Curriculum Vitae of Shelby J. Haberman

Contact Information

Postal address	50 Longwood Ave., Apt. 218 Brookline, MA 02446-5224
Telephone	(857)352-6815
Electronic mail	haberman.statistics@gmail.com
Web site	https://www.habermanstatistics.com

Education

1968	A.B.	Princeton University	Statistics (Highest Honors)
1970	Ph.D.	University of Chicago	Statistics (Hertz Fellow)

Employment

1970–72	Assistant Professor, Department of Statistics, University of Chicago
1972–75	Leonard Jimmie Savage Assistant Professor, Department of Statistics, University of Chicago
1975–81	Associate Professor, Department of Statistics, University of Chicago
1981–82	Professor, Department of Statistics, University of Chicago
1982–84	Professor, Department of Statistics, Hebrew University
1984–86	Professor, Departments of Mathematics, Managerial Economics and Decision Sciences, and Industrial Engineering and Management Sciences, Northwestern University
1986–87	Director, Center for Statistics and Probability, Northwestern University
1986–88	Chairman, Department of Statistics, Northwestern University
1986–2002	Professor, Department of Statistics, Northwestern University
2002–2005	Director, Statistical Theory and Practice, Educational Testing Service
2006–2010	Director, Statistical and Psychometric Theory and Practice, Educational Testing Service
2010–2017	Distinguished Presidential Appointee, Statistical and Psychometric Theory and Practice, Educational Testing Service
2017-2018	Distinguished Research Scientist, Edusoft
2018–	Self-employed consultant

Visiting positions

1974	Visiting Assistant Professor, Department of Statistics, University of California, Berkeley (Winter)
1977–78	Visiting Professor, Department of Statistics, Hebrew University
1981–82	Lady Davis Visiting Professor, Department of Statistics, Hebrew University
1983	Visiting Professor, Department of Statistics, University of Chicago (July)
1984	Visiting Professor, Department of Statistics, University of Chicago (July, August)

Awards

1977–78	Guggenheim Fellow
2009	NCME Award for Technical or Scientific Contributions to the Field of Educational Measurement (with S. Sinharay and G. Puhan)
2015	NCME Award for Technical or Scientific Contributions to the Field of Educational Measurement (with S. Sinharay and K. H. Chon)
2019	NCME Award for Career Contributions to Educational Measurement

Professional activities

1974–75	Vice President (Workshops), Chicago Chapter American Statistical Association
1975–	Fellow, Institute of Mathematical Statistics
1976–82	Associate Editor, *Journal of the American Statistical Association* (Theory and Methods)
1981	Member, Nominations Committee, Institute of Mathematical Statistics
1983–	Fellow, American Statistical Association
1983–84	Board of Directors, Israel Statistical Association
1983–85	Associate Editor, *The Annals of Statistics*
1987–94	Associate Editor, *Journal of Educational Statistics*
1989–92	Research Associate, Methodology Research Center, National Opinion Research Center
1998–	Fellow, American Association for the Advancement of Science
1999–2002	Associate Editor, *Journal of the American Statistical Association* (Theory and Methods)
1999–2002	Representative of Institute of Mathematical Statistics to Section Committee of the American Association for the Advancement of Science Section on Mathematics.
2011–2019	Editorial Board, *Journal of Educational and Behavioral Statistics*
2013–	Consulting Editor, *British Journal of Mathematical and Statistical Psychology*
2019–	Editorial Board, *Journal of Educational Measurement*

Patents

Haberman, S. J., Zhang, M., & Bridgeman, B. (2016). *Systems and methods for generating automated evaluation models* (U.S. Patent No. 9,443,193).

Haberman, S. J., Lee, Y.-H., Papierman, P., Zhou, Y., & Subhedar, R. (2022). *Systems and methods for detecting unusually frequent exactly matching and nearly matching test responses* (U.S. Patent No. 11,398,161).

Publications

Dorans, N. J., & Haberman, S. J. (2022). Recent challenges to maintaining score comparability: A commentary. *Journal of Educational Measurement*, *59*, 251–264. https://doi.org/10.1111/jedm.12319.

Etzioni, R. D., Fienberg, S. E., Gilula, Z., & Haberman, S. J. (1994). Statistical models for the analysis of ordered categorical data in public health and medical research. *Statistical Methods in Medical Research*, *3*, 179–204. https://doi.org/10.1177/096228029400300205.

Gilula, Z., & Haberman, S. J. (1986). Canonical analysis of contingency tables by maximum likelihood. *Journal of the American Statistical Association*, *81*, 780–788. https://doi.org/10.1080/01621459.1986.10478335.

Gilula, Z., & Haberman, S. J. (1988). Analysis of multiway contingency tables by restricted canonical and restricted association models. *Journal of the American Statistical Association*, *83*, 760–771. https://doi.org/10.1080/01621459.1988.10478659.

Gilula, Z., & Haberman, S. J. (1994). Conditional log-linear models for analyzing categorical panel data. *Journal of the American Statistical Association*, *89*, 645–656. https://doi.org/10.1080/01621459.1994.10476789.

Gilula, Z., & Haberman, S. J. (1995a). Dispersion of categorical variables and penalty functions: Derivation, estimation, and comparability. *Journal of the American Statistical Association*, *90*, 1447–1452. https://doi.org/10.1080/01621459.1995.10476651.

Gilula, Z., & Haberman, S. J. (1995b). Prediction functions for categorical panel data. *The Annals of Statistics*, *23*, 1130–1142. https://doi.org/10.1214/aos/1176324701.

