

Rebecca Zwick
Educational Testing Service
rzwick@ets.org
805-680-8356

Education

Rutgers University Statistics Department	M.S., Statistics; 1989
Graduate School of Education, University of California, Berkeley	Ph.D., Quantitative Methods; 1983 M.A., Quantitative Methods; 1981
Antioch College, Ohio	B.A., Psychology and Education; 1974

Current Position

11/09-present Distinguished Presidential Appointee, Statistical Analysis, Data Analysis, and Psychometric Research, Educational Testing Service

- Serve as co-director of research initiative in Foundational and Applied Statistics and Psychometrics
- Conduct research on college admissions, test fairness, differential item functioning, score reporting
- Serve as associate editor for ETS research review process
- Served as psychometric advisor to National Assessment of Educational Progress (NAEP) and Cognitively Based Assessment of, for and as learning (CBAL) programs

**1996-2010 Professor, Department of Education, U.C., Santa Barbara
Affiliate, Department of Psychology**

- Research Methodology Emphasis Leader (1996-2009); developer and teacher of courses in statistics, psychometrics, test fairness, and research methods; methodology consultant for students and faculty
- Cofounder and Coordinating Committee member (1999-2009), Quantitative Methods in the Social Sciences, an interdisciplinary Ph.D. emphasis involving 10 UCSB departments, established 1999

Current National Activities in Education, Statistics, and Psychology

- Editorial Board, National Council on Measurement in Education Book Series (2011-)
- Associate Editor, Journal of Educational and Behavioral Statistics (2010-)

Honors and Awards

- Fellow of the American Educational Research Association (2016)
- Fellow of the American Statistical Association (2012)
- Outstanding Reviewer Award, Journal of Educational and Behavioral Statistics (2011 and 2013)
- ETS Spot Award for contributions to technical review of research publications (2011, 2013, 2015)
- National Council on Measurement in Education Award for Outstanding Dissemination of Educational Measurement Concepts to the Public, for "insightful...balanced scholarship...on standards and high stakes testing and use of the SAT..." (2001)
- Fellow of the Division of Measurement, Evaluation, and Statistics (Division 5) of the American Psychological Association (elected 1992)

Professional Affiliations

- American Educational Research Association (Fellow)
- American Statistical Association (Fellow)
- National Council on Measurement in Education
- Psychometric Society
- Association for the Study of Higher Education

Past Professional Activities

Professional Offices

- Vice-President for Division D (Measurement and Research Methodology) of the American Educational Research Association (AERA) and member of AERA Council (2002-4)
- Member, Executive Board of AERA (2003-4)
- Board of Directors, National Council on Measurement in Education (NCME; 2000-3)
- President (1995-6); Program Chair, AERA annual meeting (1995); Secretary Treasurer (1988-1990); Educational Statisticians Special Interest Group (SIG) of AERA:

Committees

- Board on Testing and Assessment of the National Academy of Sciences (2009-2015)
- Member, Robert L. Linn Distinguished Address Award Committee of AERA Division D (2010-2012)
- Technical Advisory Group, Programme for International Student Assessment (PISA) (2006-12)
- National Science Foundation site visit team, Pittsburgh Science of Learning Center, 2010
- College Board Psychometric Panel for the SAT/PSAT (2006-2010)
- Design and Analysis Committee, National Assessment of Educational Progress (2000-2010)
- National Research Council/Government Accountability Office expert panel on the assessment of English-Language Learners, University of California, Davis, January 2006.
- National Academy of Sciences Committee on Participation of English-Language Learners and Students with Disabilities in the National Assessment of Educational Progress (2003-4)
- Committee on Psychological Tests & Assessment of the American Psychological Association (2002-4)
- AERA News Media Interactions Committee: Founder and Chair (2002-3), Member (2003-4)
- AERA Ad Hoc Awards Committee (2002)
- NCME Publications Committee (Chair, 2001-3)
- College Board SAT Committee: Chair (2001-2003), Member (1999-2000)
- Organizer, moderator, and presenter: UCSB conference, "Rethinking the SAT," November 2001; editor of a 2004 book based on the conference
- NCME Graduate Student Issues Committee: Faculty Advisor (1999-2001)
- American Statistical Association Management Committee for the *Journal of Educational and Behavioral Statistics* (1999-2001).
- National Academy of Sciences Committee on the Technical Feasibility of Embedding Items in State and District Tests (1999)
- Technical Advisory Committee on Standard-Setting, advising ACT, Inc. and the National Assessment Governing Board on NAEP standard-setting (1997-2000)
- Board of Directors, Gevirtz Research Center, a joint research effort of UCSB's Graduate School of Education and the Santa Barbara Public Schools (1997 to 2002)
- Defense Advisory Committee on Military Personnel Testing (1991-1994)

Editing

- Advisory Board Member, *Educational Measurement: Issues and Practice* (2003-2007)
- Advisory Editor, *Journal of Educational Measurement* (2003-2007)
- Associate Editor, *Journal of Educational and Behavioral Statistics* (1991-2002)
- Editor, *Journal of Educational Measurement* (1995-1998)
- Editor, special issue of *Journal of Educational Statistics* on the National Assessment of Educational Progress (1992)
- Consulting Editor, *Psychological Assessment* (1988-1991)
- Consulting Editor, *Journal of Consulting and Clinical Psychology* (1985-1989)

Consulting

- Consultant, Research and Development Division, Educational Testing Service, (2/08-10/09)
- Technical Design Team, National Assessment of Adult Literacy (as a consultant to the American Institutes for Research, 1999-2000)
- Senior Fellow, Consortium of Universities (providing consultation to the US Defense Dept. (1998-2000)
- Consultant and workshop presenter on psychometrics and statistics, Clinical Services Research Training Program, UC San Francisco (1986-2000)

Postdoctoral Work Experience

September 1984 to June 1996 **Division of Statistics and Psychometrics Research,
Educational Testing Service, Princeton, NJ:**

12/95-6/96 **Principal Research Scientist, Research Statistics Group, ETS**

- Proposed and directed projects on the statistical assessment of test fairness, e.g.,
A Bayesian approach to the assessment of differential item functioning (DIF)
Describing and categorizing DIF in polytomous items
- Provided statistical consultation to ETS researchers and staff from testing programs

7/91-12/95 **Senior Research Scientist, Research Statistics Group, ETS**

- Proposed, obtained funding for, and directed the following projects:
Performance of graduates and undergraduates on the Test of Written English
A simulation study of methods of assessing DIF in computer-adaptive tests
- Provided statistical consultation to ETS researchers, officers, and staff from testing programs, the human resources department, and the legal department

1/90-7/91 **Director, Data Analysis and Scale Development for the National
Assessment of Educational Progress (NAEP), ETS**

- Designed and directed NAEP data analyses; contributed to NAEP funding proposals
- Hired and trained statistical staff; supervised four Ph.D.-level researchers
- Communicated results to ETS management, national panels, & U.S. Department of Education

9/84-12/89 **Research Scientist, Psychometrics Research Group, ETS**

- Designed and conducted analyses of NAEP data; contributed to proposals
- Conducted research on pathways through graduate school for minorities, time to degree in graduate school, and the validity of the GMAT for the prediction of success in doctoral study

1983-1984 **NIMH Postdoctoral Fellow, L. L. Thurstone Psychometric Laboratory,
University of North Carolina at Chapel Hill**

Research Grants (UC Santa Barbara)

- 2008-10 Comparing the Utility of SAT Scores and High School GPA in Predicting College Grades: Why Correlation Coefficients Don't Tell the Whole Story (The College Board, \$89,733)
- 2004-08 Instructional Modules in Educational Measurement and Statistics for School Personnel (National Science Foundation, \$477,988)
- 2002-04 California and the SAT: A reanalysis of University of California admissions data (Center for Studies in Higher Education, UC Berkeley, \$10,000)
- 2002-03 Special grant for editing a book based on UCSB's conference, "Rethinking the SAT" (UCSB Academic Senate, \$14,500)
- 2002 Special grant for analysis of UC admissions data for UCSB "Town Meeting" on SAT use at UC (UCSB Academic Senate, \$2,000)
- 2001-04 Why do standardized tests "overpredict" the grades of under-represented minorities? (UCSB Academic Senate Research Committee, \$4,600)
- 2001-03 SAT validity for linguistic minorities: An investigation based on regression and survival analyses (American Educational Research Assoc. Grants Program, \$25,000)
- 2000-01 New perspectives on SAT validity. (UCSB Academic Senate Research Committee, \$3,500)
- 1999-02 SAT validity for linguistic minorities (UC Linguistic Minority Research Institute, \$5,000)
- 1999-01 Grant to support development of the Quantitative Methods in the Social Sciences Ph.D. emphasis (with N. Friedkin) (UCSB Research Across the Disciplines Program, \$12,469)
- 1998-99 An investigation of alternative methods for scale anchoring and item mapping in the National Assessment of Educational Progress (ACT, Inc., \$24,900)
- 1998-99 A discussion of academic admissions issues for the general public and for academic audiences (UCSB Academic Senate Research Committee, \$2800)
- 1996-98 A Bayesian enhancement of Mantel-Haenszel differential item functioning analysis for computer-adaptive tests (Law School Admission Council, \$50,100)

Books and Book Chapters

- Zwick, R. (to appear). *Who gets in? Strategies for fair and effective college admissions*, Harvard University Press.
- Zwick, R. (to appear). Fairness issues in the assessment of college and career readiness. In K. McClarty, K. Mattern, & M. Gaertner (Eds.), *Preparing Students for College and Careers: Theory, Measurement, and Educational Practice*. Routledge.
- Zwick, R. (to appear). Admissions testing in college and graduate education. In C. Secolsky and B. Denison (Eds.), *The handbook on measurement, assessment, and evaluation in higher education* (2nd ed.), Routledge.
- Bennett, R., Zwick, R., & van Rijn, P. (to appear). Measuring well and modeling good teaching and learning practice: A review of CBAL research. In H. Jiao and R. W. Lissitz (Eds.), *Technology enhanced innovative assessment: Development, modeling, and scoring from an interdisciplinary perspective...*
- Zwick, R., & Dorans, N. J. (2016). Philosophical perspectives on fairness in educational assessment. In N. J. Dorans and L. Cook (Eds.), *Fairness in educational assessment and measurement*, Routledge.
- Zwick, R., & Bridgeman, B. (2014). Evaluating validity, fairness, and differential item functioning in multistage testing. In D. Yan, A. von Davier, & C. Lewis (Eds.), *Computerized multistage testing: Theory and applications*, CRC Press.
- Zwick, R. (2012.) El rol de los exámenes de admisión, de las notas de educación secundaria y de la situación socioeconómica en la predicción de desempeño universitario. In M. V. Santelices, J. J. Ugarte, & P. Kyllonen (Eds.), *Admisión a la Educación superior: Mediciones complementarias*, pp. 25-37, Ministerio de Educación de Chile.
- Zwick, R. (2010). Admissions Testing. In P. Peterson, E. Baker, B. McGaw, (Eds.), *International Encyclopedia of Education*, volume 4, pp. 7-14, Elsevier.
- Zwick, R. (2010). The investigation of differential item functioning in computerized adaptive tests (revised and updated), pp. 331-352. In W. J. van der Linden & C. A. W. Glas, Eds., *Elements of Adaptive Testing*, Springer.
- Zwick, R. (2006). *Higher education admissions testing*. In R. Brennan (Ed.), *Educational Measurement* (4th ed.), pp. 647-679, National Council on Measurement in Education/Praeger.
- Zwick, R., (2004). *Rethinking the SAT: The Future of Standardized Testing in University Admissions* [Editor and author of preface, pp. ix-xvi.], RoutledgeFalmer.
- Zwick, R., (2004). Is the SAT a "wealth test?" The link between educational achievement and socioeconomic status. In R. Zwick (Ed.), *Rethinking the SAT*, pp. 203-216, RoutledgeFalmer.
- Zwick, R. (2002). *Fair Game? The Use of Standardized Admissions Tests in Higher Education*, RoutledgeFalmer.
- Zwick, R. (2000). The assessment of differential item functioning in computer-adaptive tests. In W. J. van der Linden & C. A. W. Glas, Eds., *Computerized Adaptive Testing: Theory and Practice*, Kluwer.
- Zwick, R. (1993). Pairwise comparison procedures for one-way analysis of variance designs. Invited chapter in G. Keren & C. Lewis (Eds.), *A handbook for data analysis in the behavioral sciences: Statistical issues*, pp. 43-71, Erlbaum.

