Applied Psychological Measurement
http://apm.sagepub.com

Book Review: The SAGE Handbook of Quantitative Methodology for the Social Sciences

Derek Briggs

Applied Psychological Measurement 2006; 30; 447
DOI: 10.1177/0146621605280970

The online version of this article can be found at:
http://apm.sagepub.com

Published by:
SAGE Publications
http://www.sagepublications.com

The SAGE Handbook of Quantitative Methodology for the Social Sciences

Edited by David Kaplan
Thousand Oaks, CA: Sage, 2004, 528pp., $130.00 (hardcover)
ISBN 0-7619-2359-4

As the saying by Abraham Maslow goes, "If the only tool you have is a hammer, you tend to see every problem as a nail." With this in mind, a book that characterizes the heterogeneity in both problems (i.e., research questions and data structure) and the available tools to solve these problems (i.e., quantitative methodologies) is the sort of resource that should sit on the bookshelf of any social science researcher. Yet mapping the terrain of quantitative methodology in the social sciences is not so straightforward. We might suspect that if a panel of the top methodologians across social science disciplines were convened for this task, finding consensus would be elusive, with each panel member looking for validation of his or her own areas of specialization. What makes the task so challenging is the necessity, on one hand, of providing a comprehensive and accessible introduction to a broad range of methodological approaches while, on the other hand, integrating the collection of approaches within a larger taxonomy. David Kaplan has in fact convened a panel of top-notch methodologians, who take on the aforementioned challenge in the writing of *The SAGE Handbook of Quantitative Methodology for the Social Sciences (SHQM).* The result is an engrossing collection of chapters that are sure to add screwdrivers, wrenches, and the occasional buzzsaw to your toolbox. That the book falls a bit short of developing an organizing structure for the toolbox may not be so much an indictment of the editor and the authors but a reflection of the state of the field.

The *SHQM* consists of 24 chapters grouped into six sections. For a compact summary of the structure of the *SHQM,* see Table 1. The rows of Table 1 correspond to different statistical models and topics presented in the book, as organized into sections. The first section (chapters 1-3) is on the topic of scaling. These three chapters present dual scaling (Nishisato), multidimensional scaling (Heiser and Busing), and extensions to principal components analysis (Meulman, Van der Kooij, and Heiser) as fundamental methods for the quantification of qualitative observations. The second section (chapters 4-7) is on the topic of testing and measurement. These four chapters present an overview of both classical test theory and item response theory within the context of reliability and validity theory (Zumbo and Rupp), an algorithmic approach to the choice of a taxonomy of unidimensional or multidimensional item response models (Nandakumar and Ackerman), an overview of differential item functioning analysis (Roussos and Stout), and a survey of one of the unique advantages of item response modeling, computerized adaptive testing (Chang). The third section (chapters 8-11) deals with models for categorical data. These four chapters include an overview of trends in categorical data analysis (Rindskopf), followed by presentations on ordinal regression models (Johnson and Albert), latent class models (Magidson and Vermunt), and survival analysis (Willett and Singer). The fourth section (chapters 12-15) deals with models for multilevel data. These four chapters present applications of hierarchical linear models in the context of growth modeling (Hedeker), school effectiveness research

Applied Psychological Measurement, Vol. 30 No. 5, September 2006, 447–451
DOI: 10.1177/0146621605280970

Table 1

The Structure of *The SAGE Handbook of Quantitative Methodology for the Social Sciences* (*SHQM*)

SHQM Model or Topic	*SHQM* Chapter
I. SCALING	
Dual scaling	1
Multidimensional scaling and unfolding	2
Principal components analysis (with nonlinear optimal scaling transformation)	3
II. TESTING & MEASUREMENT	
Validity theory	4
Reliability theory	4
Classical test theory	4
Item response models	4-7
III. MODELS FOR CATEGORICAL DATA	
Log-linear models	8
Logit models	8
Ordinal regression	9
Latent class models	10
Survivor analysis	11
IV. MODELS FOR MULTILEVEL DATA	
Hierarchical linear and generalized linear models	12-15
V. MODELS FOR LATENT VARIABLES	
Structural equation models	16-19, 22, 23
Factor analysis	16
Dynamic factor analysis	18
Growth mixture models	19
VI. FOUNDATIONAL ISSUES	
Probabilistic models using Bayesian networks	20, 24
Significance testing	21
Exogeneity	22
Objectivity in science	23
Causal inference	24

