

## References

- Andersen, E. (1973) A goodness of fit test for the Rasch model. *Psychometrika*, 38, 123-140.
- Andrich, D. (1973) *Latent trait psychometric theory in the measurement and evaluation of essay writing ability*, Doctoral dissertation, The University of Chicago.
- Andrich, D. (1978) A rating formulation for ordered response categories. *Psychometrika*, 43, 561-573.
- Andrich, D. (1988) *Rasch Models for Measurement*. Newberry Park, CA: Sage Publications.
- Andrich, D. (2004) Understanding resistance to the data-model relationship in Rasch's paradigm: a reflection for the next generation. In E. Smith & R. Smith (Eds.) *Introduction to Rasch measurement. Theory, models and applications*. (pp. 25-47) Maple Grove, MN: JAM Press.
- Andrich, D. (2013a). The legacies of R. A. Fisher and K. Pearson in the application of the polytomous rasch model for assessing the empirical ordering of categories. *Educational and Psychological Measurement* XX(X) 1–28. Retrieved September, 2014, from <http://epm.sagepub.com/content/early/2013/02/14/0013164413477107>
- Andrich, D. (2013b) An Expanded Derivation of the Threshold Structure of the Polytomous Rasch Model That Dispels Any "Threshold Disorder Controversy." *Educational and Psychological Measurement* 73(1) 78–124.
- Andrich, D. (2014) *Advances in social measurement: A Rasch measurement theory*. Keynote address, Scientific Congress on chronic diseases, perceived health, stakes and future. Nancy, France
- Bond, T. & Fox, C. (2007) *Applying the Rasch Model. Fundamental Measurement in the Human Sciences. 2nd Edition*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Briggs, J. & Wilson, M. (2004) An introduction to multidimensional measurement using Rasch models. In E. Smith & R. Smith (Eds.) *Introduction to Rasch measurement. Theory, models and applications*. (pp.322-341) Maple Grove, MN: JAM Press.
- Choppin, B. (1985) A fully conditional estimation procedure for Rasch Model parameters. *Evaluation in Education*, 9, 29-42.
- Fischer, G. (1973). "The Linear Logistic Test Model as an Instrument in Educational Research." *Acta Psychologica*, 37, 359–374.
- Fischer, G. (1976) Some probabilistic models for measuring change. In D. N. M. De Gruijter & L. J. T. Van der Kamp (Eds.), *Advances in psychological and educational measurement* (pp.97-110). New York: Wiley.
- Fischer, G. (1995a) The linear logistic test model. In G. Fischer & I. Molenaar. (Eds.) *Rasch models – foundations, recent developments, and applications*. New York: Springer.
- Fischer, G. (1995b) Linear logistic models for change. In G. Fischer & I. Molenaar. (Eds.) *Rasch models – foundations, recent developments, and applications*. New York: Springer.
- Fischer, G. H. (1972). Conditional maximum-likelihood estimations of item parameters for a linear logistic test model. *Research Bulletin* 9, Psychological Institute University of Vienna, Vienna.

- Fischer, G. H. (1973). The linear logistic test model as an instrument in educational research. *Acta Psychologica*, 37, 359-374.
- Fischer, G. H. (1995). The linear logistic test model. In G. H. Fischer & I. W. Molenaar (Eds.), *Rasch models* (pp. 131-155). New York: Springer.
- Fischer, G. & Molenaar, I. (Eds.) (1995) *Rasch models – foundations, recent developments, and applications*. New York: Springer.
- Fisher, R.A. (1947). *The design of experiments*. (4<sup>th</sup>) Edinburgh: Oliver and Boyd.
- Fisher, W. (1992). Objectivity in measurement: a philosophical history of Rasch's separability theorem. In M. Wilson (Ed.) *Objective measurement: theory and practice. Vol. 1* (pp. 29-58). Norwood, NJ: Ablex.
- Gierl, M. & Haladyna, T. (Eds.). (2012). *Automatic Item Generation*. New York: Taylor-Francis/Routledge.
- Gorin, J. & Embretson, S. E. (2006). Item difficulty modeling of paragraph comprehension items. *Applied Psychological Measurement*, 30, 394-411.
- Guttman, L. (1950). The basis for scalogram analysis. In Stouffer et al. *Measurement and Prediction*. The American Soldier Vol. IV. New York: Wiley.
- Hahne, J. (2008). Analyzing position effects within reasoning items using the LLTM for structurally incomplete data. *Psychology Science Quarterly*, 50, 379-390.
- Hohensinn, C., Kubinger, K. D., Reif, M., Holocher-Ertl, S., Khorramdel, L., & Frebort, M. (2008). Examining Item-Position Effects in Large-Scale Assessment Using the Linear Logistic Test Model. *Psychology Science Quarterly*, 50, 391-402.
- Klein, H. (1975). *The world of measurements*. London: Allen and Unwin.
- Kubinger, K. D. (2008). On the revival of the Rasch model-based LLTM: From constructing tests using item generating rules to measuring item administration effects. *Psychology Science Quarterly*, Volume 50, 2008 (3), pp. 311-327
- Linacre, J. M. (2002). *Construction of measures from many-facet data*. *Journal of Applied Measurement*, 3, 486-512.
- Linacre, J. (2004) Estimation methods for Rasch measures. In E. Smith & R. Smith (Eds.) *Introduction to Rasch measurement. Theory, models and applications*. (pp. 25-47) Maple Grove, MN: JAM Press.
- Linacre, J. M. (1994). *Many-facet Rasch measurement* (2nd Ed.). Chicago: MESA.
- Linacre, J. M. (2012) *A user's guide to WINSTEP SMINISTEP Rasch-Model computer programs*. Retrieved July, 2013, from <http://www.winsteps.com/winsteps.htm>
- Linacre, J. & Wright, B. (2004) Construction of measures from many-facet data. In E. Smith & R. Smith (Eds.) *Introduction to Rasch measurement. Theory, models and applications*. (pp. 296-321) Maple Grove, MN: JAM Press.
- Ludlow, L. (1983) *The analysis of Rasch model residuals*. Doctoral dissertation. University of Chicago.

