ASSESSING MEASUREMENT INVARIANCE FOR APPLIED RESEARCH

Assessing Measurement Invariance for Applied Research will provide psychometricians and researchers across diverse disciplines in the social sciences with the necessary knowledge and skills to select and apply appropriate methods to assess measurement invariance. It is a user-friendly guide that describes a variety of statistical methods using a pedagogical framework emphasizing conceptual understanding, with extensive illustrations that demonstrate how to use software to analyze real data. A companion website (people.umass.edu/cswells) provides downloadable computer syntax and the data sets demonstrated in this book so readers can use them to become familiar with the analyses and understand how to apply the methods with proficiency to their own work. Evidence-supported methods that can be readily applied to real-world data are described and illustrated, providing researchers with many options from which to select given the characteristics of their data. The approaches include observed-score methods and those that use item response theory models and confirmatory factor analysis.

DR. CRAIG S. WELLS is a professor in the Research, Educational Measurement and Psychometrics program at the University of Massachusetts (UMass) Amherst, where he teaches courses on statistics and psychometrics, including structural equation modeling and psychometric modeling. He is also the associate director for the Center for Educational Assessment at UMass. His research interests pertain to statistical methods for detecting differential item functioning, and assessing the fit of item response theory models. He also has a keen interest in philosophy of science, especially how it relates to hypothesis testing. He coedited *Educational Measurement: From Foundations to Future* (2016) and was president of the Northeastern Educational Research Association in 2017.

EDUCATIONAL AND PSYCHOLOGICAL TESTING IN A GLOBAL CONTEXT

Editor
Neal Schmitt, *Michigan State University*

The Educational and Psychological Testing in a Global Context series features advanced theory, research, and practice in the areas of international testing and assessment in psychology, education, counseling, organizational behavior, human resource management, and all related disciplines. It aims to explore, in great depth, the national and cultural idiosyncrasies of test use and how they affect the psychometric quality of assessments and the decisions made on the basis of measures. Our hope is to contribute to the quality of measurement and to facilitate the work of professionals who must use practices or measures with which they may be unfamiliar or adapt familiar measures to a local context.

Published Titles:
DRAGOŞ ILIESCU,
Adapting Tests in Linguistic and Cultural Contexts, 2017

JOHN C. SCOTT, DAVE BARTRAM and DOUGLAS H. REYNOLDS (eds.),
International Applications of Web-Based Testing: Challenges and Opportunities, 2017

WILLIAM H. SCHMIDT et al.,
Schooling Across the Globe: What We Have Learned from 60 Years of Mathematics and Science International Assessments, 2018

MARIA ELENA OLIVERI and CATHY WENDLER,
Higher Education Admission Practices: An International Perspective, 2020

Forthcoming Titles:
SUMAYA LAHER,
Comparative Histories of Psychological Assessment

ASSESSING MEASUREMENT INVARIANCE FOR APPLIED RESEARCH

CRAIG S. WELLS
University of Massachusetts Amherst

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE
UNIVERSITY PRESS

University Printing House, Cambridge CB2 8BS, United Kingdom

One Liberty Plaza, 20th Floor, New York, NY 10006, USA

477 Williamstown Road, Port Melbourne, VIC 3207, Australia

314–321, 3rd Floor, Plot 3, Splendor Forum, Jasola District Centre, New Delhi – 110025, India

79 Anson Road, #06–04/06, Singapore 079906

Cambridge University Press is part of the University of Cambridge.

It furthers the University's mission by disseminating knowledge in the pursuit of education, learning, and research at the highest international levels of excellence.

www.cambridge.org
Information on this title: www.cambridge.org/9781108485227
DOI: 10.1017/9781108750561

First published 2021

A catalogue record for this publication is available from the British Library.

ISBN 978-1-108-48522-7 Hardback

This book is dedicated to my loving family:
Amanda, Madeline, and Elianna.

Contents

You can download the data files and computer syntax
described in this book at the companion website:
people.umass.edu/cswells

Figures

Cambridge University Press
978-1-108-48522-7 — Assessing Measurement Invariance for Applied Research
Craig S. Wells
Frontmatter
More Information

Cambridge University Press
978-1-108-48522-7 — Assessing Measurement Invariance for Applied Research
Craig S. Wells
Frontmatter
More Information

Tables

Preface

Why write a book about assessing measurement invariance? This question is not about justifying the importance of examining measurement invariance (I address that issue in Chapter 1) but instead about why a book on measurement invariance is useful for social science researchers. In the past 30 years, there have been several excellent texts that have described statistical methods for detecting a lack of measurement invariance (often referred to as differential item functioning), such as Holland and Wainer (1993), Camilli and Shepard (1994), Zumbo (1999), Osterlind and Everson (2009), Engelhard (2012), and Millsap's comprehensive *Statistical Approaches to Measurement Invariance* (2011) in which he provided a unifying theory of measurement invariance. Each of these books provided a valuable contribution to the literature, and we are indebted to the authors and the many other researchers who have added to the theory and practice of assessing measurement invariance. What sets this book apart from those excellent resources on this topic is that, as the title implies, this book focuses on the practical application of assessing measurement invariance, with less emphasis on theoretical development or exposition. As such, a primary emphasis of the book is to describe the methods using a pedagogical framework, followed by extensive illustrations that demonstrate how to use software to analyze real data. My intention is that you will use this book to learn about practical methods to assess measurement invariance and how to apply them to your own data. This book is intended to be a user-friendly guide for anyone who wants to assess measurement invariance in their own work. To support this goal, the computer syntax and data sets used in this book are available for download at the following website: www.people.umass.edu/cswells.