Journal Publications

- Zapata-Rivera, D., Zwick, R., & Vezzu, M. (to appear). Exploring the effectiveness of a measurement error tutorial in helping teachers understand score report results. *Educational Assessment*.
- Zwick, R. (2016, January). Transparency in college admissions is key to a fair policy on race. *Chronicle of Higher Education*. <http://chronicle.com/article/Transparency-in-College/234949>
- Zwick, R., Zapata-Rivera, D., & Hegarty, M. (2014). Comparing graphical and verbal representations of measurement error in test score reports. *Educational Assessment*, 19, 116-138.
- Zwick, R. (2012, October). The role of admissions test scores, socioeconomic status, and high school grades in predicting college achievement. *Pensamiento Educativo: Revista de Investigación Educativa Latinoamericana*. www.pel.cl.
- Zwick, R., Ye, L., & Isham, S. (2012). Improving Mantel-Haenszel DIF estimation through Bayesian updating. *Journal of Educational and Behavioral Statistics*, 37, 601-629.
- Mislevy, R. J., & Zwick, R. (2012). A framework for scaling, linking, and reporting periodic assessments. *Journal of Educational Measurement*, 49, 148-166.
- Zwick, R., & Himelfarb, I. (2011). The effect of high school socioeconomic status on the predictive validity of SAT scores and high school grade-point average. *Journal of Educational Measurement*, 48, 101-121.
- Sinharay, S., Haberman, S., & Zwick, R. (2010). Issues with self-monitoring assessments: Comments on Koretz and Beguin. *Measurement: Interdisciplinary Research and Perspectives*, 8, 191-194.
- Sklar, J., & Zwick, R. (2009). Multimedia presentations in educational measurement and statistics: Design considerations and instructional approaches, *Journal of Statistics Education*, <http://www.amstat.org/publications/jse/v17n3/sklar.html>.
- Zwick, R. & Lenaburg, L. (2009). Using discrete loss functions and weighted kappa for classification: An illustration based on Bayesian network analysis. *Journal of Educational and Behavioral Statistics*, 34, 190-200.
- Zwick, R., Sklar, J., Wakefield, G., Hamilton, C., Norman, A., & Folsom, D. (2008). Instructional tools in educational measurement and statistics (ITEMS) for school personnel: Evaluation of three web-based training modules *Educational Measurement: Issues and Practice*, 27, pp. 14-27.
- Zwick, R. (2007). College admissions in 21st century America: The role of grades, tests, and games of chance. *Harvard Educational Review*, 77, pp. 419-428.
- Zwick, R. (2007, December). Casting lots for college. Online Education Week commentary at www.edweek.org/go/zwick. Also appeared in the *Santa Barbara Independent*, www.independent.com.
- Zwick, R., & Green, J. G. (2007). New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. *Journal of Educational Measurement*, 44, 23-45.
- Zwick, R. (2006, Fall). The future of college admissions testing. *College Board Review*, pp. 14-17.
- Zwick, R. (2006, Summer). Left behind: Many disadvantaged students are 'hidden in averages,' *National CrossTalk*, Summer, pp. 12-13.
- Zwick, R., & Sklar, J. (2005). Predicting college grades and degree completion using high school grades and SAT Scores: The role of student ethnicity and first language. *American Educational Research Journal*, 42, 439-464.

Journal Publications (continued)

- Zwick, R., & Sklar, J. (2005). A note on standard errors for survival curves in discrete-time survival analysis. *Journal of Educational and Behavioral Statistics*, 30, 75-92.
- Zwick, R. (2004, Fall). Assessment literacy: Do educators know how to make use of the new avalanche of standardized test data? *National CrossTalk*, pp. 13-14.
- Zwick, R., & Schlemmer, L. (2004). SAT validity for linguistic minorities at the University of California, Santa Barbara. *Educational Measurement: Issues and Practice*, 25, 6-16, 2004.
- Zwick, R. (2003, Fall). The anti-test backlash. *National CrossTalk*, pp. 11-12.
- Zwick, R. (2003, Winter). What do we learn from high school exit exams? *National CrossTalk*, pp. 13-14.
- Zwick, R. (2002, December). Is the SAT a "wealth test?" *Phi Delta Kappan*, pp. 307-311.
- Zwick, R., & Thayer, D. T. (2002). Application of an empirical Bayes enhancement of Mantel-Haenszel DIF analysis to a computerized adaptive test. *Applied Psychological Measurement*, 26, 57-76.
- Zwick, R. (2002, Summer). Revamping the SAT: Will the modified test make the grade? *National CrossTalk*, 12-13.
- Zwick, R. (2002, Winter). Picking the perfect freshman class: Balancing the academic and nonacademic goals of admissions policy. *National CrossTalk*, p. 13.
- Zwick, R. (2001, Oct. 29). Making the grade: The SAT vs. the GPA, *San Francisco Chronicle*, p. A17.
- Zwick, R. (2001, Fall). Measuring academic potential: Is the SAT II the answer to the college admissions dilemma? *National CrossTalk*, pp. 12-13.
- Zwick, R. (2001, Summer). Making the grade: The SAT versus the GPA. *National CrossTalk*, p. 10.
- Zwick, R., Senturk, D., Wang, J., & Loomis, S. (2001). An investigation of alternative methods for item mapping in the National Assessment of Educational Progress. *Educ. Measurement: Issues and Practice*, 20, 15-25.
- Zwick, R. (2001, Spring). Can a thermometer cure a fever? The role of testing in educational reform. *National CrossTalk*, p. 12.
- Zwick, R. (2001, March/April). What causes the test-score gap in higher education? Perspectives on the Office for Civil Rights resource guide on high-stakes testing. *Change Magazine*, pp. 33-37.
- Zwick, R. (2000, Dec. 17). The standard bearers. *San Francisco Chronicle*, Sunday section, pp. 1, 4.
- Zwick, R. (2000, Fall). Standards for the standards movement: Do high school exit exams measure up? *National CrossTalk*, pp. 10-11.
- Zwick, R. (2000, Sept. 3). College Admissions: Moving the Bar. *Santa Barbara News-Press*, pp. G1-G2.
- Zwick, R. (2000). Review of Test Theory: A Unified Treatment, by Roderick P. McDonald (Erlbaum, 1999). *Journal of the American Statistical Association*, 95, 1012-1013.
- Zwick, R., Thayer D. T., & Lewis, C. (2000). Using loss functions for DIF detection: An empirical Bayes approach *Journal of Educational and Behavioral Statistics*, 25, 225-247.

