

Fourteen journals (first and second tier) that you should be browsing and/or having available to you —

First tier:

*Applied Psychological Measurement* (SAGE)

*British Journal of Mathematical and Statistical Psychology* (BPS)

*Journal of Educational Measurement* (NCME)

*Journal of Educational (and Behavioral) Statistics* (AERA/ASA)

*Journal of Mathematical Psychology* (SMP)

*Multivariate Behavioral Research* (SMEP)

*Psychological Methods* (APA/Division 5)

*Psychometrika* (Psychometric Society)

Second tier:

*Applied Measurement in Education* (LEA)

*Educational Assessment* (LEA)

*Educational Measurement: Issues and Practice* (NCME)

*Educational and Psychological Measurement* (SAGE)

*Measurement* (LEA)

*Psychological Assessment* (APA)

Acronyms: AERA (American Educational Research Association); APA (American Psychological Association); BPS (British Psychological Society); LEA (Lawrence Erlbaum Associates); NCME (National Council on Measurement in Education); SMEP (Society for Multivariate Experimental Psychology); SMP (Society for Mathematical Psychology)

Some two-hundred books on test theory/measurement/assessment that could be on your book shelf (or keep your eyes open for these in the used book stores) —

Before 1970:

Brown, W. (1911). *The essentials of mental measurement*. Cambridge: Cambridge University Press.

Stern, W. (1914). *The psychological methods of testing intelligence*. Baltimore: Warwick and York.

Whipple, G. M. (1914;1915). *Manual of mental and physical tests*, Volumes I and II. Baltimore: Warwick and York.

Ayres, L. P. (1915). *A measuring scale for ability in spelling*. New York: Russell Sage Foundation.

Binet, A. & Simon, T. (1915). *A method of measuring the development of intelligence of young children* (3rd ed.). Chicago: Chicago Medical Book.

Yerkes, R. M., Bridges, J. W., & Hardwick, R. S. (1915). *A point scale for measuring mental ability*. Baltimore: Warwick and York.

Terman, L. M. (1919). *Measurement of intelligence*. London: Har-rap.

Thorndike, E. L. (1919). *An introduction to the theory of mental and social measurements* (2nd ed.). New York: Bureau of Publications, Teachers College, Columbia University.

Yoakum, C. S. & Yerkes, R. M. (1920). *Army mental tests*. New York: Henry Holt.

Brown, W. & Thomson, G. H. (1921). *The essentials of mental measurement* (2nd ed.). Cambridge: Cambridge University Press.

- McCall, W. A. (1922). *How to measure in education*. New York: The Macmillan Company.
- Kelley, T. L. (1923). *Statistical methods*. New York: The Macmillan Company.
- Monroe, W. S. (1923). *An introduction to the theory of educational measurements*. Boston: Houghton Mifflin Company.
- Monroe, W. S. (1923). *The theory of educational measurements*. Boston: Houghton Mifflin Company.
- Monroe, W. S., DeVoss, J. C., & Kelly, F. J. (1924). *Educational tests and measurements*. Boston: Houghton Mifflin Company.
- Vincent, L. E. (1924). *A study of intelligence test elements*. New York: Bureau of Publications, Teachers College, Columbia University.
- Brown, W. & Thomson, G. (1925). *The essentials of mental measurement* (3rd ed.). Cambridge: Cambridge University Press.
- Otis, A. S. (1925). *Statistical method in educational measurement*. New York: World Book Company.
- Abelson, H. H. (1927). *The improvement of intelligence testing*. New York: Bureau of Publications, Teachers College, Columbia University.
- Kelly, T. L. (1927). *The interpretation of educational measurements*. New York: World Book.
- Spearman, C. E. (1927). *The abilities of man, their nature and measurement*. New York: Macmillan.
- Symonds, P. M. (1927). *Measurement in secondary education*. New York: The Macmillan Company.

