BUILT ENVIRONMENT SRI LANKA 2026
VOL. 16, ISSUE 02, NO. # Page Numbering
https://NSF WILL PROVIDE THE DOI AFTER COMPLETING THE SUBMISSION

Built Environment Sri Lanka Template

(Download the template and write your paper.)

Claudia J. Stanny[1] and James E. Arruda[2]

[1] Centre for University Teaching, Learning, and Assessment, University of West Florida

[2] Department of Psychology, University of West Florida

ARTICLE HISTORY
Received 01 February 2026
Accepted 02 May 2026

KEYWORDS
Architecture
Urban Design

You may include a graphical abstract to help communicate your research more effectively.

ABSTRACT

When institutions administer student evaluations of teaching (SETs) online, response rates are lower relative to paper-based administration. We analyzed average SET scores from 364 courses taught during the fall term in 3 consecutive years to determine whether administering SET forms online for all courses in the 3rd year changed the response rate or the average SET score. To control for instructor characteristics, we based the data analysis on courses for which the same instructor taught the course in each of three successive fall terms. Response rates for face-to-face classes declined when SET administration occurred only online. Although average SET scores were reliably lower in Year 3 than in the previous 2 years, the magnitude of this change was minimal (0.11 on a five-item Likert-like scale). We discuss practical implications of these findings for interpretation of SETs and the role of SETs in the evaluation of teaching quality.

1. Introduction

Student ratings and evaluations of instruction have a long history as sources of information about teaching quality (Berk, 2013). Student evaluations of teaching (SETs) often play a significant role in high-stakes decisions about hiring, promotion, tenure, and teaching awards. As a result, researchers

have examined the psychometric properties of SETs and the possible impact of variables such as race, gender, age, course difficulty, and grading practices on average student ratings (Griffin et al., 2014; Nulty, 2008; Spooren et al., 2013). They have also examined how decision makers evaluate SET scores (Boysen, 2015a, 2015b; Boysen et al., 2014; Dewar, 2011). In the last 20 years, considerable attention has been directed toward the consequences of administering SETs online (Morrison, 2011; Stowell et al., 2012) because low response rates may have implications for how decision makers should interpret SETs.

1.1. Online Administration of Student Evaluations

Administering SETs online creates multiple benefits. Online administration enables instructors to devote more class time to instruction (vs. administering paper-based forms) and can improve the integrity of the process. Students who are not pressed for time in class are more likely to reflect on their answers and write more detailed comments (Morrison, 2011; Stowell et al., 2012; Venette et al., 2010). Because electronic aggregation of responses bypasses the time-consuming task of transcribing comments (sometimes written in challenging handwriting), instructors can receive summary data and verbatim comments shortly after the close of the term instead of weeks or months into the following term.

Despite the many benefits of online administration, instructors and students have expressed concerns about online administration of SETs. Students have expressed concern that their responses are not confidential when they must use their student identification number to log into the system (Dommeyer et al., 2002). However, breaches of confidentiality can occur even with paper-based administration. For example, an instructor might recognize student handwriting (one reason some students do not write comments on paper-based forms), or an instructor might remain present during SET administration (Avery et al., 2006).

In-class, paper-based administration creates social expectations that might motivate students to complete SETs. In contrast, students who are concerned about confidentiality or do not understand how instructors and institutions use SET findings to improve teaching might ignore requests to complete an online SET (Dommeyer et al., 2002). Instructors in turn worry that low response rates will reduce the validity of the findings if students who do not complete an SET differ in significant ways from students who do (Stowell et al., 2012). For example, students who do not attend class regularly often miss class on the day that SETs are administered. However, all students (including non-attending students) can complete the forms when

they are administered online. Faculty also fear that SET findings based on a low-response sample will be dominated by students in extreme categories (e.g., students with grudges, students with extremely favorable attitudes), who may be particularly motivated to complete online SETs, and therefore that SET findings will inadequately represent the voice of average students (Reiner & Arnold, 2010).

1.1.1. Effects of Format on Response Rates and Student Evaluation Scores

The potential for biased SET findings associated with low response rates has been examined in the published literature. In findings that run contrary to faculty fears that online SETs might be dominated by low-performing students, Avery et al. (2006) found that students with higher grade-point averages (GPAs) were more likely to complete online evaluations. Likewise, Jaquett et al. (2017) reported that students who had positive experiences in their classes (including receiving the grade they expected to earn) were more likely to submit course evaluations.

Institutions can expect lower response rates when they administer SETs online (Avery et al., 2006; Dommeyer et al., 2002; Morrison, 2011; Nulty, 2008; Reiner & Arnold, 2010; Stowell et al., 2012; Venette et al., 2010). However, most researchers have found that the mean SET rating does not change significantly when they compare SETs administered on paper with those completed online. These findings have been replicated in multiple settings using a variety of research methods (Avery et al., 2006; Dommeyer et al., 2004; Morrison, 2011; Stowell et al., 2012; Venette et al., 2010).

