Contents

Preface to the First Edition 9

Preface to the Second Edition 13

1. Introduction 15

1.1 Facets of measurement 15

1.2 Purpose and plan of the book 19

2. Rasch Measurement: The Basics 21

2.1 Elements of Rasch measurement 21

2.1.1 The dichotomous Rasch model 21

2.1.2 Polytomous Rasch models 27

2.2 Rasch modeling of many-facet data 28

2.2.1 Putting the facets together 30

2.2.2 The sample data: Essay ratings 33

2.2.3 Rasch modeling of essay rating data 36

3. Rater-Mediated Assessment: Meeting the Challenge 39

3.1 Rater variability 39

3.2 Interrater reliability 42

3.2.1 The standard approach 42

3.2.2 Consensus and consistency 43

3.2.3 Limitations of the standard approach 45

3.3 A conceptual–psychometric framework 48

3.3.1 Proximal and distal facets 50

3.3.2 A measurement approach 52

Bibliografische Informationen
http://d-nb.info/1076090818

digitalisiert durch DEUTSCHE NATIONAL BIBLIOTHEK

4. Many-Facet Rasch Analysis: A First Look 55
4.1 Preparing for a many-facet Rasch analysis 55
4.2 Measures at a glance: The Wright map 57
4.3 Defining separation statistics 60
4.4 Applying separation statistics 63
4.5 Global model fit 67

5. A Closer Look at the Rater Facet: Telling Fact from Fiction 71
5.1 Rater measurement results 71
5.1.1 Estimates of rater severity 71
5.1.2 Rater fit statistics 74
5.1.3 Observed and fair rater averages 81
5.2 Studying central tendency and halo effects 82
5.2.1 Central tendency 83
5.2.2 Halo 86
5.3 Raters as independent experts 89
5.4 Interrater reliability again: Resolving the paradox 92

6. Analyzing the Examinee Facet: From Ratings to Fair Scores 95
6.1 Examinee measurement results 95
6.2 Examinee fit statistics 97
6.3 Examinee score adjustment 102
6.4 Criterion-specific score adjustment 109

7. Criteria and Scale Categories: Use and Functioning 113
7.1 Criterion measurement results 113
7.2 Rating scale structure 115
7.3 Rating scale quality 117

8. Advanced Many-Facet Rasch Measurement 123
8.1 Scoring formats 123
8.2 Dimensionality 124
8.3 Partial credit and hybrid models 127
8.4 Modeling facet interactions 132
8.4.1 Exploratory interaction analysis 133
8.4.2 Confirmatory interaction analysis 140
8.5 Summary of model variants 147

9. Special Issues 151
9.1 Rating designs 151
9.2 Rater feedback 156
9.3 Standard setting 159
9.4 Generalizability theory (G-theory) 163
9.5 MFRM software and extensions 170

10. Summary and Conclusions 173
10.1 Major steps and procedures 173
10.2 MFRM across the disciplines 179
10.3 Measurement and validation 184
10.4 MFRM and the study of rater cognition 189
10.5 Concluding remarks 191

References 193

Author Index 227

Subject Index 235