Handbook of employee selection / editors, James L. Farr, Nancy T. Tippins.
1. Employee selection. 2. Employee selection--Handbooks, manuals, etc. I. Farr, James L. II. Tippins, Nancy Thomas, 1950-
HF5549.5.S38H36 2010
658.3'112--dc22
2009052677

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the Psychology Press Web site at
http://www.psypress.com

Contents

Preface.............................................................................................................................................xix
Editors.............................................................................................................................................xxi
Contributors ................................................................................................................................. xxiii

Chapter 1  Handbook of Employee Selection: An Introduction and Overview ...................... 1
James L. Farr and Nancy T. Tippins
Brief History of Science- and Data-Based Employee Selection ........................................ 1
Structure of Handbook of Employee Selection .................................................................... 3
References ................................................................................................................................. 6

PART 1  Foundations of Psychological Measurement and Evaluation Applied to Employee Selection
John P. Campbell and Frank J. Landy, Coeditors

Chapter 2  Reliability and Validity ........................................................................................... 9
Dan J. Putka and Paul R. Sackett
Overview ......................................................................................................................................... 10
Reliability ..................................................................................................................................... 11
Role of Measurement Models ................................................................................................. 14
Estimation of Reliability .......................................................................................................... 22
Emerging Perspectives on Measurement Error ..................................................................... 36
Closing Thoughts on Reliability ............................................................................................ 37
Concept of Validity .................................................................................................................. 38
References ................................................................................................................................... 45

Chapter 3  Validation Strategies for Primary Studies ............................................................. 51
Neal W. Schmitt, John D. Arnold, and Levi Nieminen
Nature of Validity .................................................................................................................... 51
Validation in Different Contexts ............................................................................................. 53
Special Considerations in Criterion-Related Studies ........................................................... 59
Concerns About the Quality of the Data: Cleaning the Data ............................................... 60
Modes of Decision-Making and the Impact on Utility and Adverse Impact ....................... 61
Scientific or Long-Term Perspective: Limitations of Existing Primary Validation Studies, Including the Current Meta-Analytic Database ........................................................................ 64
Conclusions ........................................................................................................................... 67
References ................................................................................................................................. 67

viii Contents

Chapter 4 Work Analysis ............................................................................................................ 73
   Kenneth Pearlman and Juan I. Sanchez
   Traditional Selection-Related Applications of Work Analysis ........................................ 73
   A Review of Major Work Analysis Methods and Approaches .................................... 75
   Key Work Analysis Practice Issues ................................................................................. 83
   Frontiers of Work Analysis: Emerging Trends and Future Challenges ....................... 88
   Synopsis and Conclusions ............................................................................................... 94
   References ...................................................................................................................... 94

Chapter 5 Current Concepts of Validity, Validation, and Generalizability ......................... 99
   Jerard F. Kehoe and Kevin R. Murphy
   Converging Trends in Psychometrics—Toward a Unified Theory of Validity ........... 100
   Reliability, Validity, and Generalizations From Test Scores ..................................... 101
   Validation Process: Linking Tests With Their Uses .................................................. 102
   Generalizing Research-Based Inferences to Implementation Decisions ................. 108
   References ................................................................................................................... 120

PART 2 Implementation and Management of Employee Selection Systems in Work Organizations
   Jerard F. Kehoe and Robert E. Ployhart, Coeditors

Chapter 6 Attracting Job Candidates to Organizations .................................................. 127
   Ann Marie Ryan and Tanya Delany
   Reaching Potential Applicants ..................................................................................... 127
   Sourcing Globally ......................................................................................................... 130
   Technology as a Means of Improving Reach ............................................................. 133
   Considering Strategic Talent Management ................................................................. 134
   Maintaining Interest .................................................................................................... 136
   Maintaining Interest Around the Globe ...................................................................... 139
   Technology’s Role in Maintaining Interest ................................................................. 140
   Maintaining the Interest of Targeted Talent ............................................................... 141
   Accepting Offers .......................................................................................................... 142
   Conclusions .................................................................................................................. 145
   References ................................................................................................................... 146

Chapter 7 Test Administration and the Use of Test Scores .............................................. 151
   Jeff W. Johnson and Frederick L. Oswald
   Use of Test Scores ........................................................................................................ 151
   Decisions to Make Before Collecting Test Scores ..................................................... 151
   Collection of Test Scores ............................................................................................ 153
   Computation of Test Scores ....................................................................................... 159

## Contents

Making Selection Decisions ................................................................. 163  
Conclusions ......................................................................................... 166  
References ........................................................................................... 166

