# Preface

More than half a century ago, a far-reaching revolution started in behavioral, educational, and social measurement, which to date has also had an enormous impact on a host of other disciplines ranging from biomedicine to marketing. At that time, item response theory (IRT) began finding its way into these sciences. In many respects, IRT quickly showed important benefits relative to the then conventional approach for developing measuring instruments that was based on "classical" procedures.

Since the 1950s and the influential early work by F. Lord in IRT (for example, Lord [1952, 1953]), more than 60 years have passed that have been filled with major methodological advances in this field and more generally in behavioral and social measurement. The intervening decades have also witnessed an explosion of interest in IRT and item response modeling (IRM) across those disciplines as well as the clinical, biomedical, marketing, business, communication, and cognate sciences. These developments are also a convincing testament to the rich opportunities that this measurement approach offers to empirical scholars interested in assessing various latent constructs, traits, abilities, dimensions, or variables, as well as their interrelationships. The latent variables are only indirectly measurable, however, through their presumed manifestations in observed behavior. This is in particular possible via use of multiple indicators or multi-item measuring instruments, which have become highly popular in the behavioral and social sciences and well beyond them.

This book has been conceptualized mainly as an introductory to intermediate level discussion of IRT and IRM. To aid in the presentation, the book uses the software package Stata. This package offers, in addition to its recently developed IRT command, many and decisive benefits of general purpose statistical analysis and modeling software. After discussing fundamental concepts and relationships of special relevance to IRT, its applications in practical settings with Stata are illustrated using examples from the educational, behavioral, and social sciences. These examples can be readily "translated", however, to similar utilizations of IRM also in the clinical, biomedical, business, marketing, and related disciplines.

We find that several features set our book apart from others currently available in the IRT field. One is that unlike a substantial number of treatments of IRT (in particular older ones), we capitalize on the diverse connections of this field to the comprehensive methodology of latent variable modeling as well as related applied statistics frameworks. In many aspects, it would be fair to view this book as predominantly handling IRT and IRM, somewhat informally stated, as part of the latent variable modeling methodology. In particular, the discussion throughout the book benefits as often as possible from the

conceptual relationships between IRT and factor analysis, specifically, nonlinear factor analysis. Relatedly, whenever applicable, the important links between IRM and other statistical modeling approaches are also pointed out, such as the generalized linear model and especially logistic regression. Another distinguishing feature of the book is that it is free of misconceptions about and incorrect treatments of classical test theory (Zimmerman 1975). Regrettably, they can still be found in some measurement literature and inhibit significantly in our opinion progress in social and behavioral measurement. In addition, these misconceptions contribute to a compartmentalization approach that seems to have been at times followed especially when disseminating or teaching IRT in circles with limited or no prior familiarity with it. That approach and resulting restrictive focus of interest is in our view highly undesirable. The reason is that such an approach has the potential of creating long-term disservice to the cause of behavioral and social measurement. In this connection, we would also like to point out that unlike many previous treatments, this book presents its discussion and developments without any juxtaposition of IRT to classical test theory. This is because IRT does not need this kind of "comparison" and related misconceptions to convince scholars of what it can deliver under its assumptions (see also Raykov and Marcoulides [2016b]). A third characteristic of the book is that it demonstrates the straightforward, user-friendly, and highly effective Stata applications for IRT modeling. We hope to gain in this way many new enthusiasts for this methodological field as well as IRT software across these and related disciplines. Last but not least, our book aims to provide a coherent discussion of IRT and IRM independently of software. The goal is thereby to highlight as often as possible and in as much detail as deemed necessary important concepts and relationships in IRT before moving on to its applications. This was necessary because in our experience, many individuals seem to find some features of this modeling approach more difficult to deal with and use to their advantage than what may be seen as "conventional" applied statistical concepts and relationships. These features include in particular the inherent nonlinearity in studied item-trait relationships as well as produced estimates (predictions) of individual trait levels and measures of uncertainty associated with them. That difficulty in appreciating characteristic properties of IRT may have arguably resulted from insufficient discussion and clarification of them in some alternative accounts or presentations.

This book could be considered aimed mainly at students and researchers with limited or no prior exposure to IRT. However, we are confident that it will also be of interest to more advanced students and scientists who are already familiar with IRT, in particular owing to the above mentioned features in which the book does not overlap with the majority of others available in this field. In addition, a main goal was to enable readers to pursue subsequently more advanced studies of this comprehensive and complex methodological field and its applications in empirical research, as well as to follow more technically oriented literature on IRT and IRM. Relatedly, even though the book uses primarily examples stemming from the educational and behavioral sciences, their treatment, as well as more generally of this measurement field, allows essentially straightforward applications of the used methods and procedures also in other social science settings. These include the clinical, nursing, psychiatry, biomedicine, criminology,

organizational, marketing, and business disciplines (for example, Raykov and Calantone [2014]).

Our book has been influenced substantially by deeply enriching interactions with a number of colleagues over the past years. Special thanks are due to K. L. MacDonald and R. Raciborski for their many instructive inputs on Stata uses and applications in relation to examples used in the book, as well as on IRT and its empirical utilizations in more general terms. The importance of the contributions also of Y. Marchenko and C. Huber cannot be overstated, who provided instrumental support during our work on the book. We are especially indebted to C. Huber for helpful comments and criticism on an earlier version, which contributed markedly to its improvement. His assistance during the book-production phase was similarly invaluable, as was that of the book editor and the production assistant. We also wish to express our particular gratitude to M. D. Reckase, B. O. Muthén, D. M. Dimitrov, M. Edwards, C. Lewis, R. Steyer, S. Rabe-Hesketh, A. Skrondal, and A. Maydeu-Olivares for valuable discussions on IRT and IRM and related applied statistics and measurement approaches. We are similarly thankful to C. Falk, R. J. Wirth, N. Waller, R. Bowles, I. Moustaki, R. D. Bock, S. H. C. duToit, G. T. M. Hult, and J. Jackson for insightful discussions on IRT applications and software. We are also grateful to a number of our students in the courses we taught over the last few years who offered very useful feedback on the lecture notes we first developed for them, from which this book emerged. Last but not least, we are more than indebted to our families for their continued support in lots of ways that cannot be counted. The first author is indebted to Albena and Anna; the second author is indebted to Laura and Katerina.

*Tenko Raykov and George A. Marcoulides*