(Rumberger and Palardy), experimental designs in field settings (Seltzer), and meta-analysis (Konstantopoulos and Hedges). The fifth section (chapters 16-19) focuses on models for latent variables. All the topics in these chapters can be cast within the framework of structural equation modeling (SEM): an overview of methods for determining the number of factors in exploratory and confirmatory factor analysis (Hoyle and Duvall), extending factor analysis to handle repeated measures over time (Nesselroade and Molenaar), using SEM in the context of drawing casual inferences from experimental designs (Hancock), and using SEM in conjunction with mixture modeling and latent class analysis in the context of longitudinal data (Muthén). The sixth and final section (chapters 20-24) considers cross-cutting foundational issues. These chapters contrast the frequentist and Bayesian perspectives on

probability (Neapolitan and Morris; Gigerenzer, Krauss, and Vitouch); describe common misconceptions about significance testing (Gigerenzer, Krauss, and Vitouch); define and illustrate a range of assumptions typically made about exogeneity in statistical modeling (Kaplan); discuss the philosophical origins of the concept of scientific objectivity and what this implies about the importance of replication, parameter identification, and parsimony in statistical modeling (Mulaik); and present a view of causal inference as seen through the lens of probabilistic modeling with Bayesian networks (Neapolitan and Morris; Spirtes, Scheines, Glymour, Richardson, and Meek).

A notable strength of the *SHQM* is the generally consistent structure of each chapter. Chapters begin by providing a context, rationale, and background for the methodological approach being introduced. This is followed by a conceptual presentation of the one or more statistical models that underlie the approach. The application of the model(s) is then illustrated with data, and great care is given to assist the reader in parameter interpretation. The final section of each chapter typically concludes by summarizing the key features of the methodological approach while briefly noting more complicated extensions. Each chapter includes a reference section that guides the reader to further information on the chapter topic. The average chapter length is a very manageable 20 pages, with a healthy balance of prose, tables, figures, formulas, and citations.

The chapters in the *SHQM* were written with the intent of being more broadly accessible than the standard article published in *Applied Psychological Measurement.* The target audience for this textbook is identified in the book's preface as applied statisticians, empirical researchers, and graduate students. An implicit assumption seems to be that the reader has had a solid foundation in statistics (i.e., a 1-year introductory sequence at the graduate level). A good endorsement of the accessibility of these chapters comes from this author's experience reading them as a reviewer. For topics where this author had little to no prior background (e.g., scaling, ordinal regression, latent class models), he emerged with enough of an understanding of the approach such that he could engage in a literate conversation with a more knowledgeable colleague. Even when the conceptual overview of a particular model left him with unanswered questions, most of these questions were answered when applications of the model were subsequently illustrated with empirical or simulated data.

A weakness of many of the chapters in the *SHQM* is a lack of critical appraisal. When an empirical researcher pulls the *SHQM* from his or her shelf to get up to speed on, say, state-of-the-art methods for meta-analysis, it is just as important to provide guidelines for when the method is *not* appropriate as it is for the converse. For example, most of the statistical models presented involve parameter estimation through maximum likelihood techniques. Because the consistency and efficiency of maximum likelihood estimates are asymptotic properties, a minimum requirement for the use of these models is that the researcher has access to a sizable sample. Beyond the size of the sample, it matters even more how the sample was selected. If probability sampling methods have not been used, then it is likely that the independence or conditional independence assumptions needed to form a joint likelihood function will be violated. The resulting statistical inferences from the model become problematic. From chapter to chapter, although the authors were careful to list in passing the assumptions of a given model, there was seldom an attempt to connect the plausibility of these assumptions to the ways in which empirical data are likely to be generated. The reader would have been better served if the structure of each chapter included at least one paragraph noting limitations of the model(s). Such a paragraph could have provided references to more detailed expositions of these limitations, along with criticisms, controversies, and alternatives associated with the methodological approach. For example, when presented with a taxonomy of item response models, it seems important to reference both the statistical and the philosophical distinctions between the Rasch model and the two-parameter logistic model (cf. Thissen & Wainer, 2002). Likewise, it was surprising to find no mention of the Neyman-Rubin causal model (cf. Holland, 1986) as a complement or

alternative to the probabilistic directed acyclic graph (DAG) models presented in the chapter on causal inference by Spirtes, Scheines, Glymour, Richardson, and Meek.