- Thorndike, E. L. (1927). *The measurement of intelligence*. New York: Bureau of Publications, Teachers College, Columbia University.
- Hull, C. L. (1928). *Aptitude testing*. New York: World Book Company.
- Ruch, G. M. (1929). *The objective or new type examination*. Chicago: Scott, Foresman and Company.
- Barthelme, H. M. (1931). *The validity of intelligence test elements*. New York: Bureau of Publications, Teachers College, Columbia University.
- Thurstone, L. L. (1931). *The reliability and validity of tests: Derivation and interpretation of fundamental formulae concerned with reliability and validity of tests and illustrative problems*. Ann Arbor, MI: Edwards Brothers.
- Brigham, C. C. (1932). *A study of error*. New York: College Entrance Examination Board.
- Thurstone, L. L. (1932). *The theory of multiple factors*. Ann Arbor, MI: Edwards Brothers.
- Smith, M. (1934). *The relationship between item validity and test validity*. New York: Bureau of Publications, Teachers College, Columbia University.
- Flanagan, J. C. (1935). *Factor analysis in the study of personality*. Palo Alto, CA: Stanford University Press.
- Thurstone, L. L. (1935). *Vectors of the mind*. Chicago: University of Chicago Press.
- Hawkes, H. E., Lindquist, E. F., & Mann, C. R. (1936). *The construction and use of achievement examinations*. Boston: Houghton

- Mifflin Company.
- Bingham, W. V. (1937). *Aptitudes and aptitude testing*. New York: Harper and Brothers.
- Orleans, J. S. (1937). *Measurement in education*. New York: Thomas Nelson and Sons.
- Terman, L. M. & Merrill, M. A. (1937). *Measuring intelligence*. Boston: Houghton Mifflin.
- Smith, B. O. (1938). *Logical aspects of educational measurement*. New York: Columbia University Press.
- Freeman, F. N. (1939). *Mental tests: Their history, principles, and applications* (Revised ed.). Cambridge, MA: The Riverside Press (Houghton Mifflin).
- Holzinger, K. J. & Swineford, F. (1939). *A study in factor analysis: The stability of the bi-factor solution*. Supplementary Educational Monographs, Vol. 48. Chicago: University of Chicago Press.
- McCall, W. A. (1939). *Measurement*. New York: Macmillan.
- Horst, P. (Ed.). (1941). *The prediction of personal adjustment*. New York: Social Science Research Council.
- Thomson, G. H. (1946). *The factorial analysis of human ability*. Boston: Houghton Mifflin Company.
- Adkins, D. C. *et al.* (1947). *Construction and analysis of achievement tests*. U. S. Government Printing Office.
- Kelly, T. L. (1947). *Fundamentals of statistics*. Cambridge, MA: Harvard University Press.
- Mursell, J. L. (1947). *Psychological testing*. New York: Longmans, Green and Company.

- Ross, C. C. (1947). *Measurement in today's schools*. New York: Prentice-Hall.
- Thurstone, L. L. (1947). *Multiple factor analysis*. Chicago: University of Chicago Press.
- Bray, C. W. (1948). *Psychology and military proficiency*. Princeton, NJ: Princeton University Press.
- Donahue, W. T., Coombs, C. H., & Travers, R. M. W. (1949). *The measurement of student adjustment and achievement*. Ann Arbor, MI: University of Michigan Press.
- Goodenough, F. L. (1949). *Mental testing, its history, principles, and applications*. New York: Rinehart and Company.
- Thorndike, R. L. (1949). *Personnel section (test and measurement techniques)*. New York: Wiley.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: Wiley. (Reprinted by LEA, 1987)
- Stouffer, S. A., Guttman, L., Suchman, E. A., Lazarsfeld, P. F., Star, S. A., & Clausen, J. A. (Eds.). (1950). *Measurement and prediction*. New York: Wiley.
- Eells, K., Davis, A., Havighurst, R., Herrick, V., & Tyler, R. (1951). *Intelligence and cultural differences*. Chicago: University of Chicago Press.
- Lindquist, E. F. (Ed.). (1951). *Educational measurement*. Washington, DC: American Council on Education.
- Spearman, C.E. & Jones, L. W. (1951). *Human abilities*. London: Macmillan.
- Stevens, S. S. (Ed.). (1951). *Handbook of experimental psychology*. New York: Wiley.

