

## IRT References

|              |   |
|--------------|---|
| Books.....   | 1 |
| Papers ..... | 2 |
| Other.....   | 4 |

### Books

Baker, F. B. (1992). *Item response theory: parameter estimation techniques*. Reading, NY: Marcel Dekker.

Bock, R. D. (1970). Estimating multinomial response relations. In R. C. Bose, *et al.* (Eds.), *Contributions to statistics and probability*. Chapel Hill, NC: University of North Carolina Press, 111-132.

Bock, R. D. (1993). Different DIFs. In: P. W. Holland & H. Wainer (Eds.), *Differential item functioning*. Hillsdale, NJ: Erlbaum, 115-122.

Bock, R. D. (1997). The nominal categories model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory*. New York: Springer Verlag, 33-65.

Bock, R. D., & Zimowski, M. F. (1995). Multiple group IRT. In W. van der Linden & R. Hambleton (Eds.), *Handbook of item response theory*. New York: Springer-Verlag.

Haberman, S. J. (1979). *Analysis of qualitative data, Vol. 2. New developments*. New York: Academic Press.

Hambleton, R. K., & Swaminathan, H. (1985). *Item Response Theory. Principles and applications*. Boston: Kluwer.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Holland, P. W., & Rubin, D. B. (Eds.) (1982). *Test equating*. Hillsdale, NJ: Erlbaum.

Holland, P. W., & Wainer, H. (1993). *Differential Item Functioning*. Hillsdale, NJ: Erlbaum.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Erlbaum.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores* (with contributions by A. Birnbaum). Reading, MA: Addison-Wesley.

Muraki, E. (1997). The generalized partial credit model. In W. J. van der Linden & R. K. Hambleton (Eds.), *Handbook of Modern Item Response Theory*. New York: Springer Verlag, 153-164.

Petersen, N. S., Kolen, M. J., & Hoover, H. D. (1989). Scaling, norming, and equating. In R. L. Linn (Ed.), *Educational Measurement (3rd edition)*. New York: American Council on Education-Macmillan, 221-262.

Stroud, A. H., & Secrest, D. (1966). *Gaussian Quadrature Formulas*. Englewood Cliffs, NJ: Prentice-Hall.

Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of DIF using the parameters of item response models. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning*. Hillsdale, NJ: Erlbaum, 67-113.

Van der Linden, W. J., & Hambleton, R. K. (1997). *Handbook of modern item response theory*. New York: Springer-Verlag.

Wainer, H. (Ed.) (1990). *Computerized adaptive testing: a primer*. Hillsdale, NJ: Erlbaum.

## Papers

Bock, R. D. (1972). Estimating item parameters and latent ability when responses are scored in two or more nominal categories. *Psychometrika*, **37**, 29-51.

Bock, R. D., & Aitkin, M. (1981). Marginal maximum likelihood estimation of item parameters: application of an EM algorithm. *Psychometrika*, **46**, 443-445.

Bock, R. D., Gibbons, R. D., & Muraki, E. (1988). Full information item factor analysis. *Applied Psychological Measurement*, **12**, 261-280.

Bock, R. D., & Lieberman, M. (1970). Fitting a response model for  $n$  dichotomously scored items. *Psychometrika*, **35**, 179-197.

Bock, R. D., Muraki, E., & Pfiffenberger, W. (1988). Item pool maintenance in the presence of item parameter drift. *Journal of Educational Measurement*, **25**, 275-285.

Bock, R. D., Thissen, D., & Zimowski, M. F. (1997). IRT estimation of domain scores. *Journal of Educational Measurement* **34**, 197-211.

Clogg, C. C., & Goodman, L. A. (1984). Latent structure analysis of a set of multi-dimensional contingency tables. *Journal of the American Statistical Association*, **79**, 762-661.

Dempster, A. P., Laird, N. M., & Rubin, D. B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, **39**, 1-38.