Gilula, Z., & Haberman, S. J. (1998). Partition of chi-square [Online version in 2005]. In P. Armitage & T. Colton (Eds.), *The encyclopedia of biostatistics* (pp. 622–627, Vol. 1). John Wiley. https://doi.org/10.1002/9781118445112.stat04859.

Gilula, Z., & Haberman, S. J. (2000). Density approximation by summary statistics: An information-theoretic approach. *Scandinavian Journal of Statistics*, *27*, 521–534. https://doi.org/10.1111/1467-9469.00204.

Gilula, Z., & Haberman, S. J. (2001). Analysis of categorical response profiles by informative summaries. *Sociological Methodology*, *31*, 129–187. https://doi.org/10.1111/0081-1750.00094.

Gilula, Z., & Haberman, S. J. (2008). Correction note to *density approximation by summary statistics: an information-theoretic approach*. *Scandinavian Journal of Statistics*, *35*, 762–762. https://doi.org/10.1111/j.1467-9469.2008.00631.x.

Gilula, Z., Haberman, S. J., & van der Heijden, P. G. M. (2001). Multivariate analysis: Discrete variables (correspondence models). In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 10218–10221). Elsevier Science. https://doi.org/10.1016/b0-08-043076-7/00477-0.

Gilula, Z., Haberman, S. J., & van der Heijden, P. G. M. (2015). Multivariate analysis: Discrete variables (correspondence models). In J. D. Wright (Ed.), *International encyclopedia of the social and behavioral sciences* (2nd, pp. 121–124, Vol. 16). Elsevier Science. https://doi.org/10.1016/b978-0-08-097086-8.42153-7.

Goodman, L. A., & Haberman, S. J. (1990). The analysis of non-additivity in two-way analysis of variance. *Journal of the American Statistical Association*, *85*, 139–145. https://doi.org/10.1080/01621459.1990.10475317.

Guo, H., Rios, J. A., Haberman, S., Liu, O. L., Wang, J., & Paek, I. (2016). A new procedure for detection of students' rapid guessing responses using response time. *Applied Measurement in Education*, *29*, 173–183. https://doi.org/10.1080/08957347.2016.1171766.

Haberman, S. J. (1972). Algorithm AS 51: Log-linear fit for contingency tables. *Applied Statistics*, *21*, 218–225. https://doi.org/10.2307/2346506.

Haberman, S. J. (1973a). Algorithm AS 57: Printing multidimensional tables. *Applied Statistics*, *22*, 118–126. https://doi.org/10.2307/2346320.

Haberman, S. J. (1973b). The analysis of residuals in cross-classified tables. *Biometrics*, *29*, 205–220. https://doi.org/10.2307/2529686.

Haberman, S. J. (1973c). Log-linear models for frequency data: Sufficient statistics and likelihood equations. *The Annals of Statistics*, *1*, 617–632. https://doi.org/10.1214/aos/1176342458.

Haberman, S. J. (1974a). *The analysis of frequency data*. University of Chicago Press.

Haberman, S. J. (1974b). Log-linear models for frequency tables derived by indirect observation: Maximum likelihood equations. *The Annals of Statistics*, *2*, 911–924. https://doi.org/10.1214/aos/1176342813.

Haberman, S. J. (1974c). Log-linear models for frequency tables with ordered classifications. *Biometrics*, *30*, 589–600. https://doi.org/10.2307/2529224.

Haberman, S. J. (1975a). Direct products and linear models for complete factorial tables. *The Annals of Statistics*, *3*, 314–333. https://doi.org/10.1214/aos/1176343059.

Haberman, S. J. (1975b). How much do Gauss-Markov and least square estimates differ? a coordinate-free approach. *The Annals of Statistics*, *3*, 982–990. https://doi.org/10.1214/aos/1176343201.

Haberman, S. J. (1975c). Iterative scaling procedures for log-linear models for frequency tables derived by indirect observation. *Proceedings of the American Statistical Association (Statistical Computing Section)*, 45–50.

Haberman, S. J. (1976a). Correction to AS 51: Log-linear fit for contingency tables. *Applied Statistics*, *25*, 193. https://doi.org/10.2307/2346695.

Haberman, S. J. (1976b). Generalized residuals for log-linear models. *Proceedings of the Ninth International Biometrics Conference*, *1*, 104–172.

Haberman, S. J. (1976c). Review of *discrete multivariate analysis: theory and practice*, by Y. M. M. Bishop, S. E. Fienberg, and P. W. Holland. *The Annals of Statistics*, *4*, 817–820. https://doi.org/10.1214/aos/1176343556.

Haberman, S. J. (1977a). Analysis of scores of Ivy League football games. In S. P. Ladany & R. Machol (Eds.), *Optimal strategies in sports* (pp. 106–108). North-Holland.

Haberman, S. J. (1977b). Discussion of *maximum likelihood from incomplete data via the e-m algorithm*, by A. P. Dempster, N. M. Laird, and D. B. Rubin. *Journal of the Royal Statistical Society, Series B*, *39*, 31–32.

Haberman, S. J. (1977c). Log-linear models and frequency tables with small expected cell counts. *The Annals of Statistics*, *5*, 1148–1169. https://doi.org/10.1214/aos/1176344001.

Haberman, S. J. (1977d). Maximum likelihood estimates in exponential response models. *The Annals of Statistics*, *5*, 815–841. https://doi.org/10.1214/aos/1176343941.

Haberman, S. J. (1977e). Product models for frequency tables involving indirect observation. *The Annals of Statistics*, *5*, 1124–1147. https://doi.org/10.1214/aos/1176344000.

Haberman, S. J. (1977f). Review of *prediction analysis of cross classifications*, by D. K. Hildebrand, J. D. Laing, and H. Rosenthal. *Journal of the American Statistical Association*, *72*, 923. https://doi.org/10.2307/2286490.

Haberman, S. J. (1978). *Analysis of qualitative data, volume i: Introductory topics*. Academic Press.

Haberman, S. J. (1979). *Analysis of qualitative data, volume ii: New developments*. Academic Press. https://doi.org/10.1016/b978-0-12-312502-6.50005-0.

Haberman, S. J. (1980a). Discussion of *regression models for ordinal data*, by P. McCullagh. *Journal of the Royal Statistical Society, Series B*, *42*, 136–137.

Haberman, S. J. (1980b). A note on the conditional moments of a multivariate normal distribution confined to a convex set. *Journal of Multivariate Analysis*, *10*, 398–404. https://doi.org/10.1016/0047-259x(80)90060-3.

Haberman, S. J. (1981a). Comment on *An exponential family of probability distributions for directed graphs*, by P. W. Holland and S. Leinhardt. *Journal of the American Statistical Association*, *76*, 60–61. https://doi.org/10.2307/2287041.

Haberman, S. J. (1981b). Tests for independence in two-way contingency tables based on canonical correlation and on linear-by-linear interaction. *The Annals of Statistics*, *9*, 1178–1186. https://doi.org/10.1214/aos/1176345635.

Haberman, S. J. (1982a). Analysis of dispersion of multinomial responses. *Journal of the American Statistical Association*, *77*, 568–580. https://doi.org/10.2307/2287713.

Haberman, S. J. (1982b). Measures of association [Online versions in 2004, 2006, and 2014]. In S. Kotz & N. L. Johnson (Eds.), *Encyclopedia of statistical sciences* (pp. 130–137, Vol. 1). John Wiley. https://doi.org/10.1002/0471667196.ess0066.

Haberman, S. J. (1984). Adjustment by minimum discriminant information. *The Annals of Statistics*, *12*, 971–988. https://doi.org/10.1214/aos/1176346715.

Haberman, S. J. (1985a). Log-linear fit for contingency tables [Reprint with minor changes of Haberman (1972)]. In P. Griffiths & I. D. Hill (Eds.), *Applied statistics algorithms* (pp. 88–97). Ellis Horwood.

Haberman, S. J. (1985b). Printing multidimensional tables [Reprint with minor changes of Haberman (1973a)]. In P. Griffiths & I. D. Hill (Eds.), *Applied statistics algorithms* (pp. 101–109). Ellis Horwood.

Haberman, S. J. (1986). Correction for *adjustment by minimum discriminant information*, by S. J. Haberman. *The Annals of Statistics*, *14*, 358. https://doi.org/10.1214/aos/1176349863.

Haberman, S. J. (1987). Comment on *The calculation of posterior distributions by data augmentation*, by M. A. Tanner and W. H. Wong. *Journal of the American Statistical Association*, *82*, 546. https://doi.org/10.2307/2289461.

Haberman, S. J. (1988a). A stabilized Newton-Raphson algorithm for log-linear models for frequency tables derived by indirect observation. *Sociological Methodology*, *18*, 193–211. https://doi.org/10.2307/271049.

Haberman, S. J. (1988b). A warning on use of chi-square statistics with frequency tables with small expected cell counts. *Journal of the American Statistical Association*, *83*, 555–560. https://doi.org/10.1080/01621459.1988.10478632.

Haberman, S. J. (1989). Concavity and estimation. *The Annals of Statistics*, *17*, 1631–1661. https://doi.org/10.1214/aos/1176347385.

Haberman, S. J. (1990a). Comment on *A critical look at accumulation analysis and related methods*, by M. Hamada and C. F. J. Wu. *Technometrics*, *32*, 131–132. https://doi.org/10.1080/00401706.1990.10484626.

Haberman, S. J. (1990b). Review of *Eigenvalue Techniques for Qualitative Data*, by A. Israëls. *Journal of Official Statistics*, *6*, 217–219.

Haberman, S. J. (1991). Discussion of *Measures, models, and graphical displays in the analysis of cross-classified data*, by L. A. Goodman. *Journal of the American Statistical Association*, *86*, 1121–1123.

Haberman, S. J. (1995a). Computation of maximum-likelihood estimates in association models. *Journal of the American Statistical Association*, *90*, 1438–1446. https://doi.org/10.2307/2291536.

Haberman, S. J. (1995b). Review of *Statistical Applications Using Fuzzy Sets*, by K. G. Manton, M. A. Woodbury, and H. D. Tolley. *Journal of the American Statistical Association*, *90*, 1131–1133. https://doi.org/10.2307/2291362.

Haberman, S. J. (1996). *Advanced statistics, volume 1: Description of populations*. Springer-Verlag. https://doi.org/10.1007/978-1-4757-4417-0.

Haberman, S. J. (2000). Comments on *Invariance and factorial models*, by P. McCullagh. *Journal of the Royal Statistical Society, Ser. B*, *62*, 249–250.

Haberman, S. J. (2001). Multivariate analysis: Discrete variables (loglinear models). In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 10028–10233). Elsevier Science. https://doi.org/10.1016/b0-08-043076-7/00475-7.

Haberman, S. J. (2004a). *Joint and conditional maximum likelihood estimation for the Rasch model for binary responses* (ETS Research Report No. RR-04-20). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2004.tb01947.x.

Haberman, S. J. (2004b). *Probability prediction and classification* (ETS Research Report No. RR-04-19). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2004.tb01946.x.

Haberman, S. J. (2004c). *Statistical and measurement properties of features used in essay assessment* (ETS Research Report No. RR-04-21). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2004.tb01948.x.

Haberman, S. J. (2005a). *Identifiability of parameters in item response models with unconstrained ability distributions* (ETS Research Report No. RR-05-24). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2005.tb02001.x.

Haberman, S. J. (2005b). *Interpretations of reliability* (ETS Research Report No. RR-05-29). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2005.tb02006.x.

Haberman, S. J. (2005c). *Latent-class item response models* (ETS Research Report No. RR-05-28). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2005.tb02005.x.

Haberman, S. J. (2005d). *When can subscores have value?* (ETS Research Report No. RR-05-08). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2005.tb01985.x.

Haberman, S. J. (2006a). *Adaptive quadrature for item response models* (ETS Research Report No. RR-06-29). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2006.tb02035.x.

Haberman, S. J. (2006b). Bias in estimation of misclassification rates. *Psychometrika*, *71*, 387–394. https://doi.org/10.1007/s11336-004-1145-6.

Haberman, S. J. (2006c). *An elementary test of the normal 2pl model against the normal 3pl alternative* (ETS Research Report No. RR-06-14). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2006.tb02020.x.

Haberman, S. J. (2006d). *Joint and conditional estimation for implicit models for tests with polytomous item scores* (ETS Research Report No. RR-06-03). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2006.tb02009.x.

Haberman, S. J. (2007a). Electronic essay grading. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (pp. 205–233, Vol. 26). North-Holland. https://doi.org/10.1016/s0169-7161(06)26007-3.

Haberman, S. J. (2007b). *The information a test provides on an ability parameter* (ETS Research Report No. RR-07-18). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2007.tb02060.x.

Haberman, S. J. (2007c). The interaction model. In M. von Davier & C. H. Carstensen (Eds.), *Multivariate and mixture distribution Rasch models* (pp. 201–216). Springer Science. https://doi.org/10.1007/978-0-387-49839-3_13.

Haberman, S. J. (2008a). *Asymptotic limits of item parameters in joint maximum-likelihood estimation for the Rasch model* (ETS Research Report No. RR-08-04). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02090.x.

Haberman, S. J. (2008b). *Continuous exponential families: An equating tool* (ETS Research Report No. RR-08-05). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02091.x.

Haberman, S. J. (2008c). *Linking with continuous exponential families: Single-group designs* (ETS Research Report No. RR-08-61). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02147.x.

Haberman, S. J. (2008d). *Outliers in assessments* (ETS Research Report No. RR-08-41). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02127.x.

Haberman, S. J. (2008e). *Reliability of scaled scores* (ETS Research Report No. RR-08-70). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02156.x.

Haberman, S. J. (2008f). *Subscores and validity* (ETS Research Report No. RR-08-64). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02150.x.

Haberman, S. J. (2008g). When can subscores have value? *Journal of Educational and Behavioral Statistics*, *33*, 204–229. https://doi.org/10.3102/1076998607302636.

Haberman, S. J. (2009a). *Linking parameter estimates derived from an item response model through separate calibrations* (ETS Research Report No. RR-09-40). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2009.tb02197.x.

Haberman, S. J. (2009b). *Use of generalized residuals to examine goodness of fit of item response models* (ETS Research Report No. RR-09-15). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2009.tb02172.x.

Haberman, S. J. (2009c). Using exponential families for equating. In A. A. von Davier (Ed.), *Statistical models for test equating, scaling, and linking* (pp. 125–140). Springer. https://doi.org/10.1007/978-0-387-98138-3_8.

Haberman, S. J. (2010a). *Limits on the accuracy of linking* (ETS Research Report No. RR-10-22). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2010.tb02229.x.

Haberman, S. J. (2010b). Model selection. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (pp. 289–295, Vol. 7). Elsevier. https://doi.org/10.1016/b978-0-08-044894-7.01377-4.

Haberman, S. J. (2010c). *Raw-to-scale conversions with continuous exponential families* (ETS Research Memorandum No. RM-10-14). Educational Testing Service.

Haberman, S. J. (2010d). Using exponential families for equating. In A. A. von Davier (Ed.), *Statistical models for test equating, scaling, and linking* (pp. 125–140). Springer. https://doi.org//10.1007/978-0-387-98138-3_8.

Haberman, S. J. (2011a). The contributions of Paul Holland. In N. J. Dorans & S. Sinharay (Eds.), *Looking back: Proceedings of a conference in honor of paul holland* (pp. 3–17). Springer. https://doi.org/10.1007/978-1-4419-9389-2_1.

Haberman, S. J. (2011b). *Use of e-rater in scoring of the TOEFL iBT writing test* (ETS Research Report No. RR-11-25). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02261.x.

Haberman, S. J. (2013). *A general program for item-response analysis that employs the stabilized Newton-Raphson algorithm* (ETS Research Report No. RR-13-32). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2013.tb02339.x.

Haberman, S. J. (2014). *A program for adjustment by minimum discriminant information* (ETS Research Memorandum No. RM-14-01). Educational Testing Service.

Haberman, S. J. (2015a). Multivariate analysis: Discrete variables (loglinear models). In J. D. Wright (Ed.), *International encyclopedia of the social & behavioral sciences* (pp. 125–130, Vol. 16). Elsevier. https://doi.org/10.1016/b978-0-08-097086-8.42153-7.

Haberman, S. J. (2015b). *A program for equivalent-group equating with continuous exponential families* (ETS Research Memorandum No. RM-15-05). Educational Testing Service.

Haberman, S. J. (2015c). Pseudo-equivalent groups and linking. *Journal of Educational and Behavioral Statistics*, *40*, 254–273. https://doi.org/10.3102/1076998615574772.

Haberman, S. J. (2016a). Exponential family distributions relevant to IRT. In W. J. van der Linden (Ed.), *Handbook of item response theory, volume 2: Statistical tools* (pp. 47–69). CRC Press.

Haberman, S. J. (2016b). Models with nuisance and incidental parameters. In W. J. van der Linden (Ed.), *Handbook of item response theory, volume 2: Statistical tools* (pp. 151–170). CRC Press.

Haberman, S. J. (2016c). *A program for iterative proportional fitting of two-way arrays* (ETS Research Memorandum No. RM-16-04). Educational Testing Service.

Haberman, S. J. (2016d). *A program for nonparametric raw-to-scale conversion* (ETS Research Memorandum No. RM-16-03). Educational Testing Service.

Haberman, S. J. (2017). *A program for nonparametric equivalent-group equating* (ETS Research Memorandum No. RM-17-01). Educational Testing Service.

Haberman, S. J. (2018). Contribution to *Special* section in memory of Stephen E. Fienberg (1942—2016) AOAS Editor-in-Chief 2013—2015. *The Annals of Applied Statistics*, *12*, v. https://doi.org/10.1214/17-AOAS122ED.

Haberman, S. J. (2019a). *Cross-validation and U-statistics* (ETS Research Report No. RR-19-27). Educational Testing Service. https://doi.org/10.1002/ets2.12263.

Haberman, S. J. (2019b). *Measures of agreement versus measures of prediction accuracy* (ETS Research Report No. RR-19-20). Educational Testing Service. https://doi.org/10.1002/ets2.12258.

Haberman, S. J. (2020a). *Application of best linear prediction and penalized best linear prediction to ETS tests* (ETS Research Report No. RR-20-08). Educational Testing Service. https://doi.org/10.1002/ets2.12290.

Haberman, S. J. (2020b). Rejoinder: A brief response to the commentaries by Robert Mislevy and David Thissen. *Journal of Educational Measurement*, *57*, 403–404. https://doi.org/10.1111/jedm.12281.

Haberman, S. J. (2020c). Statistical theory and assessment practice. *Journal of Educational Measurement*, *57*, 374–385. https://doi.org/10.1111/jedm.12282.

Haberman, S. J., & Dorans, N. J. (2011). *Sources of score scale inconsistency* (ETS Research Report No. RR-11-10). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02246.x.

Haberman, S. J., Guo, H., Liu, J., & Dorans, N. (2008). *Consistency of SAT I: Reasoning Test score conversions* (ETS Research Report No. RR-08-67). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02153.x.

Haberman, S. J., Holland, P. W., & Sinharay, S. (2006). *Limits on log cross-product ratios for item response models* (ETS Research Report No. RR-06-10). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2006.tb02016.x.

Haberman, S. J., Holland, P. W., & Sinharay, S. (2007). Limits on log odds ratios for unidimensional item response theory models. *Psychometrika*, *72*, 551–561. https://doi.org/10.1007/s11336-007-9009-0.

Haberman, S. J., Jiang, W., & Spencer, B. D. (2021). *Development of methodology for evaluation of model-based estimates of the population size for states* (NORC Working Paper Series No. WP2021.03). National Opinion Research Center.

Haberman, S. J., & Lee, Y.-H. (2017). *A statistical procedure for testing unusually frequent exactly matching responses and nearly matching responses* (ETS Research Report No. RR-17-23). Educational Testing Service. https://doi.org/10.1002/ets2.12150.

Haberman, S. J., & Lee, Y.-H. (2020). *A program to group test takers identified by multiple security analyses* (ETS Research Memorandum No. RM-20-01). Educational Testing Service.

Haberman, S. J., Lee, Y.-H., & Qian, J. (2009). *Jackknifing techniques for evaluation of equating accuracy* (ETS Research Report No. RR-09-39). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2009.tb02196.x.

Haberman, S. J., Liu, Y., & Lee, Y.-H. (2019). *Distractor analysis for multiple-choice tests: An empirical study with international language assessment data* (ETS Research Report No. RR-19-39). Educational Testing Service. https://doi.org/10.1002/ets2.12275.

Haberman, S. J., & Qian, J. (2004). *The best linear predictor for true score from a direct estimate and several derived estimates* (ETS Research Report No. RR-04-35). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2004.tb01962.x.

Haberman, S. J., & Qian, J. (2007). Linear prediction of a true score from a direct estimate and several derived estimates. *Journal of Educational and Behavioral Statistics*, *32*, 6–23. https://doi.org/10.3102/1076998606298036.

Haberman, S. J., & Sinharay, S. (2008). *Sample-size requirements for automated essay scoring* (ETS Research Report No. RR-08-32). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02118.x.

Haberman, S. J., & Sinharay, S. (2010a). The application of the cumulative logistic regression model to automated essay scoring. *Journal of Educational and Behavioral Statistics*, *35*, 586–602. https://doi.org/10.3102/1076998610375839.

Haberman, S. J., & Sinharay, S. (2010b). *How can multivariate item response theory be used in reporting of subscores?* (ETS Research Report No. RR-10-09). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2010.tb02216.x.

Haberman, S. J., & Sinharay, S. (2010c). Reporting of subscores using multidimensional item response theory. *Psychometrika*, *75*, 209–227. https://doi.org/10.1007/s11336-010-9158-4.

Haberman, S. J., & Sinharay, S. (2011). *How does the knowledge of subgroup membership of examinees affect the prediction of true subscores?* (ETS Research Report No. RR-11-43). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02279.x.

Haberman, S. J., & Sinharay, S. (2013a). Does subgroup membership information lead to better estimation of true subscores? *British Journal of Mathematical and Statistical Psychology*, *66*, 452–469. https://doi.org/10.1111/j.2044-8317.2012.02061.x.