- Thomson, G. H. (1951). *The factorial analysis of human ability*. London: London University Press.
- Lindquist, E. F. (1953). *Design and analysis of experiments in psychology and education*. Boston: Houghton Mifflin.
- Stephenson, W. (1953). *The study of behavior*. Chicago: University of Chicago Press.
- Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). New York: McGraw-Hill. (Original published in 1936)
- Torgersson, W. S. (1958). *Theory and methods of scaling*. New York: Wiley.
- Wechsler, D. W. (1958). *The measurement and appraisal of adult intelligence* (4th ed.). Baltimore, MD: Williams & Wilkins.
- Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.
- Thurstone, L. L. (1959). *The measurement of values*. Chicago: The University of Chicago Press.
- Gulliksen, H. & Messick, S. (Eds.). (1960). *Psychological scaling: Theory and applications*. New York: Wiley.
- Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danish Institute for Educational Research. (An expanded edition with forward and afterward by B. D. Wright published in 1980, University of Chicago Press)
- Harris, C. W. (Ed.). (1963). *Problems in measuring change*. Madison, WI: University of Wisconsin Press.
- Horst, P. (1963). *Matrix algebra for social scientists*. New York: Holt, Rinehart, & Winston.
- Coombs, C. H. (1964). *A theory of data*. New York: Wiley.

- Cronbach, L. & Gleser, G. (1965). *Psychological tests and personnel decisions* (2nd ed.). Urbana, IL: University of Illinois Press.
- Horst, P. (1965). *Factor analysis of data matrices*. New York: Holt, Rinehart & Winston.
- Ghiselli, E. E. (1966). *The validity of occupational aptitude tests*. New York: Wiley.
- Magnusson, D. (1967). *Test theory*. Reading, MA: Addison-Wesley.
- Lazarsfeld, P. F. & Henry, N. W. (1968). *Latent structure analysis*. Boston: Houghton Mifflin.
- Lord, F. M. & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Before 1990:

- Bormuth, J. R. (1970). *On the theory of achievement test items*. Chicago: University of Chicago Press.
- Dubois, P. H. (1970). *A history of psychological testing*. Boston: Allyn & Bacon.
- Mokken, R. J. (1970). *A theory and procedure of scale analysis*. The Hague: Mouton.
- Lawley, D. N. & Maxwell, A. E. (1971). *Factor analysis as a statistical method* (2nd ed.). London: Butterworth.
- Thorndike, R. L. (Ed.). (1971). *Educational measurement* (2nd ed.). Washington, DC: American Council on Education.
- Cronbach, L. J., Gleser, G. C., Nanda, H., & Rajaratnam, N. (1972). *The dependability of behavioral measurements: Theory of generalizability for scores and profiles*. New York: Wiley.



- Ebel, R. L. (1972). *Essentials of educational measurement*. Englewood Cliffs, NJ: Prentice-Hall.
- Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.
- Harman, H. H. (1976). *Modern factor analysis* (3rd ed.). Chicago: University of Chicago Press. (Original edition published 1960)
- Popham, W. J. (1978). *Criterion-referenced measurement*. Englewood Cliffs, NJ: Prentice Hall.
- Allen, M. J. & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole. (Reissued 2002; Long Grove, IL: Waveland Press)
- Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.
- Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.
- van der Kamp, L. J. T., Langerak, W. F., de Gruijter, D. N. M. (Eds.). (1980). *Psychometrics for educational debates*. Chichester: Wiley.
- Carlson, D. (Ed.). (1981). *New directions for testing and measurement: Testing in the states*. San Francisco, CA: Jossey-Bass.
- Anastasi, A. (1982). *Psychological testing* (5th ed.). New York: Macmillan. (First edition published 1954).
- Berk, R. A. (Ed.). (1982). *Handbook for detecting biased test items*. Baltimore: Johns Hopkins University Press.
- Holland, P. W. & Rubin, D. B. (Eds.). (1982). *Test equating*. New York: Academic Press.
- Wigdor, A. K. & Garner, W. R. (Eds.). (1982). *Ability testing: Uses, consequences and controversies, Part II*. Washington, DC: National Academy Press.