- Luecht, R. M. (2012). An Introduction to Assessment Engineering for Automatic Item Generation. In M. Gierl & T. Haladyna (Eds.). *Automatic Item Generation*. New York: Taylor-Francis/Routledge.
- Luecht, R. M. (2013). Assessment engineering task model maps, task models and templates as a new way to develop and implement test specifications. *Journal of Applied Testing Technology*, 14 ([www.testpublishers.org/journal-of-applied-testing-technology](http://www.testpublishers.org/journal-of-applied-testing-technology)).
- Mair, P. & Hatzinger, R., (2007) Extended Rasch modeling: the eRm package for the application of IRT models in R. *Journal of Statistical Software*. 20, 9. Retrieved July, 2007, from <http://www.jstatsoft.org/v20/i09/v20i09.pdf>
- Mair, P., & Hatzinger, R. (2007). eRm: extended Rasch models. R package version 0.9.5: <http://r-forge.r-project.org/>.
- Masters, G. (1982) A Rasch model for partial credit scoring. *Psychometika*, 47, 149-174.
- Mead, R. (1976) *Fit of data to the Rasch model through the analysis of residuals*. Doctoral dissertation. University of Chicago.
- Newstead, S. E., Bradon, P., Handley, S. J., Dennis, I., & Evans, J. S. B. T. (2006). Predicting the difficulty of complex logical reasoning problems. *Thinking & Reasoning*, 12, 62-90. Munich: Spektrum.
- Pollitt, A. et al. (2004). *Let's stop marking exams*. Paper presented at the IAEA Conference, Philadelphia. <http://www.lifeinbits.org/camexam/htdocs/papers/2004LSMEAP.pdf>.
- Poinstingl, H. (2008). The LLTM as the basis for item generating rules of a new reasoning test: Family Relations Test. *Psychology Science Quarterly*, 50.
- R Core Team (2014). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org/>.
- Rasch, G. (1960) *Probabilistic models for some intelligence and attainment tests*. Copenhagen: Danmarks Paedagogiske Institut. (reprinted 1980 with Foreword, Afterword, and References, Chicago: The University of Chicago Press)
- Rasch, G. (1977) On specific objectivity: an attempt at formalizing the request for generality and validity of scientific statements. *Danish Yearbook of Philosophy* 14: 58-94.
- Scheiblechner, H. (1972). The learning and solving of complex reasoning items. *Zeitschrift fur Experimentelle und Angewandete Psychologie*, 3, 456-506.
- Smith, E. & Smith, R. (2004) *Introduction to Rasch measurement. Theory, models and applications*. Maple Grove, MN: JAM Press.
- Smith, E. (2004) Metric development and score reporting in Rasch measurement. In E. Smith & R. Smith (Eds.) *Introduction to Rasch measurement. Theory, models and applications*. (pp. 25-47) Maple Grove, MN: JAM Press.
- Sonnleitner, P. (2008). Using the LLTM to evaluate an item generating system for reading comprehension. *Psychology Science Quarterly*, 50, 345-362.
- Stone, M. (2004) Substantive scale construction. In E. Smith & R. Smith (Eds.) *Introduction to Rasch measurement. Theory, models and applications*. (pp.201-225) Maple Grove, MN: JAM Press.

- Tavernor, R. (2007) *Smoot's ear: the measure of humanity*. New Haven: Yale University Press.
- Thurstone, L. (1926) The scoring of individual performance. *Journal of Educational Psychology*, 17, 446-457.
- Thurstone, L. (1928) Attitudes can be measured. *American Journal of Sociology* 33: 529-554.
- Wright, B. (1968) Sample-free test calibration and person measurement. In *Proceedings of the 1967 ETS invitational conference on testing problems*. (pp. 85-101) Princeton: ETS.
- Wright, B. (1977) Solving measurement problems with the Rasch model. *Journal of Educational Measurement* 14, 2, pp. 97-116.
- Wright, B. (1980) Foreword. In G. Rasch. *Probabilistic models for some intelligence and attainment tests*. Chicago: The University of Chicago Press.
- Wright, B. & Masters, G. (1982) *Rating scale analysis*. Chicago: MESA Press.
- Wright, B. & Panchapakesan, N. (1969) A procedure of sample-free item analysis. *Educational and Psychological Measurement* 29, pp. 23-48.
- Wright, B. & Stone, M. (1979) *Best test design*. Chicago: MESA Press
- Wu, Hsin-Yi. (1998) Software based on S-P chart analysis and its applications. *Proceedings of the National Science Council ROC(D)*. Vol. 8, No. 3, 1998. pp. 108-120. Downloaded September, 2007, from [nr.stpi.org.tw/ejournal/ProceedingD/v8n3/108-120.pdf](http://nr.stpi.org.tw/ejournal/ProceedingD/v8n3/108-120.pdf)