Journal Publications (continued)

- Zwick, R., Thayer D. T., & Lewis, C. (1999). An empirical Bayes approach to Mantel-Haenszel DIF analysis. *Journal of Educational Measurement*, 36, 1-28.
- Hambleton, R. K., Brennan, R. L., Brown, W., Dodd, B., Forsyth, R. A., Mehrens, W. A. Nellhaus, J., Reckase, M., Rindone, D., van der Linden, W., & Zwick, R. (2000). A response to "Setting reasonable and useful performance standards" in the National Academy of Sciences' *Grading the Nation's Report Card*. *Educational Measurement: Issues and Practice*, 19, 5-14.
- Zwick, R. (1999, December). Eliminating standardized tests in college admissions: The new affirmative action? *Phi Delta Kappan*, 320-324.
- Zwick, R. (1999, February 10). Back door affirmative action. *Education Week*, pp. 56, 35.
- Zwick, R., Thayer, D. T., & Mazzeo, J. (1997). Descriptive and inferential procedures for assessing DIF in polytomous items. *Applied Measurement in Education*, 10, 321-344.
- Zwick, R. (1997). Effect of adaptive administration on the variability of the Mantel-Haenszel measure of differential item functioning. *Educ. and Psych. Measurement*, 57, 412-421.
- Rudas, T., & Zwick, R. (1997). Estimating the importance of differential item functioning. *Journal of Educational and Behavioral Statistics*, 22, 31-45.
- Zwick, R., & Thayer, D. T. (1996). Evaluating the magnitude of differential item functioning in polytomous items. *Journal of Educational and Behavioral Statistics*, 21, 187-201.
- Zwick, R., Thayer, D. T., & Wingersky, M. (1995). Effect of Rasch calibration on ability and DIF estimation in computer-adaptive tests. *Journal of Educational Measurement*, 32, 341-363.
- Zwick, R., Thayer, D. T., & Wingersky, M. (1994). A simulation study of methods for assessing differential item functioning in computerized adaptive tests. *Applied Psychological Measurement*, 18, 121-140.
- Zwick, R., Donoghue, J., & Grima, A. (1993). Assessment of differential item functioning for performance tasks. *Journal of Educational Measurement*, 30, 233-251.
- Braun, H., & Zwick, R. (1993). Empirical Bayes analysis of families of survival curves: Applications to the analysis of degree attainment. *Journal of Educ. Statistics*, 18, 285-303.
- Zwick, R. (1993). The validity of the GMAT for the prediction of grades in doctoral study in business and management: An empirical Bayes approach. *Journal of Educational Statistics*, 18, 91-107.
- Beaton, A. E., & Zwick, R. (1992). Overview of the National Assessment of Educational Progress. *Journal of Educational Statistics*, 17, 95-109.
- Zwick, R. (1992). Statistical and psychometric issues in the measurement of educational achievement trends: Examples from the National Assessment of Educational Progress. *Journal of Educational Statistics*, 17, 205-218.
- Zwick, R. (1991). Effects of item order and context on estimation of NAEP reading proficiency. *Educational Measurement: Issues and Practice*, 10, 10-16.
- Zwick, R. (1990). When do item response function and Mantel-Haenszel definitions of differential item functioning coincide? *Journal of Educational Statistics*, 15, 185-197.