- Anderson, S. B. & Helmick, J. S. (Eds.). (1983). *On educational testing*. San Francisco: Jossey-Bass.
- Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic Books.
- Gorsuch, R. L. (1983). *Factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum. (Original published 1974)
- Hambleton, R. K. (1983). *Applications of item response theory*. Vancouver: Educational Research Institute of British Columbia.
- Hulin, C. L., Drasgow, F., & Parsons, C. K. (1983). *Item response theory: Applications to psychological measurement*. Homewood, IL: Dow-Jones Irwin.
- Peterson, J. J. (1983). *The Iowa testing programs: The first fifty years*. Iowa City, IA: University of Iowa Press.
- Wainer, H. & Messick, S. (Eds.). (1983). *Principals of modern psychological measurement: A Festschrift for Frederic M. Lord*. Hillsdale, NJ: Erlbaum.
- Weiss, D. J. (Ed.). (1983). *New horizons in testing: Latent trait theory and computer adaptive testing*. New York: Academic Press.
- Berk, R. A. (Ed.). (1984). *A guide to criterion-reference test construction*. Baltimore, MD: Johns Hopkins University Press.
- Berk, R. A. (1984). *Screening and diagnosis of children with learning disabilities*. Springfield, IL: Charles C. Thomas.
- Cronbach, L. L. (1984). *Essentials of psychological testing* (4th ed.). New York: Harper and Row. (First edition published 1949)
- Andrich, D. (1985). *A latent-trait model for items with response dependencies: Implications for test construction and analysis*. Orlando, FL: Academic Press.

- Embretson, S. E. (Ed.). (1985). *Test design: Developments in psychology and psychometrics*. New York: Academic Press.
- Hambleton, R. K. & Swaminathan, H. (1985). *Item response theory: Principles and applications*. Boston: Kluwer.
- Holland, J. L. (1985). *Making vocational choices: A theory of vocational personalities and work environments*. Englewood Cliffs, NJ: Prentice-Hall.
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, & Winston.
- Plake, B. S. & Witt, J. C. (Eds.). (1986). *The future of testing*. Hillsdale, NJ: Erlbaum.
- Sternberg, R. & Detterman, D. (Eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex Publishing.
- Langeheine, R. & Rost, J. (Eds.). (1988). *Latent trait and latent models*. New York: Plenum Press.
- Wainer, H. & Braun, H. I. (Eds.). (1988). *Test validity*. Hillsdale, NJ: Erlbaum.
- Linn, R. L. (Ed.). (1989). *Educational measurement* (3rd ed.). New York: Macmillan.
- Linn, R. L. (Ed.). (1989). *Intelligence: Measurement, theory, and public policy*. Urbana, IL: University of Illinois Press.

After 1990:

- Frederiksen, N., Glaser, R. L., Lesgold, A. M., & Shafto, M. G. (Eds.). (1990). *Diagnostic monitoring of skill and knowledge acquisition*. Hillsdale, NJ: LEA.

- Hopkins, K. S., Stanley, J. C., & Hopkins, B. R. (1990). *Educational and psychological measurement and evaluation* (7th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Reynolds, C. R. & Kamphaus, R. W. (Eds.). (1990). *Handbook of psychological and educational assessment of children: Intelligence and achievement*. New York: Guilford Press.
- Goldberg, L. R., Grenier, J. R., Guion, R. M., Sechrest, L. B., & Wing, H. (1991). *Questionnaires used in the prediction of trustworthiness in pre-employment selection decisions*. Washington, DC: American Psychological Association.
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.
- Mehrens, W. A. & Lehmann, I. J. (1991). *Measurement and evaluation in education and psychology* (4th ed.). Fort Worth, TX: Holt, Rinehart, & Winston.
- Shavelson, R. J. & Webb, N. M. (1991). *Generalizability theory*. Newbury Park, CA: Sage.
- Snow, R. E. & Wiley, D. E. (Eds.). (1991). *Improving inquiry in social science: A volume in honor of Lee J. Cronbach*. Hillsdale, NJ: Erlbaum.
- Wigdor, A. K. & Green, B. F. (Eds.). (1991). *Performance assessment in the workplace: Volume I*. Washington, DC: National Academy Press.
- Bennett, R. E. & Ward, W. C. (Eds.). (1993). *Construction versus choice in cognitive measurement*. Hillsdale, NJ: Erlbaum.
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies*. New York: Cambridge University Press.

- Frederiksen, N., Mislevey, R. J., & Bejar, I. I. (Eds.). (1993). *Test theory for a new generation of tests*. Hillsdale, NJ: Erlbaum.
- Hagenaars, J. A. (1993). *Loglinear models with latent variables*. Newbury Park: Sage.
- Holland, P. W. & Wainer, H. (Eds.). (1993). *Differential item functioning*. Hillsdale, NJ: LEA.
- Schuler, H., Farr, J. L., & Smith, M. (Eds.). (1993). *Personnel selection and assessment: Individual and organizational perspectives*. Hillsdale, NJ: Erlbaum.
- Shepard, L. A., Glaser, R., Linn, R., & Bohnstedt, G. (1993). *Setting performance standards for student achievement*. Stanford, CA: National Academy of Education.
- Camilli, G. & Shepard, L. A. (1994). *Methods for identifying biased test items*. Thousand Oaks, CA: Sage.
- Fischer, G. H. & Laming, D. (Eds.). (1994). *Contributions to mathematical psychology, psychometrics, and methodology*. New York: Springer.
- Hernstein, R. J. & Murray, C. (1994). *The bell curve: Intelligence and class structure in American life*. New York: The Free Press.
- Laveault, D., Zumbo, B. D., Gessaroli, M. E., & Boss, M. W. (Eds.). (1994). *Modern theories of measurement: Problems and issues*. Ottawa, Canada: University of Ottawa.
- Nunnally, J. & Bernstein, I. (1994). *Psychometric theory* (3rd ed.). New York: McGraw-Hill. (1st ed. 1967; 2nd ed. 1978).
- Fischer, G. & Molenaar, I. (Eds.). (1995). *Rasch models: Foundations, recent developments, and applications*. New York: Springer.

- Linn, R. L. & Gronlund, N. E. (1995). *Measurement and assessment in teaching* (7th ed.). Englewood Cliffs, NJ: Prentice-Hall.
- Nichols, P. D., Chipman, S. F., & Brennan, R. L. (Eds.). (1995). *Cognitively diagnostic assessment*. Hillsdale, NJ: Erlbaum.
- Gould, S. J. (1996). *The mismeasure of man: Revised and expanded*. New York: W. W. Norton. (Original published 1981)
- Heinen, T. (1996). *Latent class and discrete latent trait models: Similarities and differences*. Thousand Oaks, CA: Sage.
- Phillips, G. W. & Goldstein, A. (Eds.). (1996). *Technical issues in large-scale performance assessment*. Washington, DC: National Center for Education Statistics.
- Samejima, F. (Ed.). (1996). *The graded response model*. New York: Springer.
- van der Linden, W. J. & Hambleton, R. K. (Eds.). (1997). *Handbook of modern item response theory*. New York: Springer.
- Marley, A. A. J. (Ed.). (1997). *Choice, decision, and measurement: Essays in honor of R. Duncan Luce*. Mahwah, NJ: LEA.
- Rost, J. & Langeheine, R. (Eds.). (1997). *Applications of latent trait and latent class models in the social sciences*. Münster: Waxmann.
- Sands, W. A., Waters, B. K., & McBride, J. R. (Eds.). (1997). *Computerized adaptive testing: From inquiry to operation*. Washington, DC: American Psychological Association.
- Willingham, W. W. & Cole, N. S. (Eds.). (1997). *Gender and fair assessment*. Mahwah, NJ: LEA.
- Guion, R. (1998). *Assessment, measurement, and prediction for personnel decisions*. Mahwah, NJ: LEA.

