

REFERENCES

- Abrami, P. C. (1989). How Should We Use Student Ratings to Evaluate Teaching? Research in Higher Education, 30, 2, 221-27.
- Abrami, P. C., D' Apollonia, S., & Cohen, P. A. (1990). The validity of student ratings of instruction: What we know and what we don't. Journal of Educational Psychology, 82, 219-231.
- Aleamoni, L. M.; & Graham, M. H. (1974). The Relationship between CEQ ratings and instructor's rank, class size and course level. Journal of Educational Measurement;11, 3, 189-202, F 74.
- Aleamoni, L. M. & Hexner, P. Z. (1981). A Review of the Research on Student Evaluation and a Report on the Effect of Different Sets of Instructions on Student Course and Instructor Evaluation. Instructional Science, 9, 1, 67-84.
- American Evaluation Association. (1993). Guiding principles for evaluators. <http://www.eval.org/EvaluationDocuments/aeaprin6.html>
- Andrews, H. A. (1985). Evaluating for excellence: Addressing the need for responsible and effective faculty evaluation. Stillwater, OK: New Forums Press.
- Atkinson, R. C. & Tuzin, D. (1992). Equilibrium in the research university. Change, 3, 21-31.
- Bailiff, J. & Khan, S. (1987). The university and the rediscovery of teaching: A system-level model. In Kurfis, J. (Ed.) *To improve the academy: Resources for student, faculty and institutional development* (pp. 75-82). Stillwater, OK: New Forum Press.
- Ballantyne, C. (November, 1998). What students think: An innovative look at student evaluations of teaching. Paper presented at the annual meeting of the American Evaluation Association, Chicago.
- Barnes, L. B. & Barnes M. W. (1993). Academic discipline and generalizability of student evaluations of instruction. Research in Higher Education, 34, 135-149.
- Bess, J. L. (1982). University organizations: A matrix analysis of the academic professions. New York, NY: Human Sciences Press Inc.
- Biglan, A. (1973). The characteristics of subject matter in different academic areas. Journal of Applied Psychology, 57, 195-203.

- Bok, D. (1992). The improvement of teaching. New York, NY: American Council of Learned Societies.
- Boyer, E. L. (1987). College: The undergraduate experience in America. New York, NY: Harper & Row.
- Braskamp, L. A., Brandenburg, D. C. & Ory, J. C. (1984). Evaluating teaching effectiveness. London: Sage Publications, Inc.
- Braskamp, L. A. & Ory, J. C. (1994). Assessing faculty work: Enhancing individual and institutional performance. San Francisco, CA: Jossey Bass.
- Brodie, D. A. (1998). Do students report that easy professors are excellent teachers? The Canadian Journal of Higher Education, 23, 1, 1-20.
- Brodie, D.A. (1999). Has publication bias inflated the reported correlation between student achievement and ratings of instructors? Paper presented at the Annual Meeting of the American Educational Research Association, Montreal CA, April.
- Bukalski, P. J. & Zirpola, D. J. (1993). Guide for non-tenured faculty members: Annual evaluation, promotion, and tenure (University Film and Video Association Monograph No. 6). Atlanta, GA: Georgia State University, Department of Communication.
- Bureau of National Affairs (1964). The Civil Rights Act of 1964: text, analysis, legislative history: What it means to employers, businessmen, unions, employees, minority groups. Bureau of National Affairs, Washington: DC.
- Burke, P. (1992). The Fabrication of Louis XIV. Yale University Press, New Heaven and London. New Haven, CT.
- Cahn, S. (1987, October 14). Faculty members should be evaluated by their peers, not by their students. Chronicle of Higher Education, p. B2.
- Cambridge, B. L. (1996). The paradigm shifts: Examining the quality of teaching through assessment of student learning. Innovative Higher Education, 20, 4, 287-297.
- Campbell, D. T. and Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait matrix. Psychological Bulletin, 56, 81-105.
- Cashin, W. E. (1988). Student ratings of teaching: A summary of the research. IDEA paper No. 20. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.