### Chapter 8  
**Technology and Employee Selection** ........................................... 171  
*Douglas H. Reynolds and David N. Dickter*  
Introduction: The Changing Relationship Between Industrial-Organizational Psychology and Information Technology .................................................. 171  
Critical Issues in Technology-Based Selection ...................................... 176  
Implementation Issues ........................................................................ 186  
Future Directions ................................................................................ 189  
References ........................................................................................... 192

### Chapter 9  
**Strategy, Selection, and Sustained Competitive Advantage** ............ 195  
*Robert E. Ployhart and Jeff A. Weekley*  
Why Personnel Selection Must Show Business-Unit-Level Value ............ 196  
SHRM .................................................................................................. 197  
Alignment of Selection and Strategy .................................................... 200  
Selection’s Potential Contribution to Business Unit Value .................... 203  
Selection’s Symbiotic Relationship With Other HR Activities ............... 208  
Conclusions ......................................................................................... 209  
Acknowledgments ................................................................................ 209  
References ........................................................................................... 210

### Chapter 10  
**Managing Sustainable Selection Programs** .................................. 213  
*Jerard F. Kehoe, Stefan T. Mol, and Neil R. Anderson*  
Organization Context for Selection ....................................................... 213  
Defining Selection System Sustainability .............................................. 214  
HR Technology .................................................................................... 225  
Conclusions ......................................................................................... 233  
References ........................................................................................... 233

### Chapter 11  
**The Business Value of Employee Selection** ................................... 235  
*Wayne F. Cascio and Lawrence Fogli*  
Traditional Model of Employee Selection .............................................. 235  
Challenges to the Business Value of the Traditional Approach .............. 236  
Dynamic, Contemporary Approach to Selection as an Organizational Process ......................................................................................... 237  
Importance of Social Context and Interpersonal Processes in Selection Decisions ..................................................................................... 240  
Conclusions: How Should We Value the Success of Selection in Organizations? .......................................................................................... 248  
References ........................................................................................... 250

PART 3  Categories of Individual Difference Constructs for Employee Selection

David Chan and Leaetta Hough, Coeditors

Chapter 12  Cognitive Abilities

Deniz S. Ones, Stephan Dilchert, Chockalingam Viswesvaran, and Jesús F. Salgado

History, Current Usage, and Acceptability of Cognitive Ability Measures in Employee Selection........................................... 255
Definitions and Theoretical Underpinnings ................................................................. 258
Structure ......................................................................................................................... 258
Measurement .................................................................................................................... 259
Criterion-Related Validity Evidence ........................................................................ 261
Group Differences on Cognitive Ability Measures..................................................264
Future Challenges for Research and Practice ..........................................................269
Epilogue ....................................................................................................................... 270
References .................................................................................................................. 270

Chapter 13  Physical Performance Tests

Deborah L. Gebhardt and Todd A. Baker

Job Analysis for Arduous Jobs ................................................................................. 277
Physical Performance Tests ...................................................................................... 281
Validity of Physical Performance Tests ....................................................................285
Test Scoring and Administration .............................................................................287
Legal Issues .............................................................................................................. 291
Benefits of Physical Testing ......................................................................................294
References ................................................................................................................ 294

Chapter 14  Personality: Its Measurement and Validity for Employee Selection

Leaetta Hough and Stephan Dilchert

Structure of Personality Variables ........................................................................... 299
Measurement Methods ............................................................................................. 301
Validity of Personality Constructs and Other Factors That Affect Their Usefulness .................................................... 308
Conclusions .............................................................................................................. 312
References ................................................................................................................ 312

Chapter 15  Values, Styles, and Motivational Constructs

David Chan

Values ....................................................................................................................... 321
Cognitive Styles ....................................................................................................... 324
Motivational Constructs .......................................................................................... 327
Practical Considerations and Future Research Challenges ...................................... 332
Epilogue ................................................................................................................. 335
References ............................................................................................................. 336

## Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Practical Intelligence, Emotional Intelligence, and Social Intelligence</td>
<td>339</td>
</tr>
<tr>
<td></td>
<td><em>Filip Lievens and David Chan</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Definitions and Conceptualizations</td>
<td>339</td>
</tr>
<tr>
<td></td>
<td>Measurement Approaches</td>
<td>344</td>
</tr>
<tr>
<td></td>
<td>Conceptual Framework for Examining Practical, Emotional, and Social Intelligence</td>
<td>349</td>
</tr>
<tr>
<td></td>
<td>Strategies for Future Research</td>
<td>351</td>
</tr>
<tr>
<td></td>
<td>Epilogue</td>
<td>354</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>355</td>
</tr>
<tr>
<td>17</td>
<td>Decisions in Developing and Selecting Assessment Tools</td>
<td>363</td>
</tr>
<tr>
<td></td>
<td><em>Nancy T. Tippins, Jone M. Papinchock, and Emily C. