilot of 10 courses. The pilot had an 83% response rate. Continuing into 2003 the pilot expanded to 80 courses (with an 81% response rate) and then expanded to all of their offerings (with a 64% response rate). Currently they maintain a response rate around 70%. Duke Law started online course evaluations in 2003 when their scantron machine broke and the expense of replacing was too great. They proposed a goal of 70% response rate and used the same form online. The first term averaged a 66% response rate (with 29% of the 82 courses reaching the 70% goal). In spring 2004 the average was 60% (with 30% of the 119 courses reaching the 70% goal). In fall 2004 the average was 52% (with 8% of the 93 courses reaching the 70% goal). In spring 2005, after dropping non-law students from the pool, the average was 67% (with 41% of the 117 courses reaching the 70% goal). The school is considering several penalties for failure to fill out an evaluation — withholding registration, withholding grades, or withholding free printing.
This paper reports the findings of 2 studies done at Northern Arizona State University. The first study looked at historic data from 2000 - 2002 to examine student responses to online course evaluations in 1108 course sections. This group had an average response rate of 31%. A follow-up questionnaire was sent to 50 faculty in the group to explore what strategies improved response rate. These results informed the second study on 39 online course sections and 21 sections of a required freshman face-to-face course. The second study used some basic strategies (no penalty strategies) in the implementation of the online course evaluations: 2 weeks before the end of the course the URL to evaluation was posted in the course management system, an announcement containing a statement of course evaluation value and due date was sent in a method appropriate to the class (email, online syllabus or discussion board), and a reminder email was sent 1 week before the class ended containing the URL and due date. The 39 online course sections averaged a 74% response rate and the 21 face-to-face courses averaged a 67% response rate. In addition, 11 sections of the face-to-face course used paper evaluations and received a 83% response rate. These suggestions are very similar to the emerging findings from the TLT Group’s BeTA project.

Marquette University moved from a copyrighted instrument, IAS, to their own instrument, MOCES. Because of the copyright concerns the new instrument has re-worded items that maintain the intent of the IAS items. In spring semester of 2008 a pilot was conducted in 124 course sections with 3837 students. They evaluated the effectiveness of an online approach versus paper and pencil and the software used to deliver the evaluations. Response rates online were lower in 3 of the 5 pilot departments, comparable in 1 and higher in 1 when compared to 3 semester averages of paper and pencil forms. A “power analysis” of the response rates revealed the rates were high enough of 95% confidence in the results. There was no significant difference in the mean ratings for the 4 core questions between the old IAS form and the MOCES online form.

The YFCY distributed its survey that assesses student development during the first year in college using 3 methods: online, online or paper, and paper. In a pool of 57 schools, 16 used the alternative methods of distribution. The study found no significant difference in responses between the methods. The response rate overall was 21%. The online only method response rate was 17% and the online or paper group had a 24% response rate.

The Rose-Hulman Institute of Technology piloted an online course evaluation in 2002 with a small group of faculty. Over the academic year the pilot had a 70% response rate. 77% of students preferred the online mode and faculty reacted positively to the pilot. In 2003 the entire campus adopted the online form. Over the 3 terms, the online evaluations had response rates of
86%, 78% and 67%. In 2004 the 3 terms had 75%, 71% and 67%. Historically paper evaluations had an 85 - 87% response rate. They are investigating various incentive possibilities.


Drexel University studied whether significant differences exist in student responses to course evaluations given on paper and online in 3 courses. Response rates in the 3 classes for paper and online (respectively) were 37% and 45%, 44% and 50%, 70% and 37%. In comparing students who responded to the evaluations across the 3 courses the study found that women were more likely than men to respond, students who earned higher grades were more likely to respond, and students with a higher overall GPA were more likely to respond. For two courses the online evaluations had a slightly higher average item rating. For the other course 2 significant differences were found: students doing the online evaluation were less likely to participate actively and contribute thoughtfully during class and to attend class when compared to the paper evaluation group. But the responses overall were not significantly different.


This article summarizes the scholarly research on student course evaluations in three main sections: The validity and reliability of course evaluations, Online vs. paper course evaluations, Student perceptions of course evaluations, Effects of allowing students access to course evaluation data and Recommendations for improving response rates. The three recommendations for improving response rates are to 1) make evaluation a part of the course (most effective), 2) send reminder notices and 3) offer a small incentive.
Appendix C - Tips to ESCI Online Instructors

Your department has elected to participate in the ESCI Online Pilot program for XXX quarter 20XX. All of the end of quarter evaluation survey data for your sections will be administered and collected online. The ESCI Online server's address is https://esci.id.ucsb.edu

How Does the ESCI Online System Work for Students?
The survey window for XXX 20XX will be from 12:01 AM Monday March 2 through 11:59 PM Friday March 13. Your students will receive an email invitation to participate when the system opens to accept responses. They will receive periodic reminder emails to participate until they have completed all of the surveys for which they are responsible. The reminders go out on an accelerating schedule. The email reminder schedule is included below for your convenience:

<table>
<thead>
<tr>
<th>Reminder Type</th>
<th>Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>OPEN</td>
<td>2015-03-02</td>
</tr>
<tr>
<td>EMAIL #1 (OPEN)</td>
<td>2015-03-02</td>
</tr>
<tr>
<td>FIRST REMINDER:</td>
<td>2015-03-04</td>
</tr>
<tr>
<td>SECOND REMINDER:</td>
<td>2015-03-06</td>
</tr>
<tr>
<td>THIRD REMINDER:</td>
<td>2015-03-09</td>
</tr>
<tr>
<td>FOURTH REMINDER:</td>
<td>2015-03-11</td>
</tr>
<tr>
<td>FINAL REMINDER:</td>
<td>2015-03-12</td>
</tr>
<tr>
<td>CLOSE</td>
<td>2015-03-13</td>
</tr>
</tbody>
</table>

In addition to the automated email invitation and reminders, this quarter we also have automated reminders in GauchoSpace. For students enrolled in classes and sections participating in ESCI Online, once the survey window is open, when they log into GauchoSpace they will have a reminder that will appear on their “dashboard” page. The figure below illustrates what it looks like.

The icon will be green during week one of the survey period, and will be red during week two. This additional element helped to substantially increase response rates previously. Students can click on the link to go directly to the ESCI Online site.
Below are some suggestions to help encourage students to participate more fully in the ESCI Online process. In particular, mentioning participation in class in person, and why the evaluation information is useful to you, is very important and students do find it compelling (a point that is also mentioned quite a bit in the survey results from previous Winter quarters). I also encourage you to use the link provided below to periodically monitor the response rate for your course(s), and follow up with students if the response rate is low. “Normal” response rates for the traditional paper surveys vary between 18% and 100% with a mean of 73%. Larger courses have lower response rates, typically, with the paper surveys.

**Encouraging Students to Complete ESCI Online Evaluations**

In addition to the automated email reminders that the ESCI server will be sending students, and the GauchoSpace reminder system, we encourage you to reinforce those reminders by reaching out to your students to also encourage them to participate. This approach has been one of the most successful ways of improving response rates in our pilot.

- **In-class announcement about the ESCI Online Surveys**
  - During class, instructors can emphasize the importance of completing the online surveys by describing how they are used to improve the course and teaching:
    - Valuable feedback for the instructor
    - Allows the course to be improved each quarter
    - When possible, offer concrete examples of how you have used past feedback to improve the course.
    - Constructive comments give insight into what is working and what, specifically, can be improved. Be sure to look at your course survey to verify that open-ended comments are allowed.
  
- **Send out an electronic announcement**
  Instructors can use either Instructor Announcements in GauchoSpace, or Peachmail, to reach students who may not have been present for the in-class announcement. This message could:
  - Reiterate the value of the surveys
  - Specify the dates that the survey will be available and
  - Provide the url to access the surveys (https://esci.id.ucsb.edu/)
  - Instructors may wish to send out a reminder using the GauchoSpace announcements or by e-mail about 3 days before the closing date for submission.
  - Although students will receive automated e-mail reminders to complete the ESCI evaluations, the best way to improve response rates is for the Instructor to encourage his/her students to complete the surveys personally.
Checking Student Participation
You can log into the ESCI Online server yourself and check on the response rates for your courses in real time. To do that:

- Go to: https://esci.id.ucsb.edu/
- Click on the “Login” tab
- Log in with your UCSB NetID and password
- Click on the “Instructor” tab to view current response rates for your course(s)
- Log out
- If you have a problem, please click on the “Help” tab and submit a help ticket so that we can determine what caused the problem, correct it, and then follow up with you.

Please let me know if you have any questions. Thank you for your help with this important project for the campus.

Cordially,
George Michaels
Executive Director, Instructional Development
gb@email.ucsb.edu
(805) 893-2378