

Applying the Rasch Model:
Fundamental Measurement
in the Human Sciences
Second Edition

TREVOR G. BOND

Hong Kong Institute of Education

CHRISTINE M. FOX

University of Toledo



2007

LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
Mahwah, New Jersey London

Copyright © 2007 by Lawrence Erlbaum Associates, Inc.

All rights reserved. No part of this book may be reproduced in any form, by photostat, microform, retrieval system, or any other means, without prior written permission of the publisher.

Lawrence Erlbaum Associates, Inc., Publishers
10 Industrial Avenue
Mahwah, New Jersey 07430
www.erlbaum.com
www.bondandfox.com

Cover design by Kathryn Houghtaling Lacey

**CIP information for this book can be obtained by contacting
the Library of Congress**

ISBN 978-0-8058-5461-9 (cloth)
ISBN 978-0-8058-5462-6 (paper)
ISBN 978-1-4106-1457-5 (e book)

Books published by Lawrence Erlbaum Associates are printed on acidfree paper, and their bindings are chosen for strength and durability.

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

To Ben Wright,
Provocative, persistent, and passionate
TGB & CMF

Contents

Foreword	ix
Preface	xiii
1 Why Measurement Is Fundamental	1
2 Important Principles of Measurement Made Explicit	15
3 Basic Principles of the Rasch Model	29
4 Building a Set of Items for Measurement	49
5 Invariance: A Crucial Property of Scientific Measurement	69
6 Measurement Using Likert Scales	101
7 The Partial Credit Rasch Model	123
8 Measuring Facets Beyond Ability and Difficulty	143
9 Development, Education, and Rehabilitation: Change Over Time	163
10 The Rasch Model Applied Across the Human Sciences	183
11 Rasch Modeling Applied: Rating Scale Design	219
12 Model Fit and Unidimensionality	235
13 A Synthetic Overview	261

Appendix A: Technical Aspects of the Rasch Model	277
Appendix B: Getting Started	287
Glossary	309
References	315
Author Index	331
Subject Index	335

Foreword

*Then felt I like some watcher of the skies
When a new planet swims into his ken;*

—John Keats, “On first looking into Chapman’s Homer”

The first edition of this remarkable work arrived stealthily. Those of us in the know were aware that it was to be released at the 2001 AERA Annual Meeting in Seattle. When the Exhibitor area opened I headed for the Lawrence Erlbaum Associates booth and looked for the book. I purchased the very first copy. By the end of the AERA Meeting, “Bond & Fox” had sold out and was on its way to becoming an Erlbaum best seller.

And deservedly so. Rigorous measurement has been essential to the advancement of civilization since Babylonian times, and to the advancement of physical science since the Renaissance. But education, psychology, and social science have short-changed themselves by degrading measurement to the “assignment of numerals ... according to rule” (Stevens, 1959a, p. 25) and then acting as though one rule is as good as another. The essential rule in successful measurement is used ubiquitously for money, length, area, volume, weight, temperature—the things we deem most important to supporting life, liberty, and the pursuit of happiness. That rule is “one more unit means the same amount extra no matter how much there already is.” This is exactly what Rasch measurement operationalizes for social science—and is what this book is all about.

Rasch measurement has appeared to be so complex and technical. Georg Rasch was a consultant mathematician and statistician. Though he used simple graphical techniques himself (Rasch, 1960, p. 71), he presented them in obscure algebra (1960, p. 173). In the following years, the utility of his conceptual breakthrough was lost in arguments over statistical niceties and the struggles of developing working software on the newly available computers. Benjamin Wright, the leading advocate of practical Rasch measurement, would convince audiences of the value of Rasch measurement. But even his books proved perplexing for the neophyte. It seemed that by the time the Rasch practitioner had become compe-

tent enough to write a book, those early hurdles that are always encountered in using an unfamiliar methodology were forgotten.

Trevor Bond and Christine Fox remembered, and wrote this book—now revised to excel its former self! Purists may decry the lack of technical sophistication within these pages—but that is exactly accords with the needs of the neophyte. This book filled a gaping void. Its success has motivated others to follow its lead, but so far, none has been so ruthless in rejecting the shackles of Greek letters, in-group jargon, statistical niceties, and the constraining historical precedents of a century of psychometrics.

Of course, history is important. The ideas underlying Rasch measurement go back to ancient Babylon, to the Greek philosophers, to the adventurers of the Renaissance. In fact, the fundamental mathematical relationship underlying the Rasch model was expressed by American philosopher Charles Sanders Peirce. Here is how he phrased it: “The combination of independent concurrent arguments takes a very simple form when expressed in terms of the intensity of belief, measured in the proposed way. It is this. Take the sum of all the feelings [*logarithm* of the chance”, i.e., log-odds] of belief which would be produced separately by all the arguments pro and subtract from that the similar sum for arguments con, and the remainder is the feeling of belief we ought to have on the whole. This is a procedure which men often resort to, under the name of balancing reasons.” (Peirce, 1878, p. 709)

So, why, over a century after the publication of Peirce’s article in a leading science journal of its time, is the arithmetic of log-odds regarded as obscure, technical, even fanatical, in Social Science? An answer is that Social Science has yet to regard rigorous measurement as crucial to progress.

