SUMMER 2022, VOL. 35 NO. 2; ISSN 1051-0796

RMT

RASCH MEASUREMENT TRANSACTIONS

- Overview of this Issue of RMT – Stefanie A. Wind & Leigh M. Harrell-Williams
- Item Calibration Invariance across Samples with Extreme Ability Differences – Thomas R. O'Neill and Ting Wang *(peer reviewed)*
- News from the Rasch Measurement SIG - Trent Haines
- Tutorial Announcement - Randy MacIntosh
- Book Announcement – Stefanie A. Wind

Transactions of the Rasch Measurement SIG
American Educational Research Association

Overview of The Issue

In this issue of RMT, we have included one peer-reviewed research note, news from the Rasch Measurement SIG, and several announcements that may be of interest to the Rasch community.

The issue begins with a research note from Thomas R. O'Neill and Ting Wang on item calibration invariance.

Following the research note is the latest election and award news from the new Rasch Measurement SIG Chair, Trent Haines.

Then, we provide two announcements regarding the use of R in Rasch measurement analysis that may be of interest to our readers. The first is about an online Rasch DIF tutorial in R by Randy MacIntosh. The second is about a newly released book on using R for Rasch analysis by Dr. Stefanie A. Wind and Cheng Hua.

As always, we welcome your contributions to the next issue for RMT. We would appreciate receiving your research note, conference or workshop announcement, etc. by September 1, 2022. Please contact us at the email addresses below if you wish to submit something for inclusion.

Sincerely,
Your RMT Co-editors, Leigh and Stefanie

Rasch Measurement Transactions
www.rasch.org/rmt

Editors: Leigh M. Harrell-Williams
& Stefanie A. Wind
Email submissions to: Leigh.Williams@memphis.edu
or swind@ua.edu
RMT Editors Emeritus: Richard M. Smith,
John M. Linacre, &
Ken Royal
Rasch SIG Chair: Trent Haines
Secretary: Haiying Long
Treasurer: Eli Jones
Program Chairs: Audrey Conway Roberts &
Kaiwen Man

Item Calibration Invariance across Samples with Extreme Ability Differences

The discussion of what counts as measurement was a hot topic from the 1920s to the 1960s and is still argued about in the social sciences today, although the best answers seem to have already been found. For example, fundamental measurement, a term coined by physicist Norman Campbell (Campbell, 1920), requires, at least by analogy, the possibility of a physical concatenation, like joining the ends of sticks to concatenate length or piling bricks to concatenate weight. Luce and Tukey (1964) established that conjoint additivity is a way of demonstrating the invariance property of fundamental measurement and it is now recognized as the mathematical generalization of Campbell's fundamental measurement. Rasch described the process in a social science context and referred to it as specific objectivity. He described the requirements as follows: "A person having a greater ability than another should have the greater probability of solving *any* item of the type in question, and similarly, one item being more difficult than another one means that for *any* person the probability of solving the second item correctly is the greater one" (Rasch, 1960).

This property is desirable because if you want a stable hierarchy of questions (a measurement construct) the hierarchy should not be dependent upon the ability level of the people who answered the questions. Ben Wright (1967) demonstrated this invariance property using a dichotomous Rasch model with the response from 976 law students to 48 reading comprehension items on the Law School Admission Test. He compared the item difficulty calibrations for the 48 questions based upon the top and bottom thirds of the examinees. Comparable item hierarchies resulted across the 2 subpopulations.

The analysis presented here is similar to the Wright study, only more extreme. It compares 197 medical questions from the American Board of Family Medicine's In-Training Exam (ITE; O'Neill et al., 2015, 2016), but it uses only the top and bottom 5% of the 13,542 examinees that took the examination. The middle 90% of the data was excluded! The correlation between these two sets of calibrations was 0.89 (Figure 1).

