

# Contents

|                                                                                                            |           |
|------------------------------------------------------------------------------------------------------------|-----------|
| Preface.....                                                                                               | xi        |
| Author biography .....                                                                                     | xiii      |
| <b>1 Conceptual Overview for a Theory of Standard Setting .....</b>                                        | <b>1</b>  |
| 1.1 The Prevalence of Standards and Standard Setting .....                                                 | 1         |
| 1.2 Definitions of a Standard .....                                                                        | 4         |
| 1.3 The Differing Contexts for Setting a Standard.....                                                     | 7         |
| 1.4 Who Defines the Standard? .....                                                                        | 8         |
| 1.5 Qualifications of the Persons Who Participate in the<br>Standard Setting Process .....                 | 9         |
| 1.6 The Standard Setting Process as a Data Collection Design.....                                          | 11        |
| 1.7 Calculating the Value of the Standard .....                                                            | 12        |
| 1.8 Evaluation of the Results of a Standard Setting Process.....                                           | 14        |
| 1.9 Is a Performance Standard an Instructional Goal? .....                                                 | 15        |
| 1.10 The Goals for This Book.....                                                                          | 16        |
| <b>2 A Theory of Standard Setting.....</b>                                                                 | <b>17</b> |
| 2.1 Introduction to a Theory of Standard Setting .....                                                     | 17        |
| 2.2 Basic Concepts .....                                                                                   | 18        |
| 2.2.1 Agency .....                                                                                         | 18        |
| 2.2.2 Policy Definition.....                                                                               | 19        |
| 2.2.3 Content-Specific Definition and Test Specifications .....                                            | 20        |
| 2.2.4 Kernel of the Standard Setting Process .....                                                         | 22        |
| 2.2.5 Estimate of the Intended Standard.....                                                               | 24        |
| 2.2.6 Feedback to Panelists .....                                                                          | 25        |
| 2.2.7 Approval of the Estimated Standard .....                                                             | 26        |
| 2.3 Documentation of the Standard Setting Process .....                                                    | 27        |
| 2.4 The Validity Argument .....                                                                            | 27        |
| 2.5 Some Fanciful Examples of Standard Setting.....                                                        | 28        |
| 2.5.1 How Much Oxygen Is Needed for Healthy Living? .....                                                  | 28        |
| 2.5.2 Understanding of the Functioning of Passenger<br>Automobiles.....                                    | 30        |
| 2.6 Summary and Organization of Subsequent Chapters .....                                                  | 32        |
| <b>3 Agency, Policy Definitions, Content-Specific Definitions and<br/>    Minimum Qualifications .....</b> | <b>35</b> |
| 3.1 The Agency and Its Role in Setting Performance Standards .....                                         | 35        |
| 3.1.1 Involvement .....                                                                                    | 35        |
| 3.1.2 Expertise.....                                                                                       | 36        |
| 3.1.3 Heterogeneity .....                                                                                  | 36        |

|          |                                                                                |           |
|----------|--------------------------------------------------------------------------------|-----------|
| 3.1.4    | Categorization of Agencies.....                                                | 37        |
| 3.1.4.1  | State of Michigan Legislature.....                                             | 37        |
| 3.1.4.2  | National Assessment Governing Board .....                                      | 38        |
| 3.1.4.3  | National Council of State Boards of Nursing.....                               | 39        |
| 3.1.5    | Influence of Agency Characteristics on Standard<br>Setting.....                | 40        |
| 3.2      | Policy Definition ( $D_p$ ).....                                               | 41        |
| 3.3      | Content-Specific Definition ( $D_c$ ) .....                                    | 43        |
| 3.4      | Minimum Qualifications Definition ( $\min D_c$ ) .....                         | 44        |
| 3.5      | Summary.....                                                                   | 45        |
| <b>4</b> | <b>The Construct and the Test Designed to Measure the Construct.....</b>       | <b>47</b> |
| 4.1      | The Target for Test Design and Development .....                               | 47        |
| 4.2      | The Unidimensional Model for Test Development.....                             | 49        |
| 4.3      | Domain Sampling Model for Test Development.....                                | 52        |
| 4.4      | Implications of Test Design for Standard Setting .....                         | 54        |
| 4.5      | Summary.....                                                                   | 55        |
| <b>5</b> | <b>The Kernel of the Standard Setting Process.....</b>                         | <b>57</b> |
| 5.1      | Introduction .....                                                             | 57        |
| 5.2      | Examples of Standard Setting Kernels .....                                     | 59        |
| 5.2.1    | The Bookmark Kernel .....                                                      | 60        |
| 5.2.2    | The Angoff Kernel .....                                                        | 73        |
| 5.2.3    | The Body of Work Kernel .....                                                  | 79        |
| 5.2.4    | The Contrasting Groups Kernel.....                                             | 86        |
| 5.3      | Criteria for Selecting a Standard Setting Kernel.....                          | 89        |
| 5.3.1    | Criterion Number 1 .....                                                       | 89        |
| 5.3.2    | Criterion Number 2 .....                                                       | 90        |
| <b>6</b> | <b>Panelist Selection and Training.....</b>                                    | <b>93</b> |
| 6.1      | Introduction .....                                                             | 93        |
| 6.2      | The Influence of Panelists' Characteristics on the Estimated<br>Standard ..... | 94        |
| 6.3      | Selecting and Recruiting Panelists.....                                        | 97        |
| 6.4      | Statistical Considerations for Selecting Panelists.....                        | 99        |
| 6.5      | Training of Panelists .....                                                    | 103       |
| 6.5.1    | Presenting the "Big Picture" .....                                             | 106       |
| 6.5.2    | Training on Policy and Content Definitions<br>( $D_p$ and $D_c$ ) .....        | 107       |
| 6.5.3    | Training on the Language of the Reporting Score Scale .....                    | 107       |
| 6.5.4    | Training on the Tasks Required of the Standard<br>Setting Kernel.....          | 108       |
| 6.5.5    | Planning and Timing .....                                                      | 109       |
| 6.5.6    | Evaluating Training.....                                                       | 112       |
| 6.6      | Conclusions.....                                                               | 113       |