Haberman, S. J., & Sinharay, S. (2013b). Generalized residuals for general models for contingency tables with application to item response theory. *Journal of the American Statistical Association*, *108*, 1435–1444. https://doi.org/10.1080/01621459.2013.835660.

Haberman, S. J., Sinharay, S., & Chon, K. H. (2013). Assessing item fit for unidimensional item response theory models using residuals from estimated item response functions. *Psychometrika*, *78*, 417–440. https://doi.org/10.1007/s11336-012-9305-1.

Haberman, S. J., Sinharay, S., Feinberg, R. A., & Wainer, H. (2024). *Subscores: A practical guide to their production and consumption*. Cambridge University Press. https://doi.org/10.1017\/9781009413701.

Haberman, S. J., Sinharay, S., & Lee, Y.-H. (2011). *Statistical procedures to evaluate quality of scale anchoring* (ETS Research Report No. RR-11-02). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02238.x.

Haberman, S. J., Sinharay, S., & Puhan, G. (2006). *Subscores for institutions* (ETS Research Report No. RR-06-13). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2006.tb02019.x.

Haberman, S. J., Sinharay, S., & Puhan, G. (2009). Reporting subscores for institutions. *British Journal of Mathematical and Statistical Psychology*, *62*, 79–95. https://doi.org/10.1348/000711007x248875.

Haberman, S. J., & von Davier, A. A. (2014). Considerations on parameter estimation, scoring, and linking in multistage testing. In D. Yan, A. A. von Davier, & C. Lewis (Eds.), *Computerized multistage testing: Theory and applications* (pp. 229–248). CRC Press.

Haberman, S. J., & von Davier, M. (2007). Some notes on models for cognitively based skill diagnosis. In C. R. Rao & S. Sinharay (Eds.), *Handbook of statistics* (pp. 1031–1038, Vol. 26). North-Holland. https://doi.org/10.1016/s0169-7161(06)26040-1.

Haberman, S. J., von Davier, M., & Lee, Y.-H. (2008). *Comparison of multidimensional item response models: Multivariate normal ability distributions versus multivariate polytomous distributions* (ETS Research Report No. RR-08-45). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02131.x.

Haberman, S. J., von Davier, M., & Lee, Y.-H. (2019). Comparison of multidimensional item response models: Multivariate normal ability distributions versus multivariate polytomous ability distributions. In G. R. Hancock, J. R. Harring, & G. B. Macready (Eds.), *Advances in latent class analysis: A festschrift in honor of C. Mitchell Dayton* (pp. 29–60). Information Age Publishing.

Haberman, S. J., & Yan, D. (2011). *Use of continuous exponential families to link forms via anchor tests* (ETS Research Report No. RR-11-11). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02247.x.

Haberman, S. J., & Yao, L. (2015). Repeater analysis for combining information from different assessments. *Journal of Educational Measurement*, *52*(2), 223–251. https://doi.org/10.1111/jedm.12075.

Haberman, S. J., Yao, L., & Sinharay, S. (2015). Prediction of true test scores from observed item scores and ancillary data. *British Journal of Mathematical and Statistical Psychology*, *68*, 363–385. https://doi.org/10.1111/bmsp.12052.

Holzman, P. S., Haberman, S. J., Kringlen, E., Levy, D. L., Proctor, L. R., & Yasillo, N. J. (1977). Abnormal pursuit eye movements in schizophrenia: Evidence for a genetic marker. *Archives of General Psychiatry*, *34*, 802–805. https://doi.org/10.1001/archpsyc.1977.01770190064005.

Holzman, P. S., Kringlen, E., Levy, D. L., & Haberman, S. J. (1980). Deviant eye tracking in twins discordant for psychosis: A replication. *Archives of General Psychiatry*, *37*, 627–631. https://doi.org/10.1001/archpsyc.1980.01780190025002.

Holzman, P. S., Kringlen, E., Levy, D. L., Proctor, L. R., & Haberman, S. (1978). Smooth pursuit eye movements in twins discordant for schizophrenia. *Journal of Psychiatric Research*, *14*, 111–120. https://doi.org/10.1016/0022-3956(78)90013-4.

Holzman, P. S., Meltzer, H. Y., Kringlen, E., Levy, D. L., Haberman, S. J., & Davis, J. M. (1979). Plama CPK levels in monozygotic and dizygotic twins discordant for schizophrenia. *Journal of Psychiatric Research*, *15*, 127–131. https://doi.org/10.1016/0022-3956(79)90023-2.

Kim, S., von Davier, A., & Haberman, S. J. (2006). *An alternative to equating in small samples in the non-equivalent groups anchor test design* (ETS Research Report No. RR-06-27). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2006.tb02033.x.

Kim, S., von Davier, A., & Haberman, S. J. (2007). *Investigating the effectiveness of a synthetic linking function on small sample equating* (ETS Research Report No. RR-07-37). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2007.tb02079.x.

Kim, S., von Davier, A. A., & Haberman, S. J. (2008). Small-sample equating using a synthetic linking function. *Journal of Educational Measurement*, *45*, 325–342. https://doi.org/10.1016/s0169-7161(06)26040-1.

Kim, S., von Davier, A. A., & Haberman, S. J. (2011). Practical application of a synthetic linking function on small-sample equating. *Applied Measurement in Education*, *24*, 95–114. https://doi.org/10.1080/08957347.2011.554601.

Kobrin, J., Sinharay, S., Haberman, S. J., & Chajewski, M. (2011). *An investigation of the fit of linear regression models to data from an SAT validity study* (ETS Research Report No. RR-11-19) (Also College Board Report 2011-03). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02255.x.

Krol, W. F., Klimt, C. R., Morledge, J., Haberman, S. J., & Stamler, J. (1980). Persantine-aspirin reinfarction study: Design, methods, and baseline results [American Heart Association Monograph Number 71]. *Circulation*, *62*, II-1–II-42.