The item hierarchy presented below appears stable except for perhaps the 8 easiest questions. On these questions the effects of gross-mistargeting of items to the people became evident with the high ability group almost never incorrectly answering the questions. Of the 677 people in the high ability group, two of these questions were answered correctly by *everyone*, and the other six questions had only one or two people answering incorrectly.

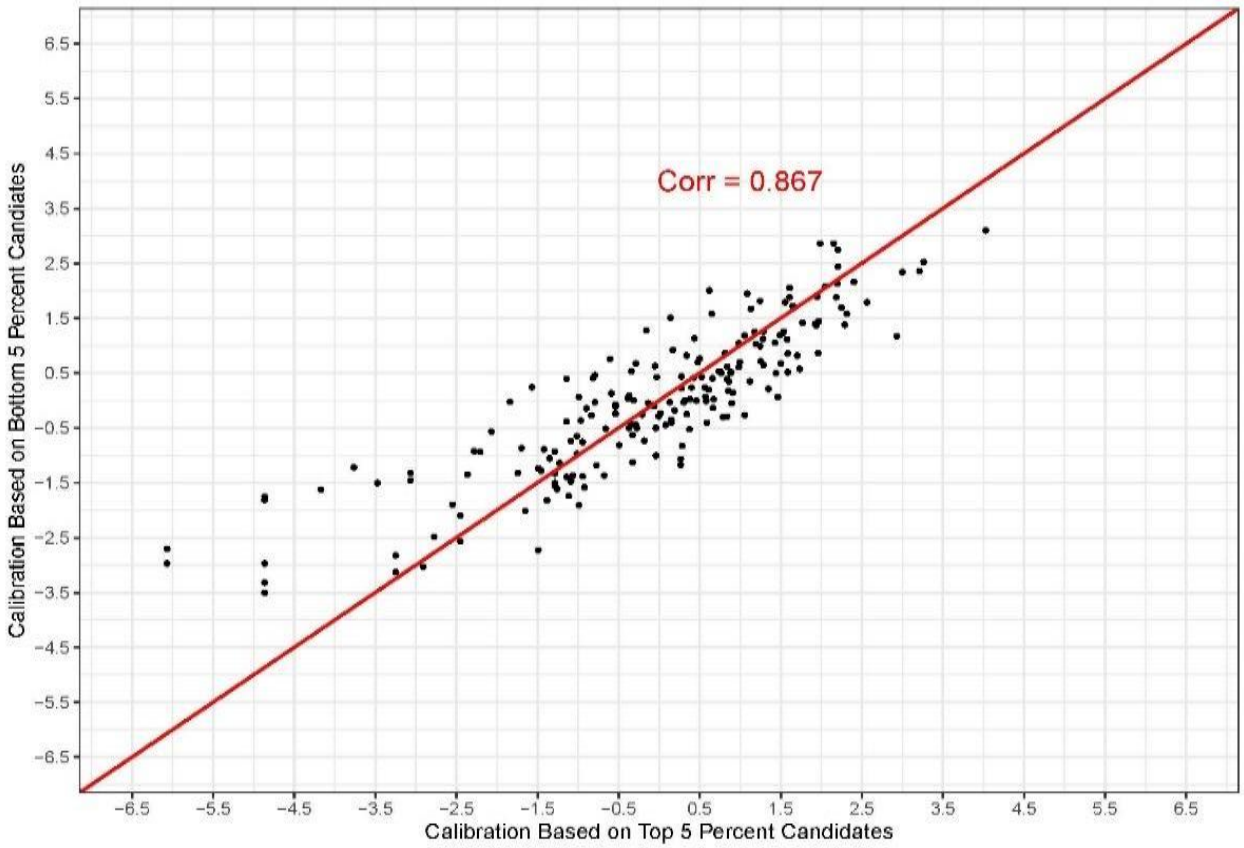


Figure 1.
Scatter plot between item calibration based on top 5 percent candidates (x-axis) and bottom 5 percent candidates (y-axis)

This article demonstrates that even when calibrating questions using groups of extremely different ability levels, the Rasch model is nevertheless very robust to these differences. When calibrating questions it is very important to ensure that there are a reasonable number of examinees for whom the questions are reasonably well-targeted. For data in which the items are substantially mis-targeted, psychometricians should consider using procedures similar to the CUTLO and CUTHI procedures in the Winsteps calibration software (Linacre, 2022). These procedures are based upon Waller's (1976) work for reducing the impact of mis-targeting questions given to examinees.