- Hakel, M. D. (Ed.). (1998). *Beyond multiple choice: Evaluating alternatives for traditional testing for selection*. Mahwah, NJ: LEA.
- Osterlind, S. J. (1998). *Constructing test items: Multiple-choice, constructed-response, performance, and other formats*. New York: Kluwer Academic.
- Pellegrino, J. W., Jones, I. R., & Mitchell, K. J. (Eds.). (1998). *Grading the nation's report card: Evaluating NAEP and transforming the assessment of educational progress*. Washington, DC: National Academy Press.
- Zenderland, L. (1998). *Measuring minds: Henry Herbert Goddard and the origins of American intelligence testing*. Cambridge: Cambridge University Press.
- AERA, APA, & NCME (1999). *Standards for educational and psychological testing*. Washington, DC: AERA.
- Cizek, G. J. (1999). *Cheating on tests: How to do it, detect it, and prevent it*. Mahwah, NJ: LEA.
- Drasgow, F. & Olson-Buchanan, J. B. (Eds.). (1999). *Innovations in computerized assessment*. Mahwah, NJ: Erlbaum.
- Embretson, S. E. & Hershberger, S. L. (Eds.). (1999). *The new rules of measurement: What every psychologist and educator should know*. Mahwah, NJ: Erlbaum.
- Feuer, M. J., Holland, P. W., Green, B. F., Bertentahl, M. W., & Hemphill, F. C. (Eds.). (1999). *Uncommon measures: Equivalence and linkage among educational tests*. Washington, DC: National Research Council.
- Heubert, J. P. & Hauser, M. H. (1999). *High stakes: Testing for*

- tracking, promotion, and graduation*. Washington, DC: National Academy Press.
- Koretz, D. M., Bertenthal, M. W., & Green, B. F. (Eds.). (1999). *Embedding questions: The pursuit of a common measure in uncommon tests*. Washington, DC: National Academy Press.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: LEA.
- Michell, J. (1999). *Measurement in psychology: Critical history of a methodological concept*. Cambridge: Cambridge University Press.
- Rust, J. & Golombok, S. (1999). *Modern psychometrics: The science of psychological assessment* (2nd ed.). London: Routledge.
- Embretson, S. E. & Reise, S. P. (2000). *Item response theory for psychologists*. Mahwah, NJ: LEA.
- Engelhard, G. & Wilson, M. (Eds.). (2000). *Objective measurement V: Theory into practice*. Norwood, NJ: Ablex Publishing.
- van der Linden, W. J. & Glas, C. A. W. (Eds.). (2000). *Computerized adaptive testing: Theory and practice*. Boston: Kluwer Academic.
- Sternberg, R. (Ed.). (2000). *Handbook of intelligence*. Cambridge: Cambridge University Press.
- Wainer, H. (Ed.). (2000). *Computerized adaptive testing: A primer* (2nd ed.). Mahwah, NJ: LEA.
- Boomsma, A., van Duijn, M. A. J., & Snijders, T. A. B. (Eds.). (2001). *Essays in item response theory*. New York: Springer.
- Brennan, R. L. (2001). *Generalizability theory*. New York: Springer.
- Cizek, G. (2001). *Standard setting: Concepts, methods, and perspectives*. Mahwah, NJ: LEA.