Exceptions to this pattern of minimal or nonsignificant differences in average SET scores appeared in Nowell et al. (2010) and Morrison (2011), who examined a sample of 29 business courses. Both studies reported lower average scores when SETs were administered online. However, they also found that SET scores for individual items varied more within an instructor when SETs were administered online versus on paper. Students who completed SETs on paper tended to record the same response for all questions, whereas students who completed the forms online tended to respond differently to different questions. Both research groups argued that scores obtained online might not be directly comparable to scores obtained through paper-based forms. They advised that institutions administer SETs entirely online or entirely on paper to ensure consistent, comparable evaluations across faculty.

Each university presents a unique environment and culture that could influence how seriously students take SETs and how they respond to

decisions to administer SETs online. Although a few large-scale studies of the impact of online administration exist (Reiner & Arnold, 2010; Risquez et al., 2015), a local replication answers questions about characteristics unique to that institution and generates evidence about the generalizability of existing findings.

1.2. Purpose of the Present Study

In the present study, we examined patterns of responses for online and paper-based SET scores at a mid-sized, regional, comprehensive university in the United States. We posed two questions: First, does the response rate or the average SET score change when an institution administers SET forms online instead of on paper? Second, what is the minimal response rate required to produce stable average SET scores for an instructor? Whereas much earlier research relied on small samples often limited to a single academic department, we gathered SET data on a large sample of courses (N = 364) that included instructors from all colleges and all course levels over 3 years. We controlled individual differences in instructors by limiting the sample to courses taught by the same instructor in all 3 years. The university offers nearly 30% of course sections online in any given term, and these courses have always administered online SETs. This allowed us to examine the combined effects of changing the method of delivery for SETs (paper-based to online) for traditional classes and changing from a mixed method of administering SETs (paper for traditional classes and online for online classes in the first 2 years of data gathered) to uniform use of online forms for all classes in the final year of data collection.

2. Acknowledgement

I would like to show my gratitude to UN-Habitat. With the help of UN-Habitat, this research managed to find and measure thermal data of actual building models. Those buildings were made of different building materials and yet followed the same building layout.

3. Disclosure statement

No potential conflict of interest was reported by the authors.

4. Funding

This work was supported by the University of Moratuwa Senate Research Committee.

5. ORCID

6. References

Avery, R. J., Bryant, W. K., Mathios, A., Kang, H., & Bell, D. (2006). Electronic course evaluations: Does an online delivery system influence student evaluations? *The Journal of Economic Education*, *37*(1), 21–37. https://doi.org/10.3200/JECE.37.1.21-37

Berk, R. A. (2012). Top 20 strategies to increase the online response rates of student rating scales. *International Journal of Technology in Teaching and Learning*, *8*(2), 98–107.

Berk, R. A. (2013). *Top 10 flashpoints in student ratings and the evaluation of teaching*. Stylus.

Boysen, G. A. (2015a). Preventing the overinterpretation of small mean differences in student evaluations of teaching: An evaluation of warning effectiveness. *Scholarship of Teaching and Learning in Psychology*, *1*(4), 269–282. https://doi.org/10.1037/stl0000042

Boysen, G. A. (2015b). Significant interpretation of small mean differences in student evaluations of teaching despite explicit warning to avoid overinterpretation. *Scholarship of Teaching and Learning in Psychology*, *1*(2), 150–162. https://doi.org/10.1037/stl0000017

Boysen, G. A., Kelly, T. J., Raesly, H. N., & Casner, R. W. (2014). The (mis)interpretation of teaching evaluations by college faculty and administrators. *Assessment & Evaluation in Higher Education*, *39*(6), 641–656. https://doi.org/10.1080/02602938.2013.860950

Buller, J. L. (2012). *Best practices in faculty evaluation: A practical guide for academic leaders*. Jossey-Bass.

Dewar, J. M. (2011). Helping stakeholders understand the limitations of SRT data: Are we doing enough? *Journal of Faculty Development*, *25*(3), 40–44.