References

Campbell, N. R. (1920). *Physics: The elements*. Cambridge University Press.

Linacre, J. M. (2022). *Winsteps Rasch Measurement computer program* (5.2.3) [Computer software]. Winsteps.com. https://www.winsteps.com/index.htm

Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, *1*(1), 1–27.

O'Neill, T. R., Li, Z., Peabody, M. R., Lybarger, M., Royal, K. D., & Puffer, J. C. (2015). The Predictive Validity of ABFM's In-Training Examination. *Family Medicine*, *47*(5), 349–356.

O'Neill, T. R., Peabody, M. R., & Song, H. (2016). The Predictive Validity of the National Board of Osteopathic Medical Examiners' COMLEX-USA Examinations With Regard to Outcomes on American Board of Family Medicine Examinations. *Academic Medicine*, *91*(11), 1568–1575. https://doi.org/10.1097/ACM.0000000000001254

Rasch, G. (1960). *Probabilistic models for some intelligence and attainment tests*. Danish Institute for Educational Research.

Waller, M. I. (1976). Estimating Parameter in the Rasch Model: Removing the Effects of Random Guessing. *ETS Research Bulletin Series*, *1976*(1), 1–20. https://doi.org/10.1002/j.2333-8504.1976.tb01094.x

Wright, B. D. (1967). *Sample-Free Test Calibration and Person Measurement. Paper Presented at the National Seminar on Adult Education Research (Chicago, February 11-13, 1968).*

Thomas R. O'Neill, PhD

Ting Wang, PhD

American Board of Family Medicine

News from the Rasch Measurement SIG

Our news includes results from the 2022 officer elections and the announcement of the recipient of the 2022 Benjamin Drake Wright Senior Scholar Award.

Election Results

Every two years, members of the AERA Rasch Measurement SIG elect new officers through the AERA-sponsored election process. The elected officer positions are the Chair, Treasurer, and Secretary. The new officers' terms began immediately after the conclusion of the 2022 AERA Annual Meeting. The newly elected officers are R. Trent Haines (Chair), Eli Andrew Jones (Treasurer), and Haiying Long (Secretary).

The SIG would like to thank the previous leaders who served during a difficult time for the research community. Those officers are Jue Wang (Chair), Dandan Liao (Treasurer), Eli Andrew Jones (Secretary), and Manqian Liao (Program Co-Chair). Their terms concluded at the end of the 2022 AERA Annual Meeting.

In addition to elected officers, the SIG has two Program Co-Chairs who are appointed by the SIG Chair to serve for two years on staggered terms. Audrey Conway Roberts continues to serve and is now in her second year as Program Co-Chair. Kaiwen Man was appointed to fill a vacancy as Program Co-Chair and will serve until the end of the 2024 AERA Annual Meeting.

Chair: R. Trent Haines, Ph.D., in an Associate Professor and the Program Director of the Graduate Program in Psychometrics in the Department of Psychology at Morgan State University. He earned his Ph.D. in Educational Psychology from the University of Kentucky in 2006. His research interests include applying the Rasch model to the development of culturally responsive measurement techniques for educational and psychosocial research. Some of his current projects include the validation of tests and measures for use in different cultures, constructing culturally appropriate rating scales, and using technology to better capture data in culturally relevant ways. Prior to his time at Morgan State University, he served in research and instructional capacities at Louisiana State University Health Sciences Center's School of Public Health and School of Nursing, the University of Kentucky, and the University of New Orleans. He

previously served the AERA Rasch SIG as Program Co-Chair (2018-2020).