As someone who has taught statistics and psychometrics for many years, I have learned that students learn better when they are actively engaged in the learning process. Therefore, I invite you to download the example data files and syntax and perform the analyses while you read along in this book. My goal is to provide you with enough instruction and practice in the most useful methods so that you can apply what you have learned to novel problems and data sets in your own work.

The audience I envisioned while writing this book is applied researchers, both graduate students and professionals, from the social sciences in which educational tests, psychological inventories, questionnaires, and surveys are administered. This includes researchers from diverse fields, such as education, psychology, public policy, higher education, marketing, sports management, sociology, and public health, to name a few. The purpose of this book is to provide researchers and psychometricians from these diverse fields with the necessary knowledge and skills to select and apply an appropriate method to assess measurement invariance for the problems they are trying to solve and the research questions they are trying to answer. Regardless of your field, I wrote this book assuming you understand basic statistics taught in an introductory statistics course (e.g., mean, standard deviation, correlation, regression, effect size, hypothesis testing, confidence intervals), as well as the concepts of reliability and validity.

This book includes chapters that describe the big ideas and the theoretical concepts of measurement invariance and various statistical procedures, while other chapters include a pedagogical framework for you to learn a new method, including guided practice opportunities. In Chapter 1 I describe the basic ideas of measurement invariance, including why it is important to assess, and I address important issues to consider when applying the methods described in this book. Chapters 2, 4, and 6 are written to teach you specific methods for assessing measurement invariance. The basic structure of each of these chapters is a description of the method, followed by detailed illustrations on how to examine measurement invariance using computer software. Chapter 2 describes observed-score methods where the groups are matched using raw scores. Chapter 4 describes detection methods that use item response theory (IRT) models. Chapter 6 describes methods that rely on confirmatory factor analysis (CFA). For those who are unfamiliar with IRT, Chapter 3 provides a description of the basic concepts of and models used in IRT. I also provide an illustration of how to use the computer program flexMIRT (Cai, 2017) to estimate the parameters for several IRT models. Chapter 5 provides a description of the underlying concepts of CFA that are needed to understand the CFA-based methods described in Chapter 6. Several CFA models are fit to data using the computer program Mplus (Muthén & Muthén, 2008–2017). Finally, Appendix A provides a brief tutorial on the computer program R, because many of the techniques described in this book rely on packages in R.

There are many statistical methods available that I could have included in this book. Given that I was writing for practitioners and not writing a methodological tome that addresses all methods that have been created, I had to strategically select the methods I wanted to address. To accomplish this goal, I decided to include methods that met four criteria. The methods must (1) have been studied empirically and shown to provide reasonably accurate assessment of measurement invariance; (2) be able to incorporate an effect size or some type of information for determining whether the lack of invariance was nontrivial; (3) be relatively simple to use, especially for practitioners who are non-psychometricians; and (4) have software readily available to implement them. Indeed, there are methods that meet these criteria that I may not have included, but there will always be more to learn and write.

One of the challenges in writing a book that encompasses methods from different areas (i.e., observed-score and IRT- and CFA-based detection methods) is that the same symbols can take on different meanings depending on the context. I have attempted to use the symbols in their original meaning, which may cause confusion. For example, the Greek letter α is often used to represent the nominal Type I error rate, common odds ratio, or item discrimination parameter in an IRT model. It is important for the reader to discern the appropriate meaning of a symbol based on the context. To help prevent possible confusion, I have tried to define the symbols when they are first introduced in a chapter.

This book would not have been written if it were not for the support and encouragement of many people. First, I would like to acknowledge my colleagues at UMass Amherst: Steve Sireci, who recommended I write the book and whose support, guidance, and mentoring has had a positive impact on my career; Ronald Hambleton, whose ideas have permeated my understanding of measurement invariance and psychometrics in general; April Zenisky, Lisa Keller, Jennifer Randall, and Scott Monroe, who have all supported me when I needed it. I also would like to acknowledge Allan Cohen, Daniel Bolt, James Wollack, and Michael Subkoviak for their tutelage and persistent support during my doctoral studies; and Ron Serlin, who taught me the value of philosophy of science and the range-null hypothesis, which I have incorporated into DIF detection methods. I would also like to thank Neal Schmitt for giving me the opportunity to write this book, as well as for providing encouraging words along the journey. I would be remiss if I did not acknowledge the hard work and

support of Emily Watton, my editor at Cambridge University Press, for her tireless effort in bringing this manuscript to press, and Clare Diston, my copyeditor, whose attention to detail and knowledge for the English language brought out the best in my writing. Last but not least, I would like to express special thanks to my wife, Amanda, for her encouragement, support, and valuable feedback.