- Cashin, W. E. (1990). Students do rate different academic fields differently. In M. Theall & J. Franklin (Eds.), Students ratings of Instruction: Issues for improving practice. New directions for teaching and learning (pp. 113-121). San Francisco: Jossey-Bass.
- Cashin, W. E. (1995). Student ratings of teaching: The research revisited. IDEA Paper No. 32. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Cashin, W. E., & Downey, R. G. (1992). Using Global Student Rating Items for Summative Evaluation. Journal of Educational Psychology, 84, 4 563-72.
- Cashin, W. E., & Perrin, B. (1978). Description of IDEA Standard Form Data Base. IDEA Technical Report No. 4. Manhattan, KS: Kansas State University, Center for Faculty Evaluation and Development.
- Cave, M., Hanney, S., Kogan, M. & Travett, G. (Eds.). (1988). The use of performance indicators in higher education: A critical analysis of developing practice. London: Kingsley Publishers, Ltd.
- Centra, J. A. (1977). Student ratings of instruction and their relationship to student learning. American Educational Research Journal; 14, 1, 17-24.
- Centra, J. A. (1979). Determining faculty effectiveness. San Francisco, CA: Jossey-Bass Inc.
- Centra, J. A. (1993). Reflective faculty evaluation: Enhancing teaching and determining faculty effectiveness. San Francisco, CA: Jossey-Bass Inc.
- Centra, J. A.; & Bonesteel, P. (1990). College teaching: An art or a science? In M. Theall and Franklin, Student ratings of instruction: Issues for improving practice. New Directions for Teaching and Learning, 43. 7-15. Fall.
- Chait, R. P. & Ford, A. T. (1982). Beyond traditional tenure: A guide to sound policies and practices. San Francisco, CA: Jossey-Bass Ltd.
- Chickering, A. W. & Gamson, Z. F. (1987). Seven principles for good practice in undergraduate education. AAHE Bulletin, pp. 3-7 March.
- Chiu, S. (2000). Use of the unbalanced nested ANOVA to examine the factors influencing student ratings of instructional quality. Unpublished doctoral dissertation. Champaign, IL: University of Illinois at Urbana-Champaign.

- Cisneros-Cohernour, E. (1998). Trade-offs: Using the feedback of student ratings for instructional improvement. Champaign, IL: University of Illinois at Urbana-Champaign.
- 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.
- Cohen, P. A. (1983). Comment on selective review of the validity of student ratings of teaching. Journal of Higher Education, *54*, 448-458.
- Cohen, P. A. (1986). An updated and expanded meta-analysis of multisection student rating validity studies. Paper presented at the Annual Meeting of the American Educational Research Association. San Francisco, CA, April.
- Cohen, P. A. (1990). Bringing research into practice. In New Directions for Teaching and Learning; Student Ratings of Instruction: Issues for Improving Practice, *43*, 123-32.
- Costin, F., Greenough, & Menges, R. J. (1971). Student ratings of college teaching: reliability, validity, and usefulness. Review of Educational Research, *41*, 5, 511-535.
- Cronbach, L.J. (1989). Construct validation after thirty years. In R.L. Linn (Ed.), Intelligence: Measurement, theory, and public policy (pp.147-171). Chicago: University of Illinois Press.
- Cronbach, L. J., & Glesser, G. C. (1965). Psychological and personnel decisions (2nd. Ed.). Urbana, IL: University of Illinois Press.
- Crooks, T. J., & Kane, M. T, (1981). The generalizability of student ratings of instructors: Item specificity and section effects. Research in Higher Education, *15*, 4, 305-13 1981.
- Cross, P. K. (1987). The need for classroom research. In J. Kurfiss (Ed), To improve the academy: Resources for student, faculty and institutional development (pp. 1-17). Stillwater, OK: New Forum Press.
- Cross, P. K. (1993). Improving the quality of instruction. In A. Levine (Ed.), Higher learning in America (pp. 287-305). Baltimore & London: The John Hopkins University Press.
- Cuban, L. (1999). How scholars trumped teachers: Change without reform in university curriculum, teaching, and research, 1890-1990. New York, NY: Teachers College Press.