Kruskal, W., & Haberman, S. J. (1968). Chromosomal effect and LSD: Samples of four. *Science*, *162*, 1508–1509. https://doi.org/10.1126/science.162.3861.1508.

Lee, Y.-H., & Haberman, S. J. (2009). *Stability of TOEFL iBT scaled scores* (ETS Research White Paper No. WP-09-01). Educational Testing Service.

Lee, Y.-H., & Haberman, S. J. (2011a). *Monitoring reuse of forms in TOEFL iBT* (ETS Research White Paper No. WP-11-01). Educational Testing Service.

Lee, Y.-H., & Haberman, S. J. (2011b). *Monitoring reuse of forms in TOEFL iBT : A follow-up study* (ETS Research White Paper No. WP-11-02). Educational Testing Service.

Lee, Y.-H., & Haberman, S. J. (2013). Harmonic regression and scale stability. *Psychometrika*, *78*, 815–829. https://doi.org/10.1007/s11336-013-9337-1.

Lee, Y.-H., & Haberman, S. J. (2015). Investigating test-taking behaviors using timing and process data. *International Journal of Testing*, *16*, 240–267. https://doi.org/10.1080/15305058.2015.1085385.

Lee, Y.-H., & Haberman, S. J. (2021). Studying score stability with a harmonic regression family: A comparison of three approaches to adjustment of examinee-specific demographic data. *Journal of Educational Measurement*, *58*, 54–82. https://doi.org/https://doi.org/10.1111/jedm.12266.

Lee, Y.-H., Haberman, S. J., & Dorans, N. J. (2019). Use of adjustment by minimum discriminant information in linking constructed-response test scores in the absence of common items. *Journal of Educational Measurement*, *56*, 452–472. https://doi.org/10.1111/jedm.12216.

Lindsey, D. T., Holzman, P. S., Haberman, S. J., & Yasillo, N. J. (1978). Smooth-pursuit eye movements: A comparison of two measurement techniques for studying schizophrenia. *Journal of Abnormal Psychology*, *87*, 491–496. https://doi.org/10.1016/0022-3956(78)90013-4.

Lockwood, J. R., Haberman, S. J., Lee, Y. H., & Yao, L. (2018). *A general framework for using data from examinee pairs for security monitoring* (ETS Research Confidential Paper No. RCP-18-02). Educational Testing Service.

Lockwood, J. R., Yao, L., Lee, Y. H., & Haberman, S. J. (2018). *An exploration of examinee coordination on the TOEFL iBT test using response times* (ETS Research Confidential Paper No. RCP-18-01). Educational Testing Service.

Lu, R., Haberman, S. J., Guo, H., & Liu, J. (2015). *Use of jackknifing to evaluate effects of anchor item selection on equating with the nonequivalent groups with anchor test (neat) design* (ETS Research Report No. RR-15-10). Educational Testing Service. https://doi.org/10.1002/ets2.12056.

Puhan, G., Sinharay, S., Haberman, S. J., & Larkin, K. (2008). *Comparison of subscores based on classical test theory methods* (ETS Research Report No. RR-08-54). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2008.tb02140.x.

Puhan, G., Sinharay, S., Haberman, S. J., & Larkin, K. (2010). The utility of augmented subscores in a licensure exam: An evaluation of methods using empirical data. *Applied Measurement in Education*, *23*, 1–20. https://doi.org/10.1080/08957347.2010.486287.

Sinharay, S., Haberman, S., & Boughton, K. (2015). Too simple to be useful: A comment on Feinberg and Wainer (2014). *Educational Measurement: Issues and Practice*, *34*(3), 6–8. https://doi.org/10.1111/emip.12080.

Sinharay, S., & Haberman, S. J. (2008). *Reporting subscores: A survey* (ETS Research Memorandum No. RM-08-18). Educational Testing Service.

Sinharay, S., & Haberman, S. J. (2009). How much can we reliably know about what examinees know? *Measurement: Interdisciplinary Research and Perspectives*, *7*, 46–49. https://doi.org/10.1080/15366360802715486.

Sinharay, S., & Haberman, S. J. (2011a). Equating of augmented subscores. *Journal of Educational Measurement*, *48*, 122–145. https://doi.org/10.1111/j.1745-3984.2011.00137.x.

Sinharay, S., & Haberman, S. J. (2011b). *Equating of subscores and weighted averages under the NEAT design* (ETS Research Report No. RR-11-01). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02237.x.

Sinharay, S., & Haberman, S. J. (2014a). An empirical investigation of population invariance in the value of subscores. *International Journal of Testing*, *14*(1), 22–48. https://doi.org/10.1080/15305058.2013.822712.

Sinharay, S., & Haberman, S. J. (2014b). How often is the misfit of item response theory models practically significant? *Educational Measurement: Issues and Practice*, *33*(1), 23–35. https://doi.org/10.1111/emip.12024.

Sinharay, S., & Haberman, S. J. (2015). Comments on *A Note on Subscores* by Samuel A. Livingston. *Educational Measurement: Issues and Practice*, *34*(2), 6–7. https://doi.org/10.1111/emip.12071.

Sinharay, S., Haberman, S. J., & Jia, H. (2011). *Fit of item response theory models: A survey of data from several operational tests* (ETS Research Report No. RR-11-29). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2011.tb02265.x.

Sinharay, S., Haberman, S. J., & Lee, Y.-H. (2011). When does scale anchoring work: A case study. *Journal of Educational Measurement*, *48*, 61–80. https://doi.org/10.1111/j.1745-3984.2011.00131.x.

