Research on Student Ratings of Instruction: Implications for Teachers and Professional Development Programs

Dennis Koyama, Ph.D.
Faculty of Liberal Arts, Sophia University, Japan
Email: dkoyama@sophia.ac.jp

Abstract

Evaluations of teaching effectiveness have taken many forms over the years, but none have been as persistent or commonplace as student ratings of instruction (SRI). SRIs have become a fundamental component of evaluating faculty effectiveness in higher education. Support for SRIs comes from end-users of the data who believe that students are uniquely positioned to evaluate faculty based on their experiences and perceptions of the instruction they received. Pragmatically, institutions tend to rely on SRI results for teacher evaluations because they reason that students learn more from faculty who are highly rated by students. However, to what degree is this enthusiasm warranted? Are SRIs reliable, valid, or trustworthy at all?

The main goals of this chapter are to present an overview of SRI research, explain ways of preparing students for SRIs (both formative and summative), and present methods for teachers to use when examining the SRI data. To these ends, this chapter will briefly review the SRI research, including evidence for the value of SRI data despite commonly held misconceptions about the possible influence of factors such as class size, GPA, gender, and professor rank. Attention is then given to understanding how to improve responses to questions that tap constructs students are unlikely to be readily able to respond to, such as "Did this course improve your critical thinking skills?" and to general agreement questions about learning, such as "The pacing of the materials was appropriate." Techniques for interpreting constructed responses from students, such as "Stop lecturing!" are also provided. Finally, the paper moves on to highlighting the connection between collecting and acting on formative classroom surveys that support

positive transfer to end-of-term SRIs and offers methods to analyze SRIs individually as well as outlining an approach to teacher development with SRI data and teacher-centered consultations by PD programs.

Keywords: student feedback, college teaching, professional development, teacher effectiveness

Introduction

> *Faculty evaluation and development cannot be considered separately. ... [[E]]valuation without development is punitive, and development without evaluation is guesswork* (Theall, 2017, p.91).

For many years, education systems around the world have embarked on and rekindled missions to reform educational practices with the aim of improving student learning (Borko, 2004). While a myriad of changes (e.g., improved facilities, smaller class sizes) might yield the desired improvements, researchers have identified teachers as one of the most influential resources that can impact students' academic success (Burroughs, et al., 2019). From this perspective, the magnitude of any positive change in student learning will ultimately depend on the synergy of teachers seeking to improve their teaching (Fullan & Miles, 1992) and the support and professional development provided by their institutions (Ball & Cohen, 1999; Borko, 2004).

Historically, universities have provided information to teachers on their teaching effectiveness through one or more of several channels of feedback, such as peer observations by colleagues, self-evaluations, retrospective assessments from alumni, and student ratings of instruction (Hobson& Talbot, 2001). This paper focuses on the latter and how student ratings of instruction[1] (SRI) can help inform teacher training and professional development (PD) programs with the goal of improving teaching practice and learning in higher education.

---

[1] I use the term *ratings* and not *evaluations* to draw an important distinction between the data (i.e., the students' ratings) and how the ratings are used (e.g., to evaluate instructional effectiveness for promotion, to support program accreditation applications, to inform curriculum revisions, to encourage professional development, et cetera).

This is a limited preview of the chapter.

To read the full-text chapter, get access by purchasing this chapter or consider buying the complete book. If your library has subscription to EBSCOhost, the chapter including other chapters of the book can be accessed through your library.

how you interact with and experience the SRI process. In the long run, improving SRIs will require specific and enduring efforts by PD programs and teachers who engage in reflective practice that focuses on actively learning about their teaching.