- Cugoano, Q. O. (1997) Thoughts and sentiments on the evil and the wicked traffic of the slavery and commerce of the human species. In E. C. Eze (Ed.). Race and the enlightenment: A reader. Pp. 79-153. Cambridge, MA: Blackwell.
- Damron, J. C. (1996). Instructor personality and the politics of the classroom. Manuscript, Douglas College, New Westminster, British Columbia, Canada V3L 5B2. (Earlier versions in the June 1994 Issue of Faculty Matters (No. 5, Pages 9-12) and the September, 1994 Issue of Update (The Newsletter of the Okanagan University College Faculty Association). Douglas College)
- Doyle, K. O., Jr. (1982). Evaluating teaching. Lexington, Mass: Lexington Books.
- Dowell, D. A., and Neal, J. A. (1982). Selective review of the validity of student ratings of teaching. Journal of Higher Education, 53, 1, 51-62.
- Dressel, P. L. (1970). Evaluation of the environment, the process, and the results of higher education. In A. S. Knowles (ED.), Handbook of college and university administration: Academic (pp. 53-80). New York, NY: McGraw Hill.
- Dunkin, M. J. (1986). Research on teaching in higher education. In Merlin C. Wittrock. (Ed.), Handbook of research on teaching. American Educational Research Association, New York: NY.
- Dwinell, P. L. & Higbee, J. L. (1993). Students' perceptions of the value of teaching evaluations. Perceptual and Motor Skills, 76, 995-100.
- Dwyer, C. A. (1993). Teaching and diversity: Meeting the challenge for innovative teaching assessment. Journal of Teacher Education, 44, 119-129.
- Eble, K. E. (1972). Professors as teachers. San Francisco, CA: Jossey-Bass, Inc.
- El Hassan, K. (1995). Students' ratings of instruction: Generalizability of findings. Studies in Educational Evaluation, 21, pp. 411-429
- Embreston (Whitely), S. (1983). Construct validity: Construct representation versus nomothetic span. Psychological Bulletin, 93, 179-197.
- Erickson, F. (1986). Qualitative methods in research on teaching. In Merlin C. Wittrock. (Ed.). Handbook of research on teaching. American Educational Research Association, New York: NY.
- Eze, E. C. (Ed.), (1997). Race and the enlightenment: A reader. Cambridge, MA: Blackwell.

- Falk, B., & Dow, K. L. (1971). The assessment of university teaching. London: Society for Research into Higher Education Ltd.
- Fayne, H. R. (1991). Practicing what we preach: Key issues in faculty evaluation. Paper presented at the annual meeting of the American Association of Colleges for Teacher Education, Atlanta, GA.
- Feldman, K. A. (1976). The superior college teacher from the student's view. Research in Higher Education, 5, 243-288.
- Feldman, K. A. (1977). Consistency and variability among college students in rating their teachers: A review and analysis. Research in Higher Education, 6, 223-274.
- Feldman, K. A. (1978). Course characteristics and college students' ratings of their teachers: What we know and what we don't. Research in Higher Education, 9, 199-242.
- Feldman, K. A. (1984). Class size and college students' evaluations of teachers and courses: A closer look. Research in Higher Education, v21 n1 p45-116.
- Feldman, K. A. (1986). The perceived instructional effectiveness of college teachers as related to their personality and attitudinal characteristics. Research in Higher Education, 24, 139-170.
- Feldman, K. A. (1987). Research productivity and scholarly accomplishment of college teachers as related to their instructional effectiveness: A review and exploration. Research in Higher Education; v26 n3 p227-98.
- Feldman, K. A. (1988). Effective college teaching from the students' and faculty's view: Matched or mismatched priorities? Research in Higher Education, 28, 4, 291-344.
- 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.
- 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, 137-172.
- 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.