The major conceptual transition in physical science was achieved with Isaac Newton’s *Philosophiæ Naturalis Principia Mathematica* [Mathematical Principles of Natural Philosophy] (1687). Prior to that work, physical science was chiefly devoted to describing phenomena. Arithmetical summaries were tentative and temporary. After Newton, his Laws of Motion were regarded as more reliable than the data. Astronomical observations that contradicted his Laws were challenged and brought into conformity with them. In contrast, Social Science has yet to advance to the point at which theory is regarded as more valid than data. But Newton’s work was possible only because of the advances in physical measurement that preceded it. Good measurement leads to profound theory.

So what is the usual situation in Social Science? “Measures” are tentative and temporary descriptions. They are considered to be good if they explain the variance of interest, or correlate well with some other “measures,” or predict well some outcome. If the “measures” are deemed unsatisfactory, then alternative methods of computing “measures” are investigated. This is the reverse of physical science. The measurement of weight and length follows strict rules. The extent to which weight explains variance, or correlates or predicts, is an experimental finding, not an aspect of measuring.

This book is the unusual situation in Social Science. It is a laborer to assist you in laying the foundation of truly scientific measurement upon which we can look

forward to seeing the building of a skyscraper of breathtaking Social Science theory—theory which will be able to withstand the storms of contradictory data.

“Enter to Grow in Wisdom. Go Forth to Apply Wisdom in Service.”¹

—John “Mike” Linacre, University of Sydney, Australia

REFERENCES

- Peirce, C.S. (1878). Illustration of the Logic of Science, by C.S. Peirce, Assistant in the United States Coast Survey. Fourth Paper: The Probability of Induction. *Popular Science Monthly*, pp. 705–718. Here from pp. 707–709.
- Stevens, S. S. (1959) Measurement, Psychophysics and Utility, Chap. 2, in C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and Theories*. New York: John Wiley.

¹ *Inscription on Flowers Hall, Huntingdon College, Montgomery, Alabama.*

Preface

At a *Quantitative Methods in the Social Sciences Colloquium* at the University of California Santa Barbara, the meeting chair introduced the topic, by referring, rather embarrassingly, to the first edition of Bond & Fox as an academic best-seller. He then provided the required empirical substantiation for his claim by circulating the latest Lawrence Erlbaum Associates catalogue with *Applying the Rasch Model* annotated just as announced: "Best Seller." Who would have imagined a book that was written primarily as a response to the perennial, frustrating question, "Can you tell me how you did that Rasch analysis?" could be so well received?

We have aimed our book at those who are studying quantitative research methods in the social sciences and, in particular, those who know they should be using the Rasch model but cannot work out the "How?" and the "Why?" We target graduate students in disciplines in education and psychology, but we intend to be relevant to those aiming to research across all human science disciplines. It seems that Rasch measurement is certainly on the uptake right across these disciplines: We aim to show our research colleagues how to get started and to give conceptually valid arguments for doing so.

Hardly a month passes without an unsolicited email from some corner of the world saying something like: I am really enjoying your Rasch analysis book, but I have a little question about page x. Last month's went like this: "After many years of dissatisfaction with traditional methods of analysing rating-scale data, I began a hunt on the basic assumption that someone somewhere must have developed newer, more satisfying methods. After finally encountering the name Rasch, and several more difficult months of trying to grasp the fundamentals . . . , I finally came across *Applying the Rasch Model*. I cannot tell you adequately what an enormous difference your book made. The graphical illustrations did wonders in furthering my understanding of measurement, the glossary was extremely helpful, and your pointing the reader to current software was a blessing."

With an audience containing colleagues such as that, it was difficult to resist putting together a second edition when the questions arose about whether one was on the way. Readers of the first edition will remember our rather heart-on-sleeve preface to that edition which gave personal accounts of how the authors succumbed to the lure of Rasch analysis and our thanks to those who helped us reach that point. Reviews of the first edition as well as solicited and unsolicited

advice from colleagues and students pointed out some inadequacies of our first effort. But somehow, we had managed to write an appropriate book at just the right time. No doubt the bulk of our readers will be pleased that we resisted most of the urging by our more mathematically literate advisors to introduce the Rasch equations earlier—in the body of the text. Instead, we have attempted to improve conceptual clarity and accuracy and to leave the formal mathematical treatment of the models to Appendix A and to the array of texts whose authors are obviously better qualified to do that.

We have argued that thermometry rather than the meter rule might be a more apt analogy for those seeking parallels between what we are trying to do in the human sciences and what measurement in the physical sciences is like. We introduce that analogy towards the end of chapter 1 and continue to develop it as needed in this second edition; to that end, chapter 1 is substantially rewritten. A little confession is in order: We yielded a little to the suggestions of our mathematically literate critics and introduce the key Rasch equation in chapter 2 along with the ICC (Item Characteristic Curve). We did this in the illustrative context of an athletics high jump competition, so readers should not fear that we have forsaken clarity in order to quieten critics. The ICCs return in chapter 3 and in the new chapter 5. Chapter 5's new title, "Invariance: A Crucial Property of Scientific Measurement", exemplifies the new approach we have taken in the second edition. We no longer treat topics such as test linking and differential item functioning as merely functional aspects of psychometrics: We set them in the context of the more crucial measurement property, invariance of item, and person measures. We argue that measures must be invariant; if not, we want theoretically-driven explanations as to why. Chapter 8 on the many-facets Rasch model has now been supplemented by a very useful small empirical example with data set and control lines provided for those who grow confident after managing to repeat the analyses given in the earlier chapters. New examples in chapter 10 look at the growing use of student satisfaction data to hold universities and colleges accountable.

Collaboration with the *Winsteps* developer allows for the inclusion of a comprehensive CD with the second edition. Included is an introductory version of the latest *Winsteps* software, specifically targeted for the beginning Rasch analyst. While the software will take the same large data files as the commercially available version, the output files are constrained to show those figures and tables that are, in our opinion, best suited for the readers of our second edition. Data sets and more comprehensive versions of analysis controls are also provided. Appendix B takes a new approach to helping colleagues on the road to independence: We provide a step-by-step approach to analyzing the BLOT data set. Other material new to the second edition includes a focus on multidimensionality and Rasch factor analysis of residuals in chapter 12.

Needless to say, we remain deeply indebted to colleagues and students who have, in a variety of ways, required us to lift our game for the second edition. First and foremost amongst colleagues in this regard has been Mike Linacre, known by most Rasch users for his software *Winsteps* and *Facets*, but known to us for his unstinting support of our efforts. He graciously shared with us his retyped

marginal notes from his dog-eared hardback first edition, a chapter-by-chapter commentary on strengths and weaknesses augmented by suggestions, advice and exemplars too many to acknowledge individually. He proactively worked with us on the free software CD and answered many technical and mathematical questions at a moment's notice. Thank you, Mike, you have been a generous model colleague. Kelly Bradley and her graduate students at Kentucky University have worked through the bulk of our second edition chapters. At short notice, Kelly volunteered her group to read critically the drafts. They have pointed to aspects needing clarification and to inconsistencies and errors that managed to escape us as we struggled to bring some sense of order to our electronic collaboration. Thanks to Kelly as well as to Shannon Sampson, Kenneth Royal, and Jessica Cunningham in particular for providing us with a perspective we might expect from a large section of our target audience. Lawrence Erlbaum Associates commissioned reviews of our revision plan from George Engelhard (Emory University), Kathy Green (University of Denver) and Randy Penfield (University of Florida); we are indebted to them for their insights and recommendations.

Other colleagues and students have helped us in more general but no less important ways. Along with those who are mentioned in the earlier preface, the following stand out: Terry Brown (Chicago) and Gerald Noelting (Quebec) will be sorely missed—both have died in the past few months. Magdalena Mok (Hong Kong), and fellow Australians, Rosemary Callingham, Peter Congdon, David Curtis, Juho Looveer, and Brian Doig all deserve special mention for their support. Italian colleagues Luigi Tesio (Milano), Enrico Gori (Udine), Silvia Ferrario (Veruno) and Guilio Vidotto (Padova) share highly appreciated passions from Rasch measurement and *cucina italiano*. Jack Stenner, Metametrics and Gage Kingsbury, NWEA, and Moritz Heene have contributed and supported in a variety of ways. Richard Smith, JAMPRESS, remains totally supportive of the publication of the collegial Rasch enterprise. Delegates to IOMW XII 2004 (Cairns), PROMS 2005 (Kuala Lumpur) and 2006 (Hong Kong) provided a wonderful context in which the ideas important to us and to this book could be discussed fruitfully. Our thanks to LEA editor, Debra Riegert, for her support of our endeavors.

So why would we spend time writing and then revising a book about what many consider just one of many analytical techniques that abound in the social sciences? Mark Wilson's review of our first edition made our point cogently for those "who might be otherwise perplexed at the motivation for writing a whole book about a topic that some might characterize as a an 'oversimplified special case' of an item response model"

The point of view in the Rasch approach to measurement is that one must start from a philosophy that establishes certain requirements of the items and the test as a whole, and these translate into requirements for the statistical model that one uses to scale the items—these requirements can then be satisfied by using items that adhere to the structure of the Rasch model. Thus, one's philosophy of measurement leads one to use a statistical analysis model that will guide the development and selection of items—the statistical model is being used as a means of quality control of the items. This is in contrast to the most common alternate

approach where the statistical model is augmented by parameters that are designed to accommodate the characteristics of the item set—one could say that the statistical model is being used to describe the items.

Thus, in Wilson's words, one person's oversimplification is another person's strong measurement philosophy. Indeed, it is the attachment to a strong measurement philosophy that requires us to choose the Rasch model over others. Of course, others will not agree with us. But how is it that many in the human sciences continue to overlook the basic principle of measurement that children understand as necessary from just after they begin grade school?

—Trevor Bond

—Christine Fox