References

Abrami, P. C. (2001). Improving judgments about teaching effectiveness using teacher rating forms. In M. Theall, P. C. Abrami, & L. A. Mets (Eds.), *The student ratings debate: Are they valid? How can we best use them?* (pp. 59–87). Jossey-Bass. https://doi.org/10.1002/ir.4

Abrami, P. C., d'Apollonia, S., & Rosenfield, S. (2007). The dimensionality of student ratings of instruction: What we know and what we do not. In R. P. Perry, and J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective* (pp. 385-456). Springer.https://doi.org/10.1007/1-4020-5742-3_10

Al-Abbadi, I., Alkhateeb, F., Khanfar, N., Mujtaba, B., & Latif, D. (2009). Pharmacy students' perceptions of the teaching evaluation process in Jordan. *Education, Business and Society: Contemporary Middle Eastern Issues*, *2*(3), 181-190. https://doi.org/10.1108/17537980910981750

Al-Issa, A., &Sulieman, H. (2007). Student evaluations of teaching: Perceptions and biasing factors.*Quality Assurance in Education*,*15*(3), 302-317.https://doi.org/10.1108/09684880710773183

Alsmadi, A. (2005). Assessing the quality of students'ratings of faculty members at Mu'tah University. *Social Behavior and Personality: An International Journal,* 33(2), 183-188.https://doi.org/10.2224/sbp.2005.33.2.183

d'Apollonia, S., & Abrami, P. C. (1997). Navigating student ratings of instruction. *American Psychologist, 52*, 1198–1208. https://doi.org/10.1037/0003-066X.52.11.1198

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. In L. Darling-Hammond and G. Sykes (Eds.), *Teaching as the learning profession* (pp. 3-31). Jossey-Bass.

Beerens, D. (2000). *Evaluating teachers for growth creating a culture of motivation and learning*. Corwin.

Benton, S. L., & Cashin, W. E. (2012). *Idea paper# 50 student ratings of teaching: A summary of research and literature.*The IDEA Center.

Benton, S. L., & Cashin, W. E. (2014). Student ratings of instruction in college and university courses. In M. B. Paulsen (Ed.) Higher education: Handbook of theory

and research, Vol. 29, (pp. 279-326). Springer.https://doi.org/10.1007/978-94-017-8005-6_7

Beran, T., Violato, C., Kline, D., & Frideres, J. (2005). The utility of student ratings of instruction for students, faculty, and administrators: A "consequential validity" study. *Canadian Journal of Higher Education*, *35*, 49–70.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8), 3-15.https://doi.org/10.3102/0013189X033008003

Boysen, G. A. (2015a). 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. (2015b). Uses and misuses of student evaluations of teaching: The interpretation of differences in teaching evaluation means irrespective of statistical information. *Teaching of Psychology, 42,* 109–118. http://dx.doi.org/10.1177/00986 28315569922

Boysen, G. A. (2016). Using student evaluations to improve teaching: Evidence-based recommendations. *Scholarship of Teaching and Learning in Psychology*, *2*(4), 273–284.https://doi.org/10.1037/stl0000069

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,* 641–656.http://dx.doi.org/10.1080/02602938 .2013.860950

Braskamp, L. A., & Ory, J. C. (1994). *Assessing faculty work: Enhancing individual and institutional performance*. Jossey-Bass.

Braskamp, L. A., Ory, J. C., & Pieper, D. M. (1981). Student written comments: Dimensions of instructional quality. *Journal of Educational Psychology, 73*, 65–70. https://doi.org/10.1037/0022-0663.73.1.65

Burdsal, C. A., & Harrison, P. D. (2008). Further evidence supporting the validity of both a multi- dimensional profile and an overall evaluation of teaching effectiveness. *Assessment and Evaluation in Higher Education, 33*, 567–576. https://doi.org/10.1080/02602930701699049

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., & Schmidt, W. (2019). *Teaching for excellence and equity: Analyzing teacher characteristics, behaviors and student outcomes with TIMSS*. Springer.https://doi.org/10.1007/978-3-030-16151-4

Canelos, J. (1985). Teaching and course evaluation procedures: A literature review of current research. *Journal of Instructional Psychology*, *12*(4), 187-195

Centra, J. A. (1993). *Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness*. Jossey-Bass.

Centra, J. A. (2009). *Differences in responses to the student instructional report: Is it bias?*. Educational Testing Service.