- Ferren, A., & Mussell, K. (1987). Strengthening faculty development programs through evaluation. In J. Kurfiss, To improve the academy: Resources for student, faculty, and institutional development (p. 133-143). Stillwater, OK: New Forum Press.
- Finkelstein, M. J. (1984). Peer review: Documentary evidence. In Milman, J. (Ed.), Handbook of Teacher Education (pp. 73-79). London: Sage publications.
- Frey, P. W. (1976). Validity of student instructional ratings: Does timing matter? Journal of Higher Education, 42, 3, 327-336.
- Gage, N. L. & Needles, M. C. (1989). Process-product research on teaching: A review of the criticisms. The Elementary School Journal, 89, 253-300.
- Genova, W. J., Madoff, M. J., Chin, R. & Thomas, G. B. (1976). Mutual benefit evaluation of faculty and administrators in higher education. Newton, MA: Ballinger Publishing Co.
- Gillmore, G. Kane, M. T., & Naccarato, R. W. (1978). The generalizability of student ratings of instruction: Estimation of the teacher and course components. Journal of Educational Measurement, 15, 1, 1-13.
- Goodwin, L. D. & Stevens, E. A. (1993). The influence of gender on university faculty members' perceptions of "good" teaching. Journal of Higher Education, 64, 166-185.
- Greene, J. C., Caracelli, V. J., & Graham, W. E. (1989). Toward a conceptual framework for mixed-method evaluation designs. Educational Evaluation and Policy Analysis, 11(3), 255-274.
- Greenwald, A. G. (1996). Validity and usefulness of student ratings of instruction. American Psychologist, 52, 11, 1182-1186.
- Greenwald & Gillmore, (1996). No pain, no gain? : The importance of measuring course workload in student ratings of instruction, University of Washington, Manuscript submitted for publication.
- Guba. E., & Lincoln, Y. (1998). Competing paradigms in qualitative research. In N. Denzin & Y. S. Lincoln (Eds.). The landscape of qualitative research: Theories and issues. Sage Publications, Thousand Oaks: CA.
- Haskell, (1997). Academic freedom, tenure, and students evaluation of faculty: Galloping polls in the 21st Century. Educational Policy Analysis Archives (On line), 5, 6, February. Available: <http://olam.ed.asu.edu/epaa/v5n6.html>

- Hattie, John, & Marsh, H. W. (1996). The relationship between research and teaching: A meta-analysis. Review of Educational Research; v66 n4 p507-42 Win.
- Hawley, R. C. (1977). Faculty evaluation: Some common pitfalls. Independent School Journal, 5, 39-40.
- Hogan, T. (1973). Similarity of student ratings across instructors, courses, and time. Research in Higher Education; 1, 2, 149-154.
- Howard, G. S., Conway, C. G., & Maxwell, S. E. (1985). Construct validity of measures of college teaching effectiveness. Journal of Educational Psychology, 77, 187-196.
- Hunter, J. E., Schmidt, F. L., & Jackson, C. B. (1982). Advanced meta-analysis: Quantitative methods of cumulating research findings across studies. San Francisco: Sage.
- Hunter, M. C. (1982). Mastery teaching/Madeline Hunter. Thousand Oaks, CA: Corwin Press
- Jacobsen, R. H. (1989). The impact of faculty incentive grants on teaching effectiveness. Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.
- Kant, I. (1975). This fellow was quite black. . . a clear proof that what he said was stupid. In E.C. Eze. (Ed.), (1997). Race and the enlightenment: A reader. Cambridge, MA: Blackwell.
- Kinney, D. P. & Smith, S. P. (1992). Age and teaching performance. Journal of Higher Education, 63, 282-302.
- Knapper, C. K., Geis, G. L., Pascal, Ch. E. & Shore, B. M. (1977). If teaching is important...: The evaluation of instruction in higher education. Canada: Clark, Irwin & Company Limited.
- Kulik, J. A. & Kulik, C.C. (1974). Student ratings of instruction. Teaching Psychology, 1, 51-57.
- Kulik, J. A. & McKeachie, W. J. (1975) "The evaluation of teachers in higher education." In F. N. Kerlinger (ed.), Review of Research in Education. Itasca, IL: Peacock,
- Kyvik, S.; & Smeby, J.-C. (1994). Teaching and research. The Relationship between the supervision of graduate students and faculty research performance. Higher Education, v28 n2 p227-39 Sep.