Sinharay, S., Haberman, S. J., & Melendez, C. R. (2012). *Examination of the added value of the scores on two sections of an admissions test* (ETS Research Memorandum No. RM-12-21). Educational Testing Service.

Sinharay, S., Haberman, S. J., P. W., H., & Lewis, C. (2012). *A note on the choice of an anchor test in equating* (ETS Research Report No. RR-12-14). Educational Testing Service. https://doi.org/10.1002/j.2333-8504.2012.tb02296.x.

Sinharay, S., Haberman, S. J., & Puhan, G. (2007). Subscores based on classical test theory: To report or not to report. *Educational Measurement: Issues and Practice*, *26*(4), 21–28. https://doi.org/10.1111/j.1745-3992.2007.00105.x.

Sinharay, S., Haberman, S. J., & Wainer, H. (2011). Do adjusted subscores lack validity? don't blame the messenger. *Educational and Psychological Measurement*, *71*, 789–797. https://doi.org/10.1177/0013164410391782.

Sinharay, S., Haberman, S. J., & Zwick, R. (2010). Issues with self-monitoring assessments: Comments on Koretz and Béguin. *Measurement: Interdisciplinary Research and Perspectives*, *8*, 191–194. https://doi.org/10.1080/15366367.2010.526454.

Sinharay, S., Puhan, G., & Haberman, S. J. (2010). Reporting diagnostic subscores in educational testing: Temptations, pitfalls, and some solutions. *Multivariate Behavioral Research*, *45*, 553–573. https://doi.org/10.1080/00273171.2010.483382.

Sinharay, S., Puhan, G., & Haberman, S. J. (2011). An NCME instructional module on subscores. *Educational Measurement: Issues and Practice*, *30*(3), 29–40. https://doi.org/10.1111/j.1745-3992.2011.00208.x.

Sinharay, S., Puhan, G., Haberman, S. J., & Hambleton, R. K. (2018). Subscores: When to communicate them, what are their alternatives, and some recommendations. In D. Zapata-Rivera (Ed.), *Score reporting: Research and applications* (pp. 35–49). Routledge. https://doi.org/10.4324/9781351136501.

Sterling, T. D., Binks, R. G., Haberman, S. J., & Pollack, S. V. (1969). Robot data screening: An ubiquitous automatic search technique. In R. Milton & J. Nelder (Eds.), *Statistical computation* (pp. 319–333). Academic Press.

Sterling, T. D., Gleser, M., Haberman, S. J., & Pollack, S. V. (1966). Robot data screening: A solution to multivariate type problems in the biological and social sciences. *Communications of the Association for Computing Machinery*, *9*, 529–532. https://doi.org//10.1145/365719.365985.

Uhlenhuth, E. H., Glass, R. M., Kellner, R., & Haberman, S. J. (1982). Relative sensitivity of clinical measures in trials of anti-anxiety agents. In E. I. Burdock, A. Sudilovsky, & S. Gershon (Eds.), *The behavior of psychiatric patients: Quantitative techniques for evaluation* (pp. 393–409). Marcel Dekker.

Uhlenhuth, E. J., Balter, M. B., Haberman, S. J., & Lipman, R. S. (1977). Remembering life events. In J. S. Strauss, H. M. Babigian, & M. Roff (Eds.), *The origins and course of psychopathology: Methods of longitudinal research* (pp. 117–134). Plenum Press.

van Rijn, P. W., Sinharay, S., Haberman, S. J., & Johnson, M. S. (2016). Assessment of fit of item response theory models used in large-scale educational survey assessments. *Large-scale Assessments in Education*, *4*, 1–23. https://doi.org/10.1186/s40536-016-0025-3.

von Davier, M., & Haberman, S. J. (2014). Hierarchical diagnostic classification models morphing into unidimensional 'diagnostic' classification models—a commentary. *Psychometrika*, *79*, 340–346. https://doi.org/10.1007/s11336-013-9363-z.

Yao, L., Haberman, S. J., & McCaffrey, D. F. (2019). *A scoring approach for practical application of penalized best linear prediction (PBLP) with human and machine scores* (ETS Research Confidential Paper No. RCP-19-02). Educational Testing Service.

Yao, L., Haberman, S. J., McCaffrey, D. F., & Lockwood, J. R. (2020). *Large-sample properties of minimum discriminant information adjustment estimates under complex sampling designs* (ETS Research Report No. RR-20-13). Educational Testing Service. https://doi.org/10.1002/ets2.12297.

Yao, L., Haberman, S. J., & Zhang, M. (2019a). Penalized best linear prediction of true test scores. *Psychometrika*, *84*, 186–211. https://doi.org/10.1007/s11336-018-9636-7.

Yao, L., Haberman, S. J., & Zhang, M. (2019b). *Prediction of writing true scores in automated scoring of essays by best linear predictors and penalized best linear predictors* (ETS Research Report No. RR-19-13). Educational Testing Service. https://doi.org/10.1002/ets2.12248.

Yao, L., Lee, Y.-H., & Haberman, S. J. (2020). *Exploring response-time similarity between pairs of examinees in TOEFL interchangeable set design (ISD)* (ETS Research Confidential Paper No. RCP-20-01). Educational Testing Service.

Yao, L., Sinharay, S., & Haberman, S. J. (2014). *Documentation for the software package SQE* (ETS Research Memorandum No. RM-14-02). Educational Testing Service.

Zhang, M., Yao, L., Haberman, S. J., & Dorans, N. J. (2019). Assessing scoring accuracy and assessment accuracy for spoken responses. In K. Zechner & K. Evanini (Eds.), *Automated speaking assessment* (pp. 32–58). Routledge. https://doi.org/10.4324/9781315165103-3.