Centra . J. A. & Gaubatz, N. B. (2000). Is there gender bias in student evaluations of teaching?. *Journal of Higher Education*, *71*(1), 17–33.https://doi.org/10.2307/2649280

Chen, Y., & Hoshower, L. B. (2003). Student evaluation of teaching effectiveness: An assessment of student perception and motivation. *Assessment & Evaluation in Higher Education*, *28*(1), 71-88.https://doi.org/10.1080/02602930301683

Clayson, D. E. (2009). Student evaluation of teaching: Are they related to what students learn? *Journal of Marketing Education, 31*, 16–30.https://doi.org/10.1177/0273475308324086

Cohen, P. A. (1980). Effectiveness of student-rating feedback for improving college instruction: A meta-analysis of findings. *Research in Higher Education, 13*, 321–341.https://doi.org/10.1007/BF00976252

Cohen, P. A. (1981). Student ratings of instruction and student achievement: A meta-analysis of multisection validity studies. *Review of Educational Research, 51*, 281–309. https://doi.org/10.3102/00346543051003281

Davis, B. G. (2009). *Tools for Teaching*. John Wiley & Sons.

Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed-mode surveys: The tailored design method*. John Wiley & Sons.

Filene, P. (2005). *The joy of teaching: A practical guide for new college instructors*. University of North Carolina Press.

Feldman, K. A. (1989). The association between student ratings of specific instructional dimensions and student achievement: Refining and extending the synthesis of data from multisection validity studies. *Research in Higher Education, 30*, 583–645. https://doi.org/10.1007/BF00992392

Feldman, K. A. (1993). College students' views of male and female college teachers: Part II–Evidence from students' evaluations of their classroom teachers. *Research in Higher Education, 34*, 151–211. https://doi.org/10.1007/BF00992216

Franklin, J. (2001). Interpreting the numbers: Using a narrative to help others read student evaluations of your teaching accurately. *New Directions for Teaching and Learning, 87*, 85-100.https://doi.org/10.1002/tl.10001

Fullan, M. G., & Miles, M. B. (1992). Getting reform right: What works and what doesn't. *Phi Delta Kappan, 73*(10), 745-752.

Hativa, N. (2013). *Student ratings of instruction: Recognizing effective teaching*. Oron Publications.

Hativa, N. (2019). *Student ratings of instruction: Can we trust them*. Oron Publications.

Hill, H., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review, 83*(2), 371-384.https://doi.org/10.17763/haer.83.2.d11511403715u376

Hoyt, D. P., & Lee, E. (2002a). *Technical report no. 12: Basic data for the revised IDEA system*.The IDEA Center.

Hoyt, D. P., & Lee, E. J. (2002b). *Technical report #13: Disciplinary differences in student ratings*. IDEA Center.

Hobson, S. M. & Talbot, D. M. (2001). Understanding student evaluations: What all faculty should know, *College Teaching*, *49*(1), 26-31, https://doi.org/10.1080/87567550109595842

Knol, M. (2013). *Improving university lectures with feedback and consultation*. Ipskamp Drukkers B.V.

Lewis, K. G. (2001). Making sense of student written comments. *New Directions for Teaching and Learning*, *2001*(87), pp. 25-32.https://doi.org/10.1002/tl.25

Liddle, B. J. (1997). Coming out in class: Disclosure of sexual orientation and teaching evaluations. *Teaching of Psychology*, *24*(1), 32-35. https://doi.org/10.1207/s15328023top2401_6

Marsh, H. W. (1987). Students' evaluations of university teaching: Research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, *11*(3), 253-388.https://doi.org/10.1016/0883-0355(87)90001-2

Marsh, H. W. (2001). Distinguishing between good (useful) and bad workloads on student evaluations of teaching. *American Educational Research Journal, 38*, 183–212.https://doi.org/10.3102/00028312038001183

Marsh, H. W. (2007). Students' evaluations of university teaching: Dimensionality, reliability, validity, potential biases and usefulness. In R. P. Perry, and J. C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*, (pp. 319-383). Springer.https://doi.org/10.1007/1-4020-5742-3_9

Marsh, H. W., & Roche, L. A. (2000). Effects of grading leniency and low workload on students' evaluations of teaching: Popular myth, bias, validity, and innocent bystanders. *Journal of Educational Psychology, 92*, 202–22. https://doi.org/10.1037/0022-0663.92.1.202

Marzano, R. & Toth, M. (2013). *Teacher evaluation that makes a difference*. Association for Supervision and Curriculum Development.