- Lane, S., Parke, C.S., & Stone, C.A. (1998). A framework for evaluating the consequences of assessment programs." Educational Measurement: Issues and Practice, 17(2), 24-28.
- Lennon, R. T. (1956). Assumptions underlying the use of content validity. Educational and Psychological Measurement, 16, 294-304.
- Licata, Christine M. (1986). Post-tenure faculty evaluation: Treat or opportunity? Washington, DC: ASHE-ERIC.
- Loevinger, J. (1957). Objective tests as instruments of psychological theory (Monograph), Psychological Reports, 3, 635-694.
- Longman, K. A. (1978). Faculty development in a period of retrenchment. Unpublished doctoral dissertation, University of Michigan, MI.
- Mabry, L. (1999). Portfolios plus: A critical guide to alternative assessment. Thousand Oaks, CA: Corwin Press, Inc.
- Magusson, K. (2000). Personal interview.
- Mann, M. P. (1987). Developmental models of faculty careers: A critique of research and theory. In J. Kurfiss, To improve the academy: Resources for student, faculty, and institutional development (p.19-32). Stillwater, OK: New Forum Press.
- Marlin, J. W. J. (1987). Student perception of end-of-course evaluations. Journal of Higher Education, 58(6), 704-716.
- Marsh, H. W. (1982). SEEQ: A reliable, valid, and useful instrument for collecting students' evaluations of university teaching, British Journal of Educational Psychology, 52, 77, 95.
- Marsh, H. W. (1984). Students' evaluations of teaching: Dimensionality, reliability, validity, potential biases, and utility. Journal of Educational Psychology, 76, 707-754.
- Marsh, H. W. (1987). Student's evaluations of university teaching: Research findings, methodological issues, and directions for future research. International Journal of Educational Research, 11, 253-388.
- Marsh, H. W. & Bailey, M. (1993). Multidimensional students' evaluations of teaching effectiveness. Journal of Higher Education, 64, 1-18.
- Marsh, Herbert W.; & Overall, J. U. (Spring, 1981). The relative influence of course level, course type, and instructor on students' evaluations of college teaching. American Educational Research Journal; v18 n1 p103-12.

- Marsh, H. W., Overall, J. U., & Kesler, S. P. (1979). Validity of student evaluations of instructional effectiveness: A comparison of faculty self-evaluations and evaluations by their students. Journal of Educational Psychology, 71, 149-160.
- McKeachie, W. J. (1971). Research on college teaching: Memo to the faculty. Ann Arbor Michigan, MI: Center for research on Learning.
- McKeachie, W. J. (1990). Research on college teaching: The historical background. Journal of Educational Psychology, 82, 189-200.
- McKeachie, W. J. (1997). Student ratings: The validity of use. American Psychologist, 52, 11, 1218-1225.
- McKeachie, W. J., Isaacson, W. J. & Millholand, J. E. (1964). Research on the characteristics of effective college teaching (Cooperative Research Rep. No. O E 850). Ann Arbor, MI: The University of Michigan.
- McKeachie, W. J., & Lin, Y. G. (1979). A Note on Validity of Student Ratings of Teaching. Educational Research Quarterly, 4, 3, 45-47.
- Menges, R. J. (1991). Why hasn't peer evaluation of college teaching caught on? Paper presented at the meeting of the American Educational Research Association, Chicago, IL.
- Menges, R. J. (1987). Colleagues as catalyst for change in teaching. In J. Kurfiss, To improve the academy: Resources for student, faculty, and institutional development (p. 83-93). Stillwater, OK: New Forum Press.
- Menges, R. J. (1998). Shortcomings of research on evaluation and improving teaching in higher education. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego: CA, April.
- Messick, S. (1980). Test validity and the ethics of assessment. American Psychologist, 35, 1012-1027.
- Messick, S. (1989). Validity. In R. L. Linn (Ed.), Educational Measurement (3rd ed.). New York: American Council on Education and Macmillan.
- Messick, S. (October, 1994). Alternative modes of assessment: Uniform standards of validity. Paper presented at a conference on Evaluating Alternatives to Traditional Testing for Selection, Bowling Green: OH.

- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741-749.
- Messick, S. (1994). Alternative modes of assessment, uniform standards of validity. Research Report.
- Messick, S. L. (1989). "Validity". In R. L. Linn (Ed.), Educational Measurement (3rd Edition). New York: American Council on Education & MacMillan.
- Messick, S. L. "Validity". In R. L. Linn (Ed.), Educational Measurement (3rd Edition). New York: American Council on Education and MacMillan, (1989).
- Miller, R. I. (1972). Evaluating faculty performance. San Francisco, CA: Jossey- Bass Inc.
- Miller, R. I. (1974). Developing programs for faculty evaluation. San Francisco, CA: Jossey-Bass Inc.
- Miller, R. I. (1979). The assessment of college performance: A handbook of techniques and measures for institutional self-evaluation. San Francisco, CA: Jossey-Bass, Inc.
- Miller, R. I. (1987). Evaluating faculty for promotion and tenure. San Francisco: Jossey Bass Inc.
- Moffett, B. S. (1997). Development of a peer evaluation model for clinical teaching faculty. Paper presented at the Annual Meeting of the American Educational Research Association, Chicago: IL, March.
- Morehead, J. W. & Shedd, P. J. (1996). Student interviews: A vital role in the scholarship of teaching. Innovative Higher Education, 20, 4, 261-269.
- Moses, I. (1988). Academic staff evaluation and development. Queensland, Australia: University of Queensland Press.
- Moss, P.A. (1992). Shifting conceptions of validity in educational measurement: Implications for performance assessment. Review of Educational Research, 62, 229-258.
- Moss, P.A. (1996) Enlarging the dialogue in educational measurement: Voices from interpretive research traditions. Educational Researcher, 25(1), 20-28.
- Moss, P.A. (1998). The of consequences in validity theory. Educational Measurement: Issues and Practice, 17(2), 6-13.