Aside from a lack of critical appraisal, as a means of providing an introduction to a "broad array of state-of-the-art quantitative methods in the social sciences" (the aim stated by Kaplan in the book's preface, p. ix), the chapters of the *SHQM* are a worthy accomplishment. The larger challenge is the extent to which the individual chapters have been integrated within and across sections of the book. Not surprisingly, topics within sections are more seamlessly integrated than topics across sections. Section III on models for categorical data starts with an overview chapter by David Rindskopf that ties together nicely the topics presented in the remaining chapters. Similar short overview chapters in the other *SHQM* sections would have been helpful. The coherence of section IV on models for multilevel data is particularly strong because each chapter is essentially an application of the same underlying statistical model. The four chapters in this section are essentially a parsimonious repackaging of chapters 5 to 8 from the book *Hierarchical Linear Models* by Steve Raudenbush and Anthony Bryk (2002). In general, the chapters in each section of the *SHQM* are rather loosely connected, as might be expected when authors write their chapters independently.

There are three notable weaknesses in the structure of the book across sections. First, although the different sections clearly represent a broad array of statistical models, there is no specific section devoted to models for continuous data and no specific chapter that presents the general linear model (GLM) taxonomy. This taxonomy might encompass ordinary and generalized least squares regression along with extensions popular in econometrics such as instrumental variables and two-stage least squares regression. Considering the widespread use of the GLM in applied social science research, this omission is puzzling. It appears the assumption is being made that a strong foundation with the GLM already exists in the target audience, yet there is good reason to believe that both foundational and advanced topics related to the GLM are frequently misunderstood. For a trenchant perspective on this topic, see Berk (2003). Second, although intriguing connections are rather easily drawn between the foundational issues presented by Kaplan and Mulaik in section VI to the statistical models in sections I through V, for the other chapters on foundational issues, the connections to previous chapters are less apparent. For example, none of the models presented in the first five sections employs Bayesian networks, which, if anything, appears to be a foundational issue for the subsequent chapter on causal inference in the same section. Meanwhile, other foundational issues, such as sampling and parameter estimation, are missing altogether. Third, the categorizations of the topics within sections appear somewhat arbitrary. Indeed, a strong case could be made that item response models (section II) could fit just as well in section I, III, IV, or V (cf. de Boeck & Wilson, 2004). Given that these categorizations are not mutually exclusive, it becomes tantamount for the reader to understand the advantages and disadvantages of choosing one methodological approach over the other, given the data at hand. The line distinguishing a variable that is considered continuous from a variable that is considered ordinal is sometimes razor thin. What is lost by turning the problem into a nail? Comparisons across methodological approaches are sometimes made in the *SHQM* chapters (e.g., in the Willett and Singer chapter on survival analysis), but they are not made consistently.

As alluded to at the outset, many of the criticisms raised in this review are probably unavoidable for this sort of book. It is hoped that they do not obscure the fact that the *SHQM* is both well conceived and well executed, providing the reader with numerous insights and a broader sense for the available tools of the quantitative methodological trade. It is most likely that few readers will have the opportunity to read this book from cover to cover, but should they feel so inspired, they will find the effort both rewarding and thought provoking.

Derek Briggs
University of Colorado

References

Berk, R. A. (2003). *Regression analysis: A constructive critique.* Thousand Oaks, CA: Sage.

de Boeck, P., & Wilson, M. (Eds.). (2004). *Explanatory item response models: A generalized linear and nonlinear approach.* New York: Springer.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association, 81*(396), 945-960.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage.

Thissen, D., & Wainer, H. (2001). *Test scoring.* Mahwah, NJ: Lawrence Erlbaum.

Author's Address

Address correspondence to Derek Briggs, University of Colorado, Research and Evaluation Methodology, School of Education, Rm. 211, 249 UCB, Boulder, CO 80309; e-mail: Derek.Briggs@colorado.edu.