- Gibbons, R. D., & Hedeker, D. R. (1992). Full information item bi-factor analysis. *Psychometrika*, **57**, 423-436.
- Haberman, S. J. (1977). Log-linear models and frequency tables with small expected cell counts. *Annals of Statistics*, **5**, 1148-1169.
- Hendrickson, E. A., & White, P. O. (1964). Promax: A quick method for rotation to oblique simple structure. *British Journal of Mathematical and Statistical Psychology*, **17**, 65.
- Kaiser, H. F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**, 187-200.
- Kolakowski, D., & Bock, R. D. (1981). A multivariate generalization of probit analysis. *Biometrics*, **37**, 541-551.
- Linn, R. L., & Hambleton, R. K. (1991). Customized sets and customized norms. *Applied Measurement in Education*, **4**, 185-207.
- Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, **47**, 149-174.
- Masters, G. N. (1985). A comparison of latent trait and latent class analyses of Likert-type data. *Psychometrika*, **50**, 69-82.
- Mislevy, R. J. (1983). Item response models for grouped data. *Journal of Educational Statistics*, **8**, 271-288.
- Mislevy, R. J. (1984). Estimating latent distributions. *Psychometrika*, **49**, 359-381.
- Mislevy, R. J. (1986). Bayesian modal estimate in item response models. *Psychometrika*, **51**, 177-195.
- Mislevy, R. J. (1987). Exploiting auxiliary information about examinees in the estimation of item parameters. *Applied Psychological Measurement*, **11**, 81-91.
- Mislevy, R. J., Johnson, E. G., & Muraki, E. (1992). Scaling procedures in NAEP. *Journal of Educational Statistics*, **17**, 131-154.
- Muraki, E. (1990). Fitting a polytomous item response model to Likert-type data. *Applied Psychological Measurement*, **14**, 59-71.
- Muraki, E. (1992). A generalized partial credit model: application of an EM algorithm. *Applied Psychological Measurement*, **16**, 159-176.
- Muraki, E., & Engelhard, G. (1985). Full information item factor analysis: applications of EAP scores. *Applied Psychological Measurement*, **9**, 417-430.
- Naylor, J. C., & Smith, A. F. M. (1982). Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*, **31**, 214-225.

- Olsson, U., Drasgow, F., & Dorans, N. J. (1982). The polyserial correlation coefficient. *Psychometrika*, **47**(3), 337-347.
- Samejima, F. (1969). Estimation of latent trait ability using a response pattern of graded scores. *Psychometrika Monograph Supplement*, No. 17.
- Samejima, F. (1974). Normal ogive model on the continuous response level in the multidimensional latent space. *Psychometrika*, **39**, 111-121.
- Thissen, D. (1982). Marginal maximum likelihood estimation for the one-parameter logistic model. *Psychometrika*, **47**, 175-186.
- Thissen, D., & Steinberg, L. (1984). A response model for multiple-choice items. *Psychometrika*, **49**, 501-519.
- Thissen, D., & Steinberg, L. (1986). A taxonomy of item response models. *Psychometrika*, **51**, 566-577.
- Thissen, D., & Steinberg, L. (1988). Data analysis using item response theory. *Psychological Bulletin*, **104**, 385-395.
- Thissen, D., Steinberg, L., & Fitzpatrick, A. R. (1989). Multiple-choice models: The distractors are also part of the item. *Journal of Educational Measurement*, **26**, 161-176.
- Thissen, D., Steinberg, L., & Money, J. A. (1989). Trace lines for testlets: A use of multiple-categorical-response models. *Journal of Educational Measurement*, **26**, 247-260.
- Thissen, D., & Wainer, H. (1982). Some standard errors in item response theory. *Psychometrika*, **47**, 397-412.
- Tsutakawa, R. K. (1992). Prior distribution for item response curves. *British Journal of Mathematical and Statistical Psychology*, **45**, 51-71.
- Tsutakawa, R. K., & Lin, H. Y. (1986). Bayesian estimation of item response curves. *Psychometrika*, **51**, 251-267.
- Wainer, H., & Kiely, G. L. (1987). Item clusters and computerized adaptive testing: A case for testlets. *Journal of Educational Measurement*, **24**, 185-201.
- Warm, T. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, **54**, 427-450.
- Zwick, R. (1987). Assessing the dimensionality of NAEP reading data. *Journal of Educational Measurement*, **24**, 293-308.

## Other

Samejima, F. (1979). A new family of models for the multiple-choice item. *Research Report*, No. 79-4, Department of Psychology, University of Tennessee.

Zimowski, M. F. (1985). *Attributes of spatial test items that influence cognitive processing*. Un-published doctoral dissertation, University of Chicago.