- Murray, H.G., Rushton, J.P., & Paunonen, S.V. (1900). Teacher personality traits and student instructional ratings in six types of university courses. Journal of Educational Psychology, 82(2), 250.
- Neal, J. E. (1988). Faculty evaluation: Its purposes and effectiveness (Report No. ED-RI-88-062014). Washington, DC: ERIC Clearinghouse on Higher Education. (ERIC Document Reproduction Service No. ED 308 800)
- Newman, Y., & Newman, L. (1983). Characteristics of academic areas and students' evaluation of instruction. Research in Higher Education, 19, 3, 323-334.
- Nyquist, J. D. & Staton-Spicer, A. Q. (1987). Non-traditional intervention strategies for improving the teaching effectiveness of graduate assistants. In J. Kurfiss, To improve the academy: Resources for student, faculty, and institutional development (pp. 168-183). Stillwater, OK: New Forum Press.
- Ory, J. (1998). Teaching evaluation: Past, present and future. Paper presented at the Annual Meeting of the American Educational Research Association, San Diego CA., April.
- Ory, J. C., Braskamp, L. A., & Pieper, D. (1980) The congruency of student evaluative information collected by three methods. Journal of Educational Psychology, 72, 181-185.
- Ory, J. & Ryan, K. (2001). How do student ratings measure up to a new validity framework? New Directions in Institutional Research, 109, Jossey-Bass Inc., Publishers, San Francisco: CA.
- Ory, J., & Wieties, (1991). A longitudinal study of faculty selection of ICES student evaluation items. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April.
- O'Toole, J., Van Alstyne, W. W. & Chait, R. (1979). Three views: Tenure. New Rochelle, NY: Change Magazine Press.
- Overall, J. U. & Marsh (1979-80?). Students' evaluations of instruction: A longitudinal study of their stability. Journal of Educational Psychology, 72, 321-325.
- Patton, M. Q. (1980). Qualitative Evaluation Methods. Beverly Hills, Sage Publications.
- Player, M. A., Shoben, E. W., & Liebowitz, R. L. (1995). Employment discrimination law: Cases and materials, St. Paul: MN: West Publishing Co.

- Powell, S. (1987). Faculty development needs as a function of status in the academic guild. In J. Kurfiss, To improve the academy: Resources for student, faculty, and institutional development (pp. 33-40). Stillwater, OK: New Forum Press.
- Reckase, M.D. (1998) Consequential validity from the test developer's perspective. Educational Measurement: Issues and Practice, 17(2), 13-16.
- Rosenshine, B. (1987). Explicit teaching. In D.C. Berliner, & B. Rosenshine, (Eds.), Talk to teachers: A festschrift for N. L. Gage (pp. 75-92), New York: Random House.
- Roush, D. C. (1983). Strategies for effective university teaching. Materials for Teaching-Methodology Workshops of the Fullbright Exchange Program. (Available from the Latin American Scholarship Program for American Universities, 25 Mt. Auburn Street, Cambridge, MA 02138-6095).
- Ryan, K. & Johnson, T. (1998). Democratizing evaluation: Meanings and methods from practice. Paper presented at the Annual Meeting of the American Evaluation Association, Chicago, IL, November.
- Scriven, M. (1991). Evaluation Thesaurus. 4th, Sage Publications, Newbury Park: CA.
- Scriven, M. (1993). The validity of student ratings: In teacher evaluation. Evaluation & Development Group.
- Scriven, M. (1994). A unified theory approach to teacher evaluation. Studies in Educational Evaluation, 21, 111-129.
- Seldin, M. (1997). The teaching portfolio: A practical guide to improved performance and promotion/tenure decisions. (2nd. Ed). Bolton, MA: Anker Publishing.
- Seldin, P. (1993). How colleges evaluate professors 1983 versus 1993. AAHE Bulletin, 6-8, 12.
- Seldin, P. (1984). Changing practices in faculty evaluation. San Francisco: CA, Jossey-Bass.
- Shapiro, B. J. (1983). Probability and certainty in seventeenth-century England: A study of the relationships between natural science, religion, history, law and literature. Princeton, NJ. : Princeton University Press.
- Shepard, L. Evaluating test validity. Review of Educational Research, 19, 405-450.
- Shingles, R. D. (1977). Faculty ratings: procedures for interpreting student evaluations. American Educational Research Journal, 14, 4, 459-70.