Journal Publications (continued)

- Zwick, R., & Ercikan, K. (1989). Analysis of differential item functioning in the NAEP history assessment. *Journal of Educational Measurement*, 26, 55-66.
- Zwick, R., (1988). Another look at interrater agreement. *Psychological Bulletin*, 103, 374-378.
- Zwick, R. (1987). Assessing the dimensionality of NAEP reading data. *Journal of Educational Measurement*, 24, 293-308.
- Zwick, R., (1987). Some properties of the correlation matrix of dichotomous Guttman items. *Psychometrika*, 52, 515-520.
- Jones, L. V., Davenport, E., Bryson, A., Bekhuis, T., & Zwick, R. (1986). Mathematics and science test scores as related to courses taken in high school and other factors. *Journal of Educ. Measurement*, 23, 197-208.
- Zwick, R., & Cramer, E. M. (1986). A multivariate perspective on the analysis of categorical data. *Applied Psychological Measurement*, 10, 141-145.
- Zwick, R. (1986). Rank & normal scores alternatives to Hotelling's T^2 . *Multivariate Behavioral Research*, 21, 169-86.
- Zwick, R. (1986). Testing pairwise contrasts in one-way analysis of variance designs. Invited contribution to statistical series, *Psychoneuroendocrinology*, 1, 253-276.
- Zwick, R., & Attkisson, C. C. (1985). Effectiveness of a client pretherapy orientation videotape. *Journal of Counseling Psychology*, 32, 514-524.
- Zwick, R. (1985). Nonparametric one-way multivariate analysis of variance: A computational approach based on the Pillai-Bartlett trace. *Psychological Bulletin*, 97, 148-152.
- Zwick, R., & Marascuilo, L. A. (1984). The selection of pairwise multiple comparison procedures for parametric and nonparametric analysis of variance models. *Psychological Bulletin*, 95, 148-155.
- Zwick, R., & Attkisson, C. C. (1984). The use of reception checks in client pretherapy orientation research. *Journal of Clinical Psychology*, 40, 446-452.
- Zwick, R. (1983). Assessing the psychometric properties of psychodiagnostic systems: How do the Research Diagnostic Criteria measure up? *Journal of Consulting & Clinical Psychology*, 51, 117-131.
- Marascuilo, L. A., & Zwick, R. (1983). Comment on Barnard: Another look at strength and direction of attitude using contrasts. *Psychological Bulletin*, 94, 534-539.
- Zwick, R., Neuhoﬀ, V., Marascuilo, L. A., & Levin, J. R. (1982). Statistical tests for correlated proportions: Some extensions. *Psychological Bulletin*, 92, 258-271.
- Attkisson, C. C., & Zwick, R. (1982). The Client Satisfaction Questionnaire: Psychometric properties and correlations with service utilization and psychotherapy outcome. *Evaluation and Program Planning*, 5, 233-237.
- Zwick, R., & Marascuilo, L. A. (1981). Can evaluation provide useful information?. *California Journal of Teacher Education*, 8, 61-80, 1981.

Selected Technical Reports and Other Publications

An investigation of the efficacy of criterion refinement procedures in Mantel-Haenszel DIF analysis. (ETS RR 13-16.) Princeton, NJ: ETS, 2013.

Disentangling the role of high school grades, SAT scores, and SES in predicting college achievement. (ETS Research Report No. 13-09.) Princeton, NJ: ETS, 2013.

A review of ETS differential item functioning assessment procedures: flagging rules, minimum sample size requirements, and criterion refinement. (ETS Research Rept. No. 12-08.) Princeton, NJ: ETS, 2012.

Applying score design principles in the design of score reports for CBAL teachers (with D. Zapata-Rivera & W. Van Winkle). (ETS Research Memorandum 12-20). Princeton, NJ: ETS, 2012.

Teaching teachers about test score interpretation: The ITEMS Project (with J. C. Sklar. In Zapata-Rivera, D., & Zwick, R. (Eds.), *Improving Test Score Reporting: Perspectives from the ETS Score Reporting Conference*. (ETS Research Report No. 11-45.) Princeton, NJ: ETS, 2011.

Scaling and Linking Through-Course Summative Assessments (with R. Mislevy).
http://www.k12center.org/rsc/pdf/TCSA_Symposium_Final_Paper_Zwick_Mislevy.pdf, 2011.

Measurement Issues in State Achievement Comparisons. (ETS Research Report No. 10-19). Princeton, NJ: ETS, 2010.

Letter report to the U.S. Department of Education on the Race to the Top Fund (with Haertel, E., et al.). Washington, DC: The National Academies Press (2009).

“What’s the Score?” “What Test Scores Do and Don’t Tell Us,” and “What’s the Difference?” [Web-based videos and supplementary materials produced as part of the NSF-sponsored project, Instructional Tools in Educational Measurement and Statistics (ITEMS), 2005-2008].

College admissions testing. Paper commissioned by the National Association for College Admission Counseling, February 2007. At www.nacac.net.org.

California and the SAT: A reanalysis of University of California admissions data (with T. Brown & J. C. Sklar). Center for Studies in Higher Education, UC Berkeley, Research and Occasional Papers Series, July 2004. At <http://repositories.cdlib.org/cshe/CSHE-8-04>.