Dommeyer, C. J., Baum, P., & Hanna, R. W. (2002). College students' attitudes toward methods of collecting teaching evaluations: In-class versus on-line. *Journal of Education for Business*, *78*(1), 11–15. https://doi.org/10.1080/08832320209599691

Dommeyer, C. J., Baum, P., Hanna, R. W., & Chapman, K. S. (2004). Gathering faculty teaching evaluations by in-class and online surveys: Their effects on response rates and evaluations. *Assessment & Evaluation in Higher Education*, *29*(5), 611–623. https://doi.org/10.1080/02602930410001689171

Feistauer, D., & Richter, T. (2016). How reliable are students' evaluations of teaching quality? A variance components approach. *Assessment & Evaluation in Higher Education*, *42*(8), 1263–1279. https://doi.org/10.1080/02602938.2016.1261083

Gilovich, T., Griffin, D., & Kahneman, D. (Eds.). (2002). *Heuristics and biases: The psychology of intuitive judgment*. Cambridge University Press. https://doi.org/10.1017/CBO9780511808098

Griffin, T. J., Hilton, J., III, Plummer, K., & Barret, D. (2014). Correlation between grade point averages and student evaluation of teaching scores: Taking a closer look. *Assessment & Evaluation in Higher Education*, *39*(3), 339–348. https://doi.org/10.1080/02602938.2013.831809

Jaquett, C. M., VanMaaren, V. G., & Williams, R. L. (2016). The effect of extra-credit incentives on student submission of end-of-course evaluations. *Scholarship of Teaching and Learning in Psychology*, *2*(1), 49–61. https://doi.org/10.1037/stl0000052

Jaquett, C. M., VanMaaren, V. G., & Williams, R. L. (2017). Course factors that motivate students to submit end-of-course evaluations. *Innovative Higher Education*, *42*(1), 19–31. https://doi.org/10.1007/s10755-016-9368-5

Morrison, R. (2011). A comparison of online versus traditional student end-of-course critiques in resident courses. *Assessment & Evaluation in Higher Education*, *36*(6), 627–641. https://doi.org/10.1080/02602931003632399

Nowell, C., Gale, L. R., & Handley, B. (2010). Assessing faculty performance using student evaluations of teaching in an uncontrolled setting. *Assessment & Evaluation in Higher Education*, *35*(4), 463–475. https://doi.org/10.1080/02602930902862875

Nulty, D. D. (2008). The adequacy of response rates to online and paper surveys: What can be done? *Assessment & Evaluation in Higher Education*, *33*(3), 301–314. https://doi.org/10.1080/02602930701293231

Palmer, M. S., Bach, D. J., & Streifer, A. C. (2014). Measuring the promise: A learning-focused syllabus rubric. *To Improve the Academy: A Journal of Educational Development*, *33*(1), 14–36. https://doi.org/10.1002/tia2.20004

Reiner, C. M., & Arnold, K. E. (2010). Online course evaluation: Student and instructor perspectives and assessment potential. *Assessment Update*, *22*(2), 8–10. https://doi.org/10.1002/au.222

Risquez, A., Vaughan, E., & Murphy, M. (2015). Online student evaluations of teaching: What are we sacrificing for the affordances of technology? *Assessment & Evaluation in Higher Education*, *40*(1), 210–234. https://doi.org/10.1080/02602938.2014.890695

Spooren, P., Brockx, B., & Mortelmans, D. (2013). On the validity of student evaluation of teaching: The state of the art. *Review of Educational Research*, *83*(4), 598–642. https://doi.org/10.3102/0034654313496870

Stanny, C. J., Gonzalez, M., & McGowan, B. (2015). Assessing the culture of teaching and learning through a syllabus review. *Assessment & Evaluation in Higher Education*, *40*(7), 898–913. https://doi.org/10.1080/02602938.2014.956684

Stark, P. B., & Freishtat, R. (2014). An evaluation of course evaluations. *ScienceOpen Research*. https://doi.org/10.14293/S2199-1006.1.SOR-EDU.AOFRQA.v1

Stowell, J. R., Addison, W. E., & Smith, J. L. (2012). Comparison of online and classroom-based student evaluations of instruction. *Assessment & Evaluation in Higher Education*, *37*(4), 465–473. https://doi.org/10.1080/02602938.2010.545869

Tversky, A., & Kahneman, D. (1971). Belief in the law of small numbers. *Psychological Bulletin*, *76*(2), 105–110. https://doi.org/10.1037/h0031322

Uttl, B., White, C. A., & Gonzalez, D. W. (2017). Meta-analysis of faculty's teaching effectiveness: Student evaluation of teaching ratings and student learning are not related. *Studies in Educational Evaluation*, *54*, 22–42. https://doi.org/10.1016/j.stueduc.2016.08.007

Venette, S., Sellnow, D., & McIntyre, K. (2010). Charting new territory: Assessing the online frontier of student ratings of instruction. *Assessment & Evaluation in Higher Education*, *35*(1), 101–115. https://doi.org/10.1080/02602930802618336

Webb, E. J., Campbell, D. T., Schwartz, R. D., & Sechrest, L. (1966). *Unobtrusive measures: Nonreactive research in the social sciences*. Rand McNally.