Treasurer: Eli Jones, Ph.D., is an assistant professor in the Educational Psychology and Research department at the University of Memphis. He received his Ph.D. from Brigham Young University (BYU) from the Educational Inquiry, Measurement, and Evaluation department. He has also served as a postdoctoral researcher for the Network for Educator Effectiveness at the University of Missouri, as an assistant professor of research at Columbus State University (Georgia), and as a second grade public school teacher. His area of research is currently focused on the measurement properties of rater-mediated assessments in education. Specifically, his research focuses on the application of many-facet Rasch models to educator evaluation (including teacher evaluation, principal evaluation, and evaluation of teacher candidates). His research has explored the psychometric properties of observational instruments, as well as the implications of sparse rating designs on model-data fit, rater error, and stability of rater severity estimates. He previously served the Rasch SIG as Secretary (2020-2022).

Secretary: Haiying Long, Ph. D., is an associate professor in Research, Evaluation, Measurement, and Statistics program and the director of Research Methods Boot Camp at the University of Kansas. Her research focuses on rater-mediated assessments, Rasch models, Structural Equation Modeling, measurement validity, program evaluation, large-scale data analysis, creativity assessment, and STEM education. She has published over 30 articles in several educational and psychological journals, including Educational Psychology Review, Review of Research in Education, Psychology of Aesthetics, Creativity, and the Arts, Social Psychology of Education, and Creativity Research Journal and has received federal funding from National Science Foundation and the Department of Education. She is currently an associate editor for Thinking Skills and Creativity.

Program Co-Chair: Audrey Conway Roberts, Ph.D., is an assistant professor in the School of Educational Foundations, Leadership & Policy at Bowling Green State University where she teaches undergraduate and graduate courses in assessment and research methodology. She received her Ph.D. in Educational Policy Studies and

Evaluation from the University of Kentucky with a focus in quantitative methodologies. Her research interests revolve around the validation of and use of measures of school effectiveness and learning environments so that more effective, data-driven decisions may be made from this information for practitioners and policymakers.

Program Co-Chair: Dr. Kaiwen Man was appointed as an Assistant Professor in the Department of Educational Studies in Psychology, Research Methodology, and Counseling at the University of Alabama. Kaiwen has held many positions as a researcher, including at the Association of American Medical College, at the Educational Testing Service, and at the Charted Financial Analyst. His research explores questions on the boundaries and interactions of the educational statistics, biometrics, and behavioral research literature with particular attention to models for eye-tracking data, responding process data, Bayesian statistics, and data mining. His works has been published in many peer-reviewed flagship quantitative journals such as Educational and Psychological Measurement, Journal of Educational and Behavioral Statistics, and Applied Psychological Measurement. Moreover, he has received the prestigious 2022 National Council on Measurement in Education (NCME) Brenda Loyd Outstanding Dissertation Award. Furthermore, his projects have been externally-funded by the ETS Harold Gulliksen Psychometric Research Fellowship program.

Benjamin Drake Wright Senior Scholar Award

During the 2022 AERA Annual Meeting, the Benjamin Drake Wright Senior Scholar Award was presented to Larry H. Ludlow at the Rasch SIG Business Meeting. This award is presented to an individual senior scholar for outstanding programmatic research and mentoring in Rasch measurement over the course of a career and who is still active in Rasch measurement research at the time the award is granted. Dr. Ludlow will be invited to give an address at the Rasch SIG Business Meeting during the 2023 AERA Annual Meeting.