Test and item bias: An overview. *Proceedings of the Social Statistics section of the 1994 meeting of the American Statistical Association*. Washington, DC: ASA, 1995.

Zwick, R. (1992). Differences in graduate school attainment patterns across academic programs and demographic groups. In J. Jones, M. E. Goertz, & C. V. Kuh (Eds.), *Minorities in graduate education: Pipeline, policy, and practice*. Princeton: ETS.

Selected Presentations

Admission to college by lottery: Should we give it a chance? Presented at the annual conference of the Association for the Study of Higher Education, Denver, November 2015.

Crafting a college class: Methods based on academic performance, noncognitive factors, affirmative action, and more. 9th Buros/ETS invited lecture, Buros Center for Testing, Lincoln, Nebraska, September 2015.

Impact of college admissions measures on characteristics of selected students. NCME meeting, Chicago, April 2015.

Selected Presentations

Extending the empirical Bayes DIF procedure to polytomous items. NCME meeting, Phila., April 2014.

Improving DIF estimation through Bayesian updating. Presented at the ETS San Francisco Office, August 2014 and at the ETS Princeton Office, November 2014.

Implications of measurement error for studies of predictive validity. ETS Seminar Series on Statistical Topics, Princeton, NJ, May 2013.

Validity and fairness considerations for multistage testing. NCME meeting, San Francisco, April 2013.

The role of admissions test scores, socioeconomic status, and high school grades in predicting college achievement. Invited presentation at a higher education conference sponsored by the Chilean Ministry of Education, Santiago, Chile, June 2012.

Reflections on the use of writing tests in college admissions: Invited presentation at a higher education conference sponsored by the Chilean Ministry of Education, Santiago, Chile, June 2012.

Efficacy of criterion refinement in DIF analysis. AERA meeting, Vancouver, BC, April 2012.

Disentangling the role of high school grades, SAT scores, and SES in predicting college achievement. AERA meeting, Vancouver, BC, April 2012.

Comparing graphical and verbal representations of measurement error in test score reports. NCME meeting, Vancouver, BC, April 2012.

Improving DIF estimation through Bayesian updating. Invited presentation, CTB McGraw-Hill, Monterey, CA, February 2012.

Scaling, linking, and reporting through-course summative assessments (with R. Mislevy). Ledyard R. Tucker Works series, ETS, July 2011.

Reducing systematic errors in college grade predictions using an index of high school socioeconomic status. AERA meeting, New Orleans, April 2011.

Innovative opportunities and measurement challenges in through-course summative assessments: NCME conference, New Orleans, April 2011; National Conference on Student Assessment, Orlando, June 2011.

A model for scaling, linking, and reporting through-course summative assessments. ETS K12 Webinar for Assessment Professionals, March 2011.

Scaling and linking through-course summative assessments. Invitational Research Symposium on Through-Course Summative Assessments, Atlanta, February 2011.

Teaching teachers about test score interpretation: The ITEMS Project. Conference on Score Reporting, Educational Testing Service, Princeton, November, 2010.

Using Innovative Assessments in State Accountability Programs. Invited presentation, National Research Council workshop, "Best Practices for State Assessment Systems" (part 2), Washington, April, 2010.

State achievement comparisons: Is the time right? Invited presentation, National Research Council workshop, "Best Practices for State Assessment Systems," Washington DC, December 11, 2009.

Empirical Bayes DIF assessment. Invited lecture, Visiting Scholar Program, Measured Progress, Dover, New Hampshire, August 2007.

Selected Presentations (continued)

Web-based training in measurement and statistics for school personnel: The ITEMS project. Invited lecture, Visiting Scholar Program, Measured Progress, Dover, New Hampshire, August 2007.

Higher education admissions testing. Invited presentation, NCME meeting, Chicago, April 2007.

Instructional Tools in Educational Measurement and Statistics (ITEMS) for School Personnel (with J. Sklar). National Educational Computing Conference, San Diego, July 2006 and the Joint Statistical Meetings, Seattle, August 2006.

New perspectives on the correlation of SAT scores, high school grades, and socioeconomic factors. Quantitative Methods in the Social Sciences colloquium, UCSB, March 2006.

ITEMS: A professional development resource for school personnel. K12 High-Speed Network Applications Coordination Committee, Sacramento, Aug. 2005 and the Assoc. of Calif. School Administrators Curriculum, Instruction & Assessment Committee, Sacramento, Oct. 2005.

Instructional Tools in Educational Measurement and Statistics (ITEMS) for School Personnel. Teacher Professional Continuum Conference sponsored by the National Science Foundation and the Center for Science Education, Washington DC, June 2005 and Reston, Virginia, May 2006.