- Pellegrino, J. W., Chudowsky, N. E., & Glaser, R. (Eds.). (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.
- Thissen, D., Nelson, L., Rosa, K., & McLeod, L. D. (2001). *Item response theory for items scored in more than two categories*. Mahwah, NJ: LEA.
- Thissen, D. & Wainer, H. (Eds.). (2001). *Test scoring*. Mahwah, NJ: LEA.
- Irvine, S. H. & Kyllonen, P. C. (Eds.). (2002). *Item generation for test development*. Mahwah, NJ: LEA.
- Lissitz, R. W. & Shafer, W. D. (Eds.). (2002). *Assessment in educational reform*. Boston: Allyn and Bacon.
- Mills, C. N., Potenza, M. T., Fremer, J. J., & Ward, W. C. (Eds.). (2002). *Computer-based testing: Building the foundation for future assessment*. Mahwah, NJ: LEA.
- Parshall, C. G., Spray, J. A., Kalohn, J. & Davey, T. (2002). *Practical considerations in computer-based testing*. New York: Springer.
- Sijtsma, K. & Molenaar, I. (2002). *Introduction to nonparametric item response theory*. Thousand Oaks, CA: Sage.
- Tate, R. (2002). *Test dimensionality*. Mahwah, NJ: LEA.
- Chang, M., Witt, D., Haikuta, K., & Jones, J. (Eds.). (2003). *Compelling interest: Examining the evidence on racial dynamics in higher education in colleges and universities*. Stanford, CA: Stanford University Press.
- Murphy, K. (2003). *Validity generalization: A critical review*. Mahwah, NJ: LEA.

- Shermis, M. D. & Burstein, J. C. (Eds.). (2003). *Automated essay scoring*. Mahwah, NJ: LEA.
- von Davier, A. A., Holland, P. W., & Thayer, D. T. (2003). *The kernel method of test equating*. New York: Springer.
- Baker, F. & Kim, S. (2004). *Item response theory: Parameter estimation techniques*. New York: Marcel Dekker. (Revised edition; original published 1992)
- De Boeck, P. & Wilson, M. (2004). *Explanatory item response models: A generalized linear and nonlinear approach*. New York: Springer.
- Haladyna, T. M. (2004). *Developing and validating multiple-choice test items* (3rd ed.). Mahwah, NJ: LEA.
- Kaplan, D. (Ed.). (2004). *The Sage handbook of quantitative methodology for social sciences*. Thousand Oaks, CA: Sage.
- Kolen, M. J. & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). New York: Springer.
- Loehlin, J. (2004). *Latent variable models: An introduction to factor, path, and structural analysis* (4th ed.). Hillsdale, NJ: LEA.
- Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. Cambridge: Cambridge University Press.
- Dwyer, C. A. (Ed.). (2005). *Measurement and research issues in a new accountability era*. Mahwah, NJ: LEA.
- van der Linden, W. (2005). *Linear models for optimal test design*. New York: Springer.
- Phelps, R. (Ed.). (2005). *Defending standardized testing*. Mahwah, NJ: LEA.

- von Davier, M. & Carstensen, C. H. (2005). *Multivariate and mixture distribution Rasch models: Extensions and applications*. New York: Springer.
- Wilson, M. (2005). *Constructing measures: An item response modeling approach*. Mahwah, NJ: LEA.
- Brennan, R. (Ed.). (2006). *Educational measurement* (4th ed.). Washington, DC: American Council on Education and NCME.
- Downing, S. M. & Haladyna, T. M. (Eds.). (2006). *Handbook of test development*. Mahwah, NJ: LEA.
- Rao, C. R. & Sinharay, S. (Eds.). (2006). *Handbook of statistics: Psychometrics, 26*. North-Holland: Elsevier.
- Williamson, D. M., Mislevy, R. J., & Bejar, I. I. (Eds.). (2006). *Automated scoring of complex tasks in computer-based testing*. Mahwah, NJ: LEA.
- Dorans, N. J. Pommerich, M., & Holland, P. W. (Eds.). (2007). *Linking and aligning scores and scales*. New York: Springer.