|          |                                                                                                              |            |
|----------|--------------------------------------------------------------------------------------------------------------|------------|
| <b>7</b> | <b>Feedback and Auxiliary Information.....</b>                                                               | <b>115</b> |
| 7.1      | Introduction .....                                                                                           | 115        |
| 7.1.1    | Categorizations of Feedback and Auxiliary Information.....                                                   | 116        |
| 7.1.2    | Purposes for Providing Feedback .....                                                                        | 118        |
| 7.2      | Types of Feedback.....                                                                                       | 121        |
| 7.2.1    | Feedback about Understanding of the Standard Setting Kernel.....                                             | 122        |
| 7.2.2    | Feedback about the Consequences of the Use of the Estimated Cut Score.....                                   | 132        |
| 7.2.3    | Other Types of Feedback .....                                                                                | 135        |
| 7.3      | Auxiliary Information.....                                                                                   | 138        |
| 7.4      | Timing of Feedback .....                                                                                     | 141        |
| 7.5      | Impact of Feedback on the Standard Setting Process.....                                                      | 141        |
| 7.6      | Conclusions and Recommendations .....                                                                        | 142        |
| <b>8</b> | <b>Estimating the Standard .....</b>                                                                         | <b>145</b> |
| 8.1      | Introduction .....                                                                                           | 145        |
| 8.2      | Assumptions Underlying the Estimation Processes.....                                                         | 148        |
| 8.2.1    | Reporting Score Scale.....                                                                                   | 148        |
| 8.2.2    | Independence of Information from Panelists .....                                                             | 149        |
| 8.2.3    | Other Assumptions .....                                                                                      | 151        |
| 8.3      | The Value that Is Estimated.....                                                                             | 151        |
| 8.3.1    | Estimation Based on Contrasting Groups Data Collection .....                                                 | 154        |
| 8.3.2    | Estimation Based on Test Item Judgment Data .....                                                            | 174        |
| 8.3.2.1  | Effect of Item Parameter Estimation Error on Estimates of Standards.....                                     | 183        |
| 8.3.2.2  | Effect of Panelists' Errors in Estimation of Conditional Probabilities .....                                 | 185        |
| 8.3.2.3  | Including Panelists' Interpretation of $D_c$ and $\min D_c$ in the Estimation of the Intended Standard ..... | 187        |
| 8.3.3    | Estimation Using Compromise Methods.....                                                                     | 189        |
| 8.4      | Recent Developments .....                                                                                    | 196        |
| 8.5      | Cautions about Generalizability .....                                                                        | 197        |
| 8.6      | Summary and Conclusion .....                                                                                 | 198        |
| <b>9</b> | <b>Evaluating the Standard Setting Process .....</b>                                                         | <b>201</b> |
| 9.1      | Introduction .....                                                                                           | 201        |
| 9.2      | Types of Evidence.....                                                                                       | 203        |
| 9.2.1    | Conceptual Coherence Evidence .....                                                                          | 203        |
| 9.2.1.1  | The Policy Definition ( $D_p$ ).....                                                                         | 205        |
| 9.2.1.2  | The Content-Specific Definition ( $D_c$ ).....                                                               | 207        |

|           |                                                                                |            |
|-----------|--------------------------------------------------------------------------------|------------|
| 9.2.1.3   | The Operational Definition of the Intended Construct, $\zeta$ .....            | 208        |
| 9.2.2     | Procedural Evidence.....                                                       | 210        |
| 9.2.2.1   | Kernel Selection and Design of the Process .....                               | 211        |
| 9.2.2.2   | Panelist Capabilities and Training.....                                        | 213        |
| 9.2.2.3   | Implementation and Documentation.....                                          | 214        |
| 9.2.3     | Internal Consistency.....                                                      | 215        |
| 9.2.3.1   | Consistency between Panelists .....                                            | 215        |
| 9.2.3.2   | Consistency over Rounds or with Other Data....                                 | 218        |
| 9.2.3.3   | Full Replication of the Standard Setting Process .....                         | 221        |
| 9.2.3.4   | Decision Consistency.....                                                      | 222        |
| 9.2.4     | External Consistency.....                                                      | 224        |
| 9.3       | The Validity Argument.....                                                     | 226        |
| <b>10</b> | <b>Trends in Standard Setting Design and Development .....</b>                 | <b>229</b> |
| 10.1      | Introduction .....                                                             | 229        |
| 10.2      | Policy and Content Definitions .....                                           | 230        |
| 10.3      | Setting Multiple Standards.....                                                | 231        |
| 10.3.1    | Multiple Standards on the Same Test .....                                      | 231        |
| 10.3.2    | Coordinated Standards for Tests at Different Levels—Vertical Articulation..... | 234        |
| 10.4      | New Approaches to Standard Setting .....                                       | 239        |
| 10.4.1    | Embedded Standard Setting (ESS).....                                           | 239        |
| 10.4.2    | Item Descriptor (ID) Matching Method.....                                      | 241        |
| 10.5      | Standard Setting on New Test Designs .....                                     | 243        |
| 10.6      | Some Final Thoughts.....                                                       | 244        |
|           | <b>References .....</b>                                                        | <b>247</b> |
|           | <b>Index.....</b>                                                              | <b>255</b> |