Evaluating the association between socioeconomic status and educational achievement: Methodological and societal factors. Invited contribution, NCME meeting, Montreal, April 2005.

Participation of students with disabilities and English-language learners in NAEP and other large-scale assessments. Conference on Large-Scale Assessment, Boston, June 22, 2004.

UC Admissions: The controversy over "comprehensive review." Invited presentation, Meeting of the Dean's Council, Gevirtz Graduate School of Education, UCSB, May 7, 2004.

Recent findings on the prediction of college performance: The role of writing skills, language background, and socioeconomic factors. Vice-presidential address at the annual AERA meeting San Diego, April 15, 2004. Also presented at the Interdisciplinary Human Development Seminar, UCSB, May 3, 2004.

How will "No Child Left Behind" affect schools and students? Annual meeting of the American Association for the Advancement of Science, Seattle, February 15, 2004.

Statistical and psychometric definitions of test fairness, Cognitive and Perceptual Sciences Seminar, Department of Psychology, UCSB, October 31, 2003.

Six perils and pitfalls of test score interpretation. Invited presentation, Hechinger Institute on Education and the Media Seminar, Kansas City Missouri, June 1, 2003.

California and the SAT: A reanalysis of University of California admissions data. Annual colloquium, Center for Studies in Higher Education, UC Berkeley, May 2, 2003.

A reanalysis of University of California admissions data. AERA meeting, Chicago, April 25, 2003.

Organizer/ moderator, *How Can Researchers and the News Media Work Together to Improve Public Understanding of Educational Assessment?* AERA Division D Vice Presidential Invited Panel Discussion at the annual AERA meeting, Chicago, April 24, 2003.

Do SAT scores and high school grades predict college success? Quantitative Methods in the Social Sciences colloquium, UCSB. February 13, 2003.

Selected Presentations (continued)

The BOARS [Board of Admissions and Relations with Schools] Proposal for University of California Admissions Testing. Invited panelist, UC Regents meeting, San Francisco, March, 2002.

Is the SAT a wealth test? Invited presentation, College Board Western Regional Meeting, Anaheim, CA, February 2002. Also presented at the conference on "Rethinking the SAT," UCSB, November 2001.

The SAT: All questions answered. Invited panelist, College Board National Forum, Denver, October 2001.

Assessing bias in educational test items: Bayesian enhancements of a standard approach. Invited contribution to the Jerzy Neyman Seminar Series, Statistics Department, UC Berkeley, Nov. 2000. Also presented at the Quantitative Methods in the Social Sciences colloquium, UCSB, Nov. 2000.

Admissions without test scores: The new affirmative action? President's invited address, Southwest Educational Research Association, Dallas, January 2000; also presented at UCSB, June, 2001.

Will admissions without test scores increase diversity on California campuses? Invited presentation, University of California, Berkeley, October 1999.

Understanding DIF through contingency tables, complete DIF histories, curricular analysis, and cognitive research. NCME meeting, Montreal, April 1999.

An empirical Bayes approach to Mantel-Haenszel DIF analysis: Theoretical development and application to CAT data. Psychometric Society meeting, Urbana, Illinois, June 1998.

A simulation study of an empirical Bayes enhancement of Mantel-Haenszel DIF analysis in a testlet-based CAT. AERA meeting, San Diego, April 1998.

An empirical Bayes approach to Mantel-Haenszel DIF analysis for computer-adaptive tests. Invited presentations: Defense Manpower Data Center, Monterey, November 1997 and American Institutes for Research, Palo Alto, December 1997.

Statistical assessment of differential item functioning: Current methods and a possible new approach. Graduate School of Education, UCSB, June 1995.

Prediction of success in graduate school: Selected results on factors affecting Ph.D. completion and grades. Graduate School of Education, UCSB, June 1995.

Statistical assessment of the fairness of educational test items. Statistics Dept., U of Michigan, May 1995.

Stability of DIF classification: An alternative representation of the variability of the Mantel-Haenszel DIF statistic. NCME meeting, San Francisco, April 1995.

Test and item bias: An overview. Invited symposium contribution, American Statistical Association meeting, Toronto, August 1994.

Differential item functioning in new modes of assessment: Opportunities for experimental and methodological research. Invited presentation, NCME meeting, New Orleans, April 1994.

Differential item functioning analysis for computer-adaptive tests and other IRT-scored measures. Psychometric Society meeting, Berkeley, June 1993.

Toward a more explicit definition of item bias. Invited presentation, Conference in honor of Lyle Jones, University of North Carolina, Chapel Hill, November 1992.

January 13, 2017