McKeachie, W. J. (1997). Student ratings: The validity of use. *American Psychologist, 52*, 1218–1225.https://doi.org/10.1037/0003-066X.52.11.1218

Mercer, J. (2005). Challenging appraisal orthodoxies: Teacher evaluation and professional development in the United Arab Emirates. *Journal of Personnel Evaluation in Education*, *18*(4), 273.https://doi.org/10.1007/s11092-007-9024-9

Nasser, F., & Fresko, B. (2002). Faculty views of student evaluation of college teaching. *Assessment & Evaluation in Higher Education*, *27*, 187–198. http://dx.doi.org/10.1080/02602930220128751

Nguyen, T. H. (2014). Student ratings in Vietnam higher education: How are instructors' reactions?.*International Journal of Innovative Management, Information &*

Production,*5*(3), 99–109. http://www.ismeip.org/IJIMIP/contents/imip1453/11.pdf.

Paulsen, M. B., & Perna, L. W. (Eds.). (2016). *Higher education: Handbook of theory and research* (Vol. 29). Springer.

Penny, A. R. & Coe, R. (2004). Effectiveness of consultation on student ratings feedback: A meta-analysis. *Review of Educational Research, 74*(2), 215-253.

Perry, R. P., & Smart, J. C. (Eds.). (1997). *Effective teaching in higher education: Research and practice.* Agathon Press.https://doi.org/10.1007/978-3-319-26829-3

Perry, R. P., & Smart, J. C. (Eds.). (2007). *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Springer. https://doi.org/10.1007/1-4020-5742-3

Schmelkin, L. P., Spencer, K. J., & Gellman, E. S. (1997). Faculty perspectives on course and teacher evaluations. *Research in Higher Education*, *38*, 575–592.https://doi.org/10.1023/A:1024996413417

Schulze, E., & Tomal, A. (2006). The chilly classroom: Beyond gender. *College Teaching, 54*(3), 263–270.https://doi.org/10.3200/CTCH.54.3.263-270

Scriven, M. (1994). Student ratings offer useful input to teacher evaluations, *Practical Assessment, Research, and Evaluation*, *4*(1), Paper 7. https://doi.org/10.7275/1jfr-et33

Stewart, C. (2014). Transforming professional development to professional learning. *Journal of Adult Education*, *43*(1), 28-33.

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

Sudkamp, A., Kaiser, J., & Moller, J. (2012). Accuracy of teachers' judgments of students' academic achievement: A meta-analysis. *Journal of Educational Psychology, 104*, 743–762.https://doi.org/10.1037/a0027627

Susanlh, Z. B., & Kaytaz, M. (2015). Determinants of student evaluation of teaching: Evidence from Turkey.*Journal of Business & Economic Policy*, *2*(1), 121-134.

Theall, M. (2017). MVP and faculty evaluation. *New Directions for Teaching and Learning*, *2017*(152). 91-98. https://doi.org/10.1002/tl.20271

Theall, M., & Franklin, J. (Eds.). (1990). *Student ratings of instruction: Issues for improving practice (No. 43)*. Jossey-Bass.https://doi.org/10.1002/tl.37219904308

Tran, T. T. T., & Do, T. X. (2020). Student evaluation of teaching: Do teacher age, seniority, gender, and qualification matter?. *Educational Studies*, 1-28.https://doi.org/10.1080/03055698.2020.1771545

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

Wachtel , H. K. (1998). Student evaluations of college teaching effectiveness: A brief review. *Assessment and Evaluation in Higher Education, 23*(2), 191–211.https://doi.org/10.1080/0260293980230207