Trent Haines
Morgan State University
Rasch SIG Chair

Tutorial Announcement

An online Rasch DIF analysis tutorial is available using the TAM package in R and demonstrates the workflow described in Andrich and Hagquist (2012, 2015) and Hagquist and Andrich (2017). The tutorial analyzes data from the American National Election Studies public opinion surveys. It may be of interest to instructors as the tutorial, software, and data are available free of charge. The tutorial is posted at https://pie-try-ka.com/dif-in-r.html.

The tutorial is a companion describing the methodology used in Pietryka and MacIntosh (2022) where we document violations of the assumption of measurement equivalence across a number of theoretically important political science variables. TAM is authored by Thomas Kiefer, Alexander Robitzsch and Margaret Wu. One important feature of the iterative workflow demonstrated in the tutorial is that it explicitly takes into account 'artificial' DIF that is a consequence of identification constraints in item parameter estimation.

The tutorial also provides the necessary R code to produce tables and plots for the DIF analysis.

Randy MacIntosh
CSU Sacramento

References

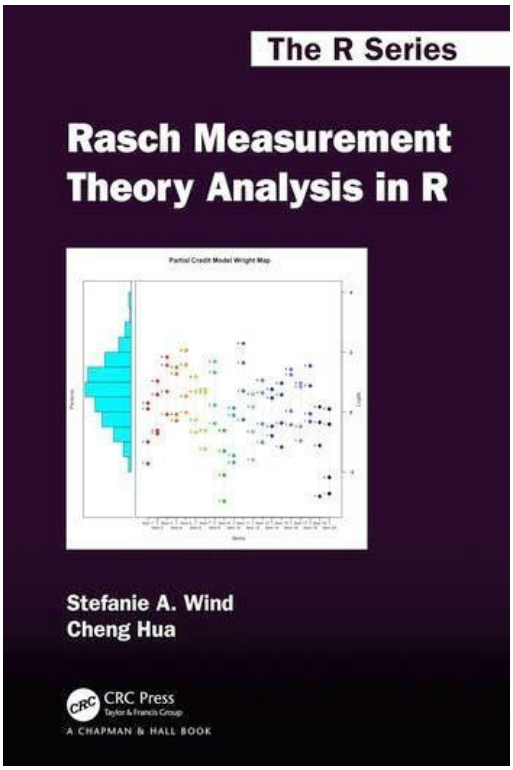
Andrich, David, and Curt Hagquist. 2012. "Real and Artificial Differential Item Functioning." Journal of Educational and Behavioral Statistics 37(3): 387–416.

Andrich, David, and Curt Hagquist. 2015. "Real and Artificial Differential Item Functioning in Polytomous Items." Educational and Psychological Measurement 75 (2): 185–207.

Hagquist, Curt, and David Andrich. 2017. "Recent Advances in Analysis of Differential Item Functioning in Health Research Using the Rasch Model." Health and Quality of Life Outcomes 15 (1): 181.

Pietryka, Matthew T., and Randall C. MacIntosh. 2022. "ANES Scales Often do not Measure What You Think They Measure." The Journal of Politics 84 (2): 1074–1090. doi.org/10.1086/715251.

Book Announcement: Rasch Measurement Theory Analysis in R



We are pleased to share the publication of *Rasch Measurement Theory Analysis in R*. This book provides researchers and practitioners with a step-by-step guide for conducting Rasch measurement theory analyses using R. It includes theoretical introductions to major Rasch measurement principles and techniques, demonstrations of analyses using several R packages that contain Rasch measurement functions, and sample interpretations of results.

Key features of the book include:

- Accessible to users with relatively little experience with R programming
- Reproducible data analysis examples that can be modified to accommodate users' own data
- Accompanying e-book website with links to additional resources and R code updates as needed
- Features dichotomous and polytomous (rating scale) Rasch models that can be applied to data from a wide range of disciplines

Interested readers can find out more about this book at the following page: https://www.routledge.com/Rasch-Measurement-Theory-Analysis-in-R/Wind-Hua/p/book/9780367776398

Stefanie A. Wind
Cheng Hua
The University of Alabama