

Using Multiple Sources of Data to Evaluate Teaching Effectiveness

Arreola, R.A. (2007). *Developing a comprehensive faculty evaluation system: A guide to designing, building, and operating large-scale faculty evaluation systems* (3rd ed.). Bolton, MA: Anker Publishing.

Berk, R.A. (2005). Survey of 12 strategies to measure teaching effectiveness. *International Journal of Teaching and Learning in Higher Education*, 17 (1), 48-62. Retrieved from: <http://www.isetl.org/ijtlhe/pdf/IJTLHE8.pdf>

Berk, R.A. (2006). *Thirteen strategies to measure college teaching: A consumer's guide to rating scale construction, assessment, and decision-making for faculty, administrators, and clinicians*. Sterling: Stylus Publishing Inc.

Chism, N.V.N (2007). *Peer review of teaching: A sourcebook* (2nd ed.). Bolton, MA: Anker Publishing.

Knapper, C., & Cranton, P. (Eds.). (2001). *Fresh approaches to the evaluation of teaching: New directions for Teaching and learning*, No. 88. San Francisco: Jossey-Bass. Retrieved from: <http://onlinelibrary.wiley.com/doi/10.1002/tl.v2001:88/issuetoc>

Mitchell, J.V., Wise, S.L., & Blake, B.S. (Eds.). (1990). *Assessment of teaching: Purposes, practices, and implications for the profession*. Hillsdale, NJ: Erlbaum.

Onwuegbuzie, A.J., Daniel, L.G., & Collins, K.M. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality and Quantity*, 43(2), 197-209. Retrieved from: <http://www.springerlink.com/content/l031k77753182234/>

Seldin, P. (1993). The use and abuse of student ratings of professors. *Chronicle of Higher Education*, 39, 40.

Seldin, P. (1999). *Changing practices in evaluating teaching: A practical guide to improved faculty performance and promotion/tenure decisions*. Bolton, MA: Anker.

Sproule, R. (2002). The underdetermination of instructor performance by data from the student evaluation of teaching. *Economics of Education Review*, 21 (3), 287-295. Retrieved from: <http://www.sciencedirect.com/science/article/pii/S0272775701000255>

Winchester, T.M., & Winchester, M. (2011). Exploring the impact of faculty reflection on weekly student evaluations of teaching. *International Journal for Academic Development*, 16 (2), 119-131. Retrieved from: <http://www.tandfonline.com/doi/abs/10.1080/1360144X.2011.568679>

Wongsurawat, W. (2011). What's a comment worth? How to better understand student evaluations of teaching. *Quality Assurance in Education: An International Perspective*, 19 (1), 67-83. Retrieved from: <http://www.emeraldinsight.com/journals.htm?issn=0968-4883&volume=19&issue=1&articleid=1905638&show=abstract>

Reliability and Validity of SETs

Abrami, P.C., d'Apollonia, S., & Cohen, P.A. (1990). Validity of student ratings of instruction: what we know and what we do not. *Journal of Educational Psychology*, 82(2), 385-486.

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*. Dordrecht, Netherlands: Springer. Retrieved from: <http://www.springerlink.com/content/m2j35292l420t007/>

Ackerman, D., Gross, B.L., & Vigneron, F. (2009). Peer observation reports and student evaluations of teaching: who are the experts? *Alberta Journal of Educational Research*, 55 (1), 18-39. Retrieved from: <http://ajer.synergiesprairies.ca/ajer/index.php/ajer/article/view/679>

Algozzine, B., Beattie, J., Bray, M., Flowers, C., Gretes, J., Howley, L., Mohanty, G., & Spooner, F. (2004). Student evaluation of college teaching: a practice in search of principles. *College Teaching*, 52, 134-141. Retrieved from: <http://www.jstor.org/stable/27559201>

Brown, M.J, Baillie, M., & Fraser, S. (2009). Rating RateMyProfessors.com: a comparison of online and official student evaluations of teaching. *College Teaching*, 57(2), 89-92. Retrieved from: <http://www.tandfonline.com/doi/abs/10.3200/CTCH.57.2.89-92>

Cashin, W.E. (1995). *Student ratings of teaching: The research revisited, IDEA Paper, no. 32*. Manhattan, KS: Center for Faculty Evaluation and Faculty Development, Kansas State University. Retrieved from:http://www.theideacenter.org/sites/default/files/Idea_Paper_32.pdf

Catano, V.M., & Harvey, S. (2011). Student perception of teaching effectiveness: development and validation of the Evaluation of Teaching Competencies Scale (ETCS). *Assessment & Evaluation in Higher Education*, 36 (6), 701-711.

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

Cohen, P.A. (1981). Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. *Review of Educational Research*, 51 (3), 281-309. Retrieved from: <http://www.jstor.org/stable/1170209>

Darby, J.A. (2007). Evaluating course evaluations: the need to establish what is being measured. *Assessment & Evaluation in Higher Education*, 32 (4), 441-455.

Denson, N., Loveday, T., & Dalton, H. (2010). Student evaluations of courses: what predicts satisfaction? *Higher Education Research and Development*, 29 (4), 339-356.