- Shulman, L. S. (1986). Paradigms and research programs in the study of teaching: A contemporary perspective. In M. C. Wittrock. (Ed.). Handbook of research on teaching. American Educational Research Association, New York: NY.
- Stake, R. E. (1995). The Art of Case Study Research. Sage Publications, Thousand Oaks: CA.
- Stake, R. E. (1996). Validity. CIRCE Publications. University of Illinois at Urbana-Champaign. <http://www.ed.uiuc.edu/CIRCE/Publications/Validity>
- Stake, R. E. (1999). Representing quality in evaluation. Paper presented at the meeting of the American Educational Research Association. Montreal, Canada, April.
- Stake, R. E. & Cisneros-Cohernour, E. J. (2000). Situational evaluation of teaching on campus. New Directions for Teaching and Learning. Jossey-Bass Inc., Publishers, San Francisco: CA.
- Statham, A., Richardson, L., & Cook, J. A. (1991). Gender and university teaching: A negotiated difference. Albany, NY: State University of New York Press.
- Tagamory, H. T., & Bishop, L. A. (1995). Student evaluations of teaching: Flaws in the instruments, Thought & Action, 11, 63-78.
- Theall, M. (1997). On drawing reasonable conclusions about student ratings of instruction: A reply to Haskell and to Stake. Education Policy Analysis Archives. 5, <http://olam.ed.asu.edu/epaa/v5n8c2.html>
- Turner, J. L. & Boice, R. (1987). Starting at the beginning: The concerns and needs of new faculty. In J. Kurfiss, To improve the academy: Resources for student, faculty, and institutional development (p. 41-55). Stillwater, OK: New Forum Press.
- Washton, N. S. (1988). Developing a faculty checklist for self-evaluation for teaching effectiveness. Paper presented at the meeting of the Professional and Organizational Development Network in Higher Education, Keystone, CO.
- Wergin, J. F., Smith, A. & Rolle, G.E. (1980, April). An empirical analysis of faculty evaluation strategies: Implications for organizational change. Paper presented at the meeting of the American Educational Research Association, Boston, MA.
- Whitman, N. & Weiss, E. (1982). Faculty evaluation: The use of explicit criteria for promotion, retention and tenure (Higher Education Report No. 2). Washington, DC: AAHE-ERIC.

- Whooley, J. E. (1990). Tenured faculty practices that do/could enhance the teaching effectiveness of non-tenured faculty. Paper presented at the annual conference of the National Council of States on In-service Education, Orlando, FL.
- Wicks, S. (1992). Peer review and quality control in higher education. British Journal of Educational Studies, 40, 57-69.
- Wigington, H.; Tollefson, N., & Rodriguez, E. (1989). Students' ratings of instructors revisited: Interactions among class and instructor variables. Research in Higher Education, 30(3), 331-44
- Wilson, R. C. (1990). Commentary: The evaluation of a faculty developer. Journal of Educational Psychology, 82, 272-274.
- Wilson, R. (2001). It's 10 a.m. do you know where your professors are? Faculty members at Boston U. fume over plan to require them to be in their offices at least 4 days a week. The Chronicle of Higher Education, section: The Faculty, p. A10 - A12, February 2.
- Yen, W.M. (1998). Investigating the consequential aspects of validity: Who is responsible and what should they do? Educational Measurement: Issues and Practice, 17(2), 5-6.