Eiszler, D.F. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education*, 43 (4), 483-502. Retrieved from:<http://www.springerlink.com/content/a0c5ct6hm1xxgh3c/>

Emery, C.R., Kramer, T.R., & Tian, R.G. (2003). Return to academic standards: a critique of students' evaluations of teaching effectiveness. *Quality Assurance in Education: An International Perspective*, 11 (1), 37-47. Retrieved from:<http://www.emeraldinsight.com/journals.htm?issn=0968-4883&volume=11&issue=1&articleid=839715&show=abstract>

Feldman, K.A. (1977). Consistency and variability among college students in rating their teachers and courses: a review and analysis. *Research in Higher Education*, 6 (3), 223–274. Retrieved from: <http://www.springerlink.com/content/mv026l431rx13590/>

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 (6), 583-645. Retrieved from: <http://www.springerlink.com/content/p6515mu760856160/>

Feldman, K.A. (1989). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. *Research in Higher Education*, 30 (2), 137-194. Retrieved from: <http://www.springerlink.com/content/q171194325658756/>

Feldman, K.A. (2007). Identifying exemplary teachers and teaching: evidence from student ratings. In R.P. Perry and J.C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, Netherlands: Springer. Retrieved from: <http://www.springerlink.com/content/r2727vn518394p70/>

Frey, P.W., Leonard, D.W., & Beatty, W.W. (1975). Student ratings of instruction: validation research. *American Educational Research Journal*, 12 (4), 435-447. Retrieved from: <http://www.jstor.org/stable/1162743>

Galbraith, C.S., Merrill, G.B., & Kline, D.M. (2012). Are student evaluations of teaching effectiveness valid for measuring student learning outcomes in business related classes? A neural network and Bayesian analyses. *Research in Higher Education*, 53(3), 353-374. Retrieved from: <http://www.springerlink.com/content/2058756205016652/>

Greenwald, A.G. (1997). Validity concerns and usefulness of student ratings of instruction. *The American Psychologist*, 52 (11), 1182-1186.

Gregory, K.M. (2012). How undergraduates perceive their professors: a corpus analysis of Rate My Professor. *Journal of Educational Technology Systems*, 40 (2), 169-193.

Hills, S.B., Naegle, N., & Bartkus, K.R. (2009). How important are items on a student evaluation? A study of item salience. *Journal of Education for Business*, 84 (5), 297-303. Retrieved from: <http://www.tandfonline.com/doi/abs/10.3200/JOEB.84.5.297-303>

Jackson, D. L., Teal, C. R., Raines, S. J., Nansel, T. R., Force, R. C., & Burdsal, C. A. (1999). The dimensions of students' perceptions of teaching effectiveness. *Educational and Psychological Measurement*, 59 (4), 580-596. Retrieved from:<http://epm.sagepub.com/content/59/4/580.abstract>

Kember, D., & Leung, D. Y. P. (2008). Establishing the validity and reliability of course evaluation questionnaires. *Assessment & Evaluation in Higher Education*, 33 (4), 341-353.

Kozub, R.M. (2008). Student evaluations of faculty: concerns and possible solutions. *Journal of College Teaching & Learning*, 5 (11), 35-40.

Kulik, J.A. (2001). Student ratings: validity, utility, and controversy. *New Directions for Institutional Research*, 109, 9-25. Retrieved from:<http://onlinelibrary.wiley.com/doi/10.1002/ir.1/abstract>

Langbein, L.I. (1994). The validity of student evaluations of teaching. *PS: Political Science & Politics*, 27 (3), 545-553. Retrieved from: <http://www.jstor.org/stable/420225>

Legg, A.M. & Wilson, J.H. (2012). RateMyProfessors.com offers biased evaluations. *Assessment & Evaluation in Higher Education*, 37 (1), 89-97.

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*. Dordrecht, Netherlands: Springer. Retrieved from:<http://www.springerlink.com/content/g7675641r0101761/>

Marsh, H.W., & Dunkin, M.J. (1997). Students' evaluations of university teaching: a multidimensional perspective. In R. Perry and J. Smart (Eds.), *Effective teaching in higher education: Research and practice*. New York: Agathon Press.

McNatt, D.B. (2010). Negative reputation and biased student evaluations of teaching: longitudinal results from a naturally occurring experiment. *The Academy of Management Learning and Education*, 9 (2), 225-242.

- Morley, D.D. (2009). SPSS macros for assessing the reliability and agreement of student evaluations of teaching. *Assessment & Evaluation in Higher Education*, 34 (6), 659-671.
- 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.
- Onwuegbuzie, A.J., Daniel, L.G., & Collins, K.M. (2009). A meta-validation model for assessing the score-validity of student teaching evaluations. *Quality and Quantity*, 43(2), 197-209. Retrieved from: <http://www.springerlink.com/content/l031k77753182234/>
- Onwuegbuzie, A.J., Witcher, A.E., Collins, K.M., Filer, J.D., Wiedmaier, C.D., & Moore, C.M. (2007). Students' perceptions of characteristics of effective college teachers: a validity study of a teaching evaluation form using a mixed-methods analysis. *American Educational Research Journal*, 44 (1), 113-160. Retrieved from:<http://www.jstor.org/stable/30069473>
- Ory, J.C. & Ryan, K. (2001). How do student ratings measure up to a new validity framework? *New Directions for Institutional Research*, 109, 27-44. Retrieved from:<http://onlinelibrary.wiley.com/doi/10.1002/ir.2/abstract>
- Patrick, C.L. (2011). Student evaluations of teaching: effects of the big five personality traits, grades, and the validity hypothesis. *Assessment & Evaluation in Higher Education*, 36 (2), 239-249.
- Pounder, J.S. (2007). Is student evaluation of teaching worthwhile? An analytical framework for answering the question. *Quality Assurance in Education*, 15 (2), 178-191. Retrieved from: <http://www.emeraldinsight.com/journals.htm?issn=0968-4883&volume=15&issue=2&articleid=1602895&show=abstract>
- Slate, J.R., LaPrairie, K.N., Shulte, D.P., & Onwuegbuzie, A.J. (2011). Views of effective college faculty: a mixed analysis. *Assessment & Evaluation in Higher Education*, 36 (3), 331-346.
- Sonntag, M.E., Bassett, J.F., & Snyder, T. (2009). An empirical test of the validity of student evaluations of teaching made on RateMyProfessors.com. *Assessment & Evaluation in Higher Education*, 34 (5), 499-504.

Sproule, R. (2000). Student evaluation of teaching: a methodological critique of conventional practices. *Educational Policy Analysis Archives*, 8 (50).

Titus, J.J. (2008). Student ratings in a consumerist academy: leveraging pedagogical control and authority. *Sociological Perspectives*, 51 (2), 397-422. Retrieved from: <http://www.jstor.org/stable/10.1525/sop.2008.51.2.397>

SET Ratings versus Student Achievement

Berk, R. A. (1988). Fifty reasons why student achievement gain does not mean teacher effectiveness. *Journal of Personnel Evaluation in Education*, 1 (4), 345-363. Retrieved from: <http://www.springerlink.com/content/p371756443080477/>

Cohen, P.A. (1981). Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. *Review of Educational Research*, 51 (3), 281-309. Retrieved from: <http://www.jstor.org/stable/1170209>

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 (6), 583-645. Retrieved from: <http://www.springerlink.com/content/p6515mu760856160/>

Feldman, K.A. (1989). Instructional effectiveness of college teachers as judged by teachers themselves, current and former students, colleagues, administrators, and external (neutral) observers. *Research in Higher Education*, 30 (2), 137-194. Retrieved from: <http://www.springerlink.com/content/q171194325658756/>

Feldman, K.A. (2007). Identifying exemplary teachers and teaching: evidence from student ratings. In R.P. Perry and J.C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, Netherlands: Springer. Retrieved from: <http://www.springerlink.com/content/r2727vn518394p70/>

Harootunian, B. (1971). Research on teaching effectiveness. In M. Mohan & R.E. Hull (Eds.), *Teaching effectiveness: Its Meaning, Assessment, and Improvement*. Englewood Cliffs, NJ: Educational Technology Publications.

Hildebrand, M. (1972). How to recommend promotion for a mediocre teacher without actually lying. *Journal of Higher Education*, 43 (1), 44-62. Retrieved from:<http://www.jstor.org/stable/1980780>

McKeachie, W.J. (2007). Good teaching makes a difference---and we know what it is. In R.P. Perry and J.C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, Netherlands: Springer. Retrieved from: <http://www.springerlink.com/content/p12lq1687231j030/>

Morsh, J., & Wilder, E. (1954). *Identifying the Effective Instructor: A Review of the Qualitative Studies, 1900-1952, Research bulletin AFPTRC-T-54-55*. San Antonio, TX: Lackland Air Force Base.

Murray, H.G. (1985). Classroom teaching behaviors related to college teaching effectiveness. In J.G. Donald & A.M. Sullivan (Eds.), *Using Research to Improve Teaching*. San Francisco: Jossey-Bass.

Murray, H.G. (2007). Low inference teaching behaviors and college teaching effectiveness: recent developments and controversies. In R.P. Perry and J.C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, Netherlands: Springer. Retrieved from:<http://www.springerlink.com/content/g44664vt4434i02r/>

Perry, R.P. (1991). Perceived control in college students: implications for instruction in higher education. In J. Smart (Ed.), *Higher Education: Handbook of Theory and Research*, 7, 1-56.

Theall, M. & Feldman, K.A. (2007). Commentary and update on Feldman's (1997) "Identifying Exemplary Teachers and Teaching: Evidence from Student Ratings." In R.P. Perry and J.C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, Netherlands: Springer. Retrieved from: <http://www.springerlink.com/content/r2727vn518394p70/>

SET Ratings versus Student Learning

Clayson, D.E. (2009). Student evaluations of teaching: Are they related to what students learn? A meta-analysis and review of the literature. *Journal of Marketing Education*, 31 (1), 16-30.

Galbraith, C.S., Merrill, G.B., & Kline, D.M. (2012). Are student evaluations of teaching effectiveness valid for measuring student learning outcomes in business related classes? A neural network and Bayesian analyses. *Research in Higher Education*, 53(3), 353-374. Retrieved from: <http://www.springerlink.com/content/2058756205016652/>

Gursoy, D. & Umbreit, W.T. (2005). Exploring student evaluations of teaching effectiveness: what factors are important? *Journal of Hospitality & Tourism Research*, 29 (1), 91-109. Retrieved from: <http://jht.sagepub.com/content/29/1/91.abstract>

Mohanty, G., Gretes, J., Flowers, C., Algozzine, B., & Spooner, F. (2005). Multi-method evaluation of instruction in engineering classes. *Journal of Personnel Evaluation in Education*, 18 (2), 139-151. Retrieved from:<http://www.springerlink.com/content/r8286hu90735j308/>

Naftulin, D.H., Ware, J.E., Donnelly, F.A. (1973). The Doctor Fox lecture: a paradigm of educational seduction. *Journal of Medical Education*, 48 (7), 630-635.

Rodin, M., & Rodin, B. (1972). Student evaluations of teachers: students rate most highly instructors from whom they learn least. *Science*, 177, 1164-1166. Retrieved from:<http://www.sciencemag.org/content/177/4055/1164.extract>

Shmanske, S. (1988). On the measurement of teaching effectiveness. *Journal of Economic Education*, 19 (4), 307-314. Retrieved from:<http://www.jstor.org/stable/1182341>

Yunker, P.J., & Yunker, J.A. (2003). Are student evaluations of teaching valid? Evidence from an analytical business core course. *Journal of Education for Business*, 78 (6), 313-317. Retrieved from:<http://www.tandfonline.com/doi/abs/10.1080/08832320309598619>

SET Ratings versus Faculty Changing/Improving Teaching Practices

Carle, A.C. (2009). Evaluating college students' evaluations of a professor's teaching effectiveness across time and instruction mode (online vs. face-to-face) using a multilevel growth modeling approach. *Computers & Education*, 53 (2), 429-435.

Cohen, P.A. (1980). Effectiveness of student rating feedback for improving college instruction: a meta-analysis of findings. *Research in Higher Education*, 13 (4), 321- 341. Retrieved from: <http://www.springerlink.com/content/r222967362184608/>

Hativa, N. (1996). University instructors' ratings profiles: stability over time, and disciplinary differences. *Research in Higher Education*, 37 (3), 341-365. Retrieved from:<http://www.springerlink.com/content/w777m55570631067/>

L'Hommedieu, R., Megnes, R.J., & Brinko, K.T. (1990). Methodological explanations for the modest effects of feedback from student ratings. *Journal of Educational Psychology*, 82 (2), 232-241.

SET Ratings versus Discipline

Cashin, W.E. (1990). Students do rate different academic fields differently. *New Directions for Teaching and Learning*, 43, 113-121. Retrieved from:<http://onlinelibrary.wiley.com/doi/10.1002/tl.37219904310/abstract>

Cashin, W.E. & Clegg, V.L. (1987). Are student ratings of different academic fields different? Paper presented at the annual meeting of the American Educational Research Association. (ERIC Document Reproduction Service No. ED286935).

Cashin, W.E. & Sixbury, G.R. (1993). *Comparative Data by Academic Field, IDEA Technical Report No. 8*. Manhattan, KS: Center for Faculty Evaluation and Development, Kansas State University.

Murray, H.G. & Renaud, R.D. (1995). Disciplinary differences in classroom teaching behaviors. *New Direction for Teaching and Learning*, 64, 31-39. Retrieved from: <http://onlinelibrary.wiley.com/doi/10.1002/tl.37219956406/abstract>

Murray, H.G., Rushton, J.P., & Paunonen, S.V. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology*, 82 (2), 250-261.

SET Ratings versus Class Size

Bedard, K. & Kuhn, P. (2008). Where class size really matters: class size and student ratings of instructor effectiveness. *Economics of Education Review*, 27 (3), 253-265. Retrieved from: <http://www.sciencedirect.com/science/article/pii/S0272775707000106>

Chau, C. (1997). A bootstrap experiment on the statistical properties of students' ratings of teaching effectiveness. *Research in Higher Education*, 38 (4), 497-517. Retrieved from: <http://www.springerlink.com/content/p188407883175564/>

Cramer, K.M., & Alexitch, L.R. (2000). Student evaluations of college professors: identifying sources of bias. *Canadian Journal of Higher Education*, 30 (2), 143-64.

Cranton, P., & Smith, R.A. (1986). A new look at the effect of course characteristics on student ratings. *American Educational Research Journal*, 23 (1), 117-128. Retrieved from: <http://www.jstor.org/stable/1163047>

Crittenden, K.S., Norr, J.L., & LeBailly, R.K. (1975). Size of university classes and student evaluations of teaching. *Journal of Higher Education*, 46 (4), 461-470. Retrieved from: <http://www.jstor.org/stable/1980673>

Fernandez, J., Mateo, M.A., & Muniz, J. (1998). Is there a relationship between class size and student ratings of teaching quality? *Educational and Psychological Measurement*, 58 (4), 596-604.

Kwan, K. (1999). How fair are student ratings in assessing the teaching performance of university teachers? *Assessment & Evaluation in Higher Education*, 24 (2), 181-195.

Marsh, H.W., Overall, J.U., & Kesler, S.P. (1979). Class size, students' evaluations, and instructional effectiveness. *American Educational Research Journal*, 16 (1), 57-70.

McPherson, M.A. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, 37 (1), 3-20.

McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

Wood, K., Linskey, A.S., & Strauss, M.A. (1974). Class size and student evaluations of faculty. *Journal of Higher Education*, 45 (7), 524-534.

Mateo, M.A., & Fernandez, J. (1996). Incidence of class size on the evaluation of university teaching quality. *Educational and Psychological Measurement*, 56 (5), 771-778.

SET Ratings versus Course Modality (online, blended, face-to-face)

Dziuban, C.D. & Moskal, P.D. (2011). A course is a course is a course: factor invariance in student evaluation of online, blended, and face-to-face learning environments. *Internet and Higher Education*, 14 (4), 236-241.

Wang, M.C, Dziuban, C.D., Cook, I.J., & Moskal, P.D. (2009). Dr. Fox rocks: using data-mining techniques to examine student ratings of instruction. In M.C. Shelly, L.D. Yore, & B. Hand (Eds.). *Quality research in literacy and science education: International perspectives and gold standards*. Dordrecht, Netherlands, Springer.

SET Ratings versus Student Gender

Centra, J.A., & Gaubatz, N.B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education*, 71 (1), 17-33.

Cramer, K.M., & Alexitch, L.R. (2000). Student evaluations of college professors: identifying sources of bias. *Canadian Journal of Higher Education*, 30 (2), 143-64.

McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

SET Ratings versus Student Expected Grades

Bharadwaj, S., Futrell, C.M., & Kantak, D.M. (1993). Using student evaluations to improve learning. *Marketing Education Review*, 3 (2), 16-21.

Clayson, D.E. (2004). A test of the reciprocity effect in the student evaluation of instructors in marketing classes. *Marketing Education Review*, 14 (2), 11-21.

Clayson, D.E., Frost, T.F., & Sheffet, M.J. (2005). Grades and the student evaluation of instruction: a test of the reciprocity effect. *Academy of Management Learning & Education*, 5 (1), 52-65.

Ewing, A.M. (2012) Estimating the impact of relative expected grade on student evaluations of teachers. *Economics of Education Review*, 31 (1), 141-154

Gigliotti, R.J., & Buchtel, F.S. (1990). Attributional bias and course evaluations. *Journal of Educational Psychology*, 82 (2), 341-351.

Goldberg, G., & Callahan, J. (1991). Objectivity of student evaluations of instructors. *Journal of Education for Business*, 66, 377-378.

Griffin, B.W. (2004). Grading leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educational Psychology*, 29, 410-425.

Krautmann, A.C., and W. Sander. (1999). Grades and student evaluations of teachers. *Economics of Education Review*, 18 (1), 59–63.

Marsh, H.W. (1980). The influence of student, course, and instructor characteristics in evaluations of university teaching. *American Educational Research Journal*, 17 (2), 219-237.

Marsh, H.W. (1982). Factors affecting students' evaluations of the same course taught by the same instructor on different occasions. *American Educational Research Journal*, 19 (4), 485-497.

Matos-Diaz, H. & Ragan, J.F. (2010). Do student evaluations of teaching depend on the distribution of expected grade? *Education Economics*, 18 (3), 317-330.

McPherson, M.A. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, 37 (1), 3-20.

McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

Patrick, C.L. (2011). Student evaluations of teaching: effects of the big five personality traits, grades, and the validity hypothesis. *Assessment & Evaluation in Higher Education*, 36 (2), 239-249.

Wright, R.E. & Palmer, J.C. (2006). A comparative analysis of different models explaining the relationship between instructor ratings and expected student grades. *Educational Research Quarterly*, 30, 3-18.

SET Ratings versus Student Perceptions of Rigor

Attiyeh, R., & Lumsden, K.G. (1972). Some modern myths in teaching economics: the U. K. experience. *American Economic Review*, 62, 429-433.

Centra, J.A. (2003). Will teachers receive higher student evaluations by giving higher grades and less course work? *Research in Higher Education*, 44, 495-518.

Clayson, D.E. & Haley, D.A. (1990). Student evaluations in marketing: what is actually being measured? *Journal of Marketing Education*, 12, 9-17

Cohen, P.A. (1981). Student ratings of instruction and student achievement: a meta-analysis of multisection validity studies. *Review of Educational Research*, 51 (3), 281-309.

Frey, P.W., Leonard, D.W., & Beatty, W.W. (1975). Student ratings of instruction: validation research. *American Educational Research Journal*, 12, 435-447. Retrieved from: <http://www.jstor.org/stable/1162743>

Marks, R.B. (2000). Determinants of student evaluations of global measures of instructor and course value. *Journal of Marketing Education*, 22, 108-119.

Pritchard, R.E. & Potter, G.C. (2011). Adverse changes in faculty behavior resulting from use of student evaluations of teaching: a case study. *Journal of College Teaching & Learning*, 8 (1), 1-7.

Steiner, S., Holley, L.C., Gerdes, K., & Campbell, H.E. (2006). Evaluating teaching: listening to students while acknowledging bias. *Journal of Social Work Education*, 42,355-376.

Titus, J.J. (2008). Student ratings in a consumerist academy: leveraging pedagogical control and authority. *Sociological Perspectives*, 51 (2), 397-422. Retrieved from:<http://www.jstor.org/stable/10.1525/sop.2008.51.2.397>

SET Ratings versus Student Perceptions of Faculty Personality (including Non-Verbal Behaviors)

Ambady, N., & Rosenthal, R. (1993). Half a minute: predicting teacher evaluations from thin slices of nonverbal behavior and physical attractiveness. *Journal of Personality and Social Psychology*, 64, 431-441

Anderson, K.J., & Smith, G. (2005). Students' preconceptions of professors: benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences*, 27 (2), 184-201.

Babad, E. (2007). Teachers' nonverbal behavior and its effects on students. In R.P. Perry and J.C. Smart (Eds.), *The scholarship of teaching and learning in higher education: An evidence-based perspective*. Dordrecht, Netherlands: Springer. Retrieved from: <http://www.springerlink.com/content/u434v619j7685l6t/>

Babad, E., Avni-Babad, D., & Rosenthal, R. (2004). Prediction of students' evaluations from brief instances of professors' nonverbal behavior in defined instructional situations. *Social Psychology of Education, 7* (1), 3-33.

Clayson, D.E., & Sheffet, M.J. (2006). Personality and the student evaluation of teaching. *Journal of Marketing Education, 28*, 149-160.

Erdle, S., Murray, H.G., & Rushton, J.P. (1985). Personality, classroom behavior, and college teaching effectiveness: a path analysis. *Journal of Educational Psychology, 77*, 394-407.

Merritt, D.J. (2008). Bias, the brain, and student evaluations of teaching. *St. John's Law Review, 82*, 235-287.

Murray, H.G. (1985). Classroom teaching behaviors related to college teaching effectiveness. In J.G. Donald & A.M. Sullivan (Eds.), *Using research to improve teaching*. San Francisco: Jossey-Bass.

Murray, H.G., Rushton, J.P., & Paunonen, S.V. (1990). Teacher personality traits and student instructional ratings in six types of university courses. *Journal of Educational Psychology, 82* (2), 250-261.

Naftulin, D.H., Ware, J.E., & Donnelly, F.A. (1973). The Doctor Fox lecture: A paradigm of educational seduction. *Journal of Medical Education, 48* (7), 630-635.

Patrick, C.L. (2011). Student evaluations of teaching: effects of the big five personality traits, grades, and the validity hypothesis. *Assessment & Evaluation in Higher Education, 36* (2), 239-249.

Titus, J.J. (2008). Student ratings in a consumerist academy: leveraging pedagogical control and authority. *Sociological Perspectives, 51* (2), 397-422. Retrieved from: <http://www.jstor.org/stable/10.1525/sop.2008.51.2.397>

Widmeyer, W.N., & Loy, J.W. (1988). When you're hot, you're hot! Warm-cold effects in first impressions of persons and teaching effectiveness. *Journal of Educational Psychology*, 80 (1), 118-121.

Williams, W.M., & Ceci, S.J. (1997). 'How'm I doing?' Problems with student ratings of instructors and courses. *Change: The Magazine of Higher Learning*, 29, 12-23.

SET Ratings versus Faculty Grading Practices (including Discussions of Grade Inflation)

Abrami, P.C., Dickens, W.J., Perry, R.P., & Leventhal, L. (1980). Do teacher standards for assigning grades affect student evaluations of instruction? *Journal of Educational Psychology*, 72, 107-118.

Boysen, G.A. (2008). Revenge and student evaluations of teaching. *Teaching of Psychology*, 35, 218-222.

Brockx, B., Spooren, P., & Mortelmans, D. (2011). Taking the grading leniency story to the edge. The influence of student, teacher, and course characteristics on student evaluations of teaching in higher education. *Educational Assessment, Evaluation and Accountability*, 23 (4), 289-306.

Centra, J.A. (2003). Will teachers receive higher student evaluations by giving higher grades and less coursework? *Research in Higher Education*, 44, 495-518.

Chacko, T.I. (1983). Student ratings of instruction: a function of grading standards. *Educational Research Quarterly*, 8 (2), 19-25.

Eiszler, D.F. (2002). College students' evaluations of teaching and grade inflation. *Research in Higher Education*, 43(4), 483-502. Retrieved from: <http://www.springerlink.com/content/a0c5ct6hm1xxgh3c/>

Ellis, L., Burke, D.M., Lomire, P. & McCormack, D.R. (2003). Student grades and average ratings of instructional quality: the need for adjustment. *Journal of Educational Research*, 97, 35-41.

- Goldman, L. (1985). The betrayal of the gatekeepers: grade inflation. *Journal of General Education*, 37 (2), 97-121.
- Greenwald, A.G., & Gillmore, G.M. (1997). Grading leniency is a removable contaminant of student ratings. *American Psychologist*, 52, 1209-1217.
- Greenwald, A.G., & Gillmore, G.M. (1997). No pain, no gain? The importance of measuring course workload in student ratings of instruction. *Journal of Educational Psychology*, 89, 743-751.
- Griffin, B.W. (2004). Grade leniency, grade discrepancy, and student ratings of instruction. *Contemporary Educational Psychology*, 29, 410-425.
- Johnson, V.E. (2003). *Grade inflation: A crisis in college education*. New York: Springer.
- Kolevzon, M.S. (1981). Grade inflation in higher education: a comparative study. *Research in Higher Education*, 15, 195-212.
- Marsh, H.W. (1987). Students' evaluations of university teaching: research findings, methodological issues, and directions for future research. *International Journal of Educational Research*, 11, 253-388.
- Marsh, H.W. & Roche, L.A. (2000). Effects of grade leniency and low workload of students' evaluations of teaching: popular myth, bias, validity, or innocent bystanders? *Journal of Educational Psychology*, 92, 202-228.
- Matos-Diaz, H. & Ragan, J.F. (2010). Do student evaluations of teaching depend on the distribution of expected grade? *Education Economics*, 18 (3), 317-330.
- Maurer, T.W. (2006). Cognitive dissonance or revenge? Student grades and course evaluations. *Teaching of Psychology*, 33, 176-179.
- McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

Moore, M., & Trahan, R. (1998). Tenure status and grading practices. *Sociological Perspectives*, 41, 775-781.

Redding, R.E. (1998). Students' evaluation of teaching fuel grade inflation. *American Psychologist*, 53, 1227-1228.

Rice, L. (1988). Student evaluation of teaching: problems and prospects. *Teaching Philosophy*, 11, 329-344.

Ryan, J.J., Anderson, J.A., & Birchler, A.B. (1980). Student evaluation: the faculty responds. *Research in Higher Education*, 12, 317-333.

Simpson, P.M., & Siguaw, J.A. (2000). Student evaluations of teaching: an exploratory study of the faculty response. *Journal of Marketing Education*, 22, 199-213.

Smith, D.L., Cook, P., & Buskist, W. (2011). An experimental analysis of the relation between assigned grades and instructor evaluations. *Teaching of Psychology*, 38 (4), 225-228.

Weinberg, B.A., Fleisher, B.M., & Hashimoto, M. (2007). Evaluating methods for evaluating instruction: the case of higher education (NBER Working Paper No. 12844). Retrieved from: <http://www.nber.org/papers/w12844>.

Wilhelm, W.B. (2004). The relative influence of published teaching evaluations and other instructor attributes on course choice. *Journal of Marketing Education*, 26, 17-30.

SET Ratings versus Faculty Gender (including Student Perceptions of Gender Roles)

Andersen, K., & Miller, E.D. (1997). Gender and student evaluations of teaching. *PS: Political Science & Politics*, 30 (2), 216-219.

Anderson, K.J. & Smith, G. (2005). Students' preconceptions of professors: benefits and barriers according to ethnicity and gender. *Hispanic Journal of Behavioral Sciences*, 27 (2), 184-201.

Arbuckle, J., & Williams, B. (2003). Student perceptions of expressiveness: age and gender effects on teacher evaluations. *Sex Roles: A Journal of Research*, 49 (9-10), 507-516).

Basow, S.A. (1995). Student evaluations of college professors: when gender matters. *Journal of Educational Psychology, 87*, 656-665.

Basow, S.A. & Montgomery, S. (2005). Student ratings and professor self-ratings of college teaching: effects of gender and divisional affiliation. *Journal of Personnel Evaluation in Education, 18* (2), 91-106. Retrieved from: <http://www.springerlink.com/content/j6307rm078586107/>

Basow, S.A., & Silberg, N.T. (1987). Student evaluations of college professors: are female and male professors rated differently? *Journal of Educational Psychology, 79*, 308-314.

Bennett, S.K. (1982). Student perceptions of and expectations of male and female instructors: evidence relating to the questions of gender bias in teaching evaluations. *Journal of Educational Psychology, 7*, 170-179.

Centra, J.A., & Gaubatz, N.B. (2000). Is there gender bias in student evaluations of teaching? *Journal of Higher Education, 71* (1), 17-33.

Freeman, H.R. (1994). Student evaluations of college instructors: effects of type of course taught, instructor gender and gender role, and student gender. *Journal of Educational Psychology, 86*, 627-630.

Kierstad, D., D'Agostino, P., & Dill, H. (1988). Sex role stereotyping of college professors: bias in student ratings of instructors. *Journal of Educational Psychology, 80*, 342-344.

Kogan, L.R., Schoenfeld-Tacher, R., & Hellyer, P.W. (2010). Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education, 15* (6), 623-636.

Langbein, L.I. (1994). The validity of student evaluations of teaching. *PS: Political Science & Politics, 27* (3), 545-553.

Laube, H., Massoni, K., Sprague, J., & Ferber, A. (2007). The impact of gender on the evaluation of teaching: what we know and what we can do. *NWSA Journal, 19* (3), 87-104.

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. Dordrecht, Netherlands: Springer.

Martin, E. (1984). Power and authority in the classroom: sexist stereotypes in teaching evaluations. *Signs*, 9, 483-492.

McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

Sidanius, J. & Crane, M. (1989). Job evaluation and gender: The case of university faculty. *Journal of Applied Social Psychology*, 19, 174-197.

Sprague, J., & Massoni, K. (2005). Student evaluations and gendered expectations: What we can't count can hurt us. *Sex Roles: A Journal of Research*, 53 (11-12), 779-793.

Statham, A., Richardson, L., & Cook, J. (1991). *Gender and university teaching: A negotiated dDifference*. Albany, NY: SUNY Press.

Tatro, C.N. (1995). Gender effects of student evaluations of faculty. *Journal of Research and Development in Education*, 28, 169-173.

SET Ratings versus Faculty Ethnicity

Carle, A.C. (2009). Evaluating college students' evaluations of a professor's teaching effectiveness across time and instruction mode (online vs. face-to-face) using a multilevel growth modeling approach. *Computers & Education*, 53 (2), 429-435.

McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

Merritt, D.J. (2008). Bias, the brain, and student evaluations of teaching. *St. John's Law Review*, 82, 235-287.

SET Ratings versus Faculty Sexual Orientation

Jennings, T. (2010). Teaching "out" in the university: an investigation into the effects of Lesbian, Bisexual, and Transgender faculty self-disclosure upon student evaluations of faculty teaching effectiveness in the USA. *International Journal of Inclusive Education*, 14 (4), 325-339.

Liddle, B.J. (1997). Coming out in class: disclosure of sexual orientation and teaching evaluations. *Teaching of Psychology*, 24 (1), 32-35.

SET Ratings versus Faculty Age

Arbuckle, J., & Williams, B. (2003). Student perceptions of expressiveness: age and gender effects on teacher evaluations. *Sex Roles: A Journal of Research*, 49 (9-10), 507-516.

McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

SET Ratings versus Faculty Rank

Kogan, L.R., Schoenfeld-Tacher, R., & Hellyer, P.W. (2010). Student evaluations of teaching: perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education*, 15 (6), 623-636.

Thyer, B.A., Myers, L.L., & Nugent, W.R. (2011). Do regular social work faculty earn better student course evaluations than do adjunct faculty or doctoral students? *Journal of Teaching in Social Work*, 31 (4), 365-377.

Methods to Improve the Fairness of SET Ratings

Abrami, P.C. (2001). Improving judgments about teaching effectiveness using teacher rating forms. *New Directions for Institutional Research*, (2001), 109: 59-87.

Abrami, P.C., d'Apollonia, S., & Cohen, P.A. (1990). Validity of student ratings of instruction: what we know and what we do not. *Journal of Educational Psychology*, 82(2), 219-231.

Cashin, W.E. & Sixbury, G.R. (1993). *Comparative Data by Academic Field, IDEA Technical Report No. 8*. Manhattan, KS: Center for Faculty Evaluation and Development, Kansas State University.

Catano, V.M. & Harvey, S. (2011). Student perception of teaching effectiveness: development and validation of the Evaluation of Teaching Competencies Scale (ETCS). *Assessment & Evaluation in Higher Education*, 36 (6), 701-711.

Cohen, E.H. (2005). Student evaluations of course and teacher: factor analysis and SSA approaches. *Assessment & Evaluation in Higher Education*, 30 (2), 123-126.

Danielsen, A.L., & White, R.A.. (1976). Some evidence on the variables associated with student evaluations of teachers. *Journal of Economic Education*, 7, 117-19.

De Witte, K., & Rogge, N. (2011). Accounting for exogenous influences in performance evaluations of teachers. *Economics of Education Review*, 30 (4), 641-653.

Denson, N., Loveday, T., & Dalton, H. (2010). Student evaluations of courses: what predicts satisfaction? *Higher Education Research and Development*, 29 (4), 339-356.

Dresel, M. & Rindermann, H. (2011). Counseling university instructors based on student evaluations of their teaching effectiveness: a multilevel test of its effectiveness under consideration of bias and unfairness variables. *Research in Higher Education*, 52 (7), 717-737.

Kogan, L.R., Schoenfeld-Tacher, R., & Hellyer, P.W. (2010). Student evaluations of teaching: Perceptions of faculty based on gender, position, and rank. *Teaching in Higher Education*, 15 (6), 623-636.

Laube, H., Massoni, K., Sprague, J., & Ferber, A. (2007). *The impact of gender on the evaluation of teaching: What we know and what we can do*. *NWSA Journal*, 19 (3), 87-104.

Lindahl, M.W., & Unger, M. (2010). Cruelty in student teaching evaluations. *College Teaching*, 58 (3), 71-76.

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*. Dordrecht, Netherlands: Springer.

Mason, P.M., Steagall, J.W., & Fabritius, M.M. (1995). Student evaluations of faculty: A new procedure for using aggregate measures of performance. *Economics of Education Review*, 14 (4), 403-16.

McCullough, B.D. & Radson, D. (2011). Analysing student evaluations of teaching: comparing means and proportions. *Evaluation & Research in Education*, 24 (3), 183-202.

McNatt, D.B. (2010). Negative reputation and biased student evaluations of teaching: longitudinal results from a naturally occurring experiment. *The Academy of Management Learning and Education*, 9 (2), 225-242.

McPherson, M.A. (2006). Determinants of how students evaluate teachers. *Journal of Economic Education*, 37 (1), 3-20.

McPherson, M.A., Jewell, R.T., & Kim, M. (2009). What determines student evaluation scores? A random effects analysis of undergraduate economics classes. *Eastern Economic Journal*, 35 (1), 37-51.

Morley, D.D. (2009). SPSS macros for assessing the reliability and agreement of student evaluations of teaching. *Assessment & Evaluation in Higher Education*, 34 (6), 659-671.

Nevo, D., McClean, R., & Nevo, S. (2010). Harnessing information technology to improve the process of students' evaluations of teaching: an exploration of students' critical success factors of online evaluations. *Journal of Information Systems Education*, 21 (1), 99-109.

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.

Stark-Wroblewski, K., Ahlering, R.F., & Brill, F.M. (2007). Toward a more comprehensive approach to evaluating teaching effectiveness: supplementing student evaluations of teaching with pre-post learning measures. *Assessment & Evaluation in Higher Education*, 32 (4), 403-415.

Wang, M.C, Dziuban, C.D., Cook, I.J., & Moskal, P.D. (2009). Dr. Fox rocks: using data-mining techniques to examine student ratings of instruction. In M.C. Shelly, L.D. Yore, & B. Hand (Eds.). *Quality research in literacy and science education: International perspectives and gold standards*. Dordrecht, Netherlands, Springer.

Wongsurawat, W. (2011). What's a comment worth? How to better understand student evaluations of teaching. *Quality Assurance in Education: An International Perspective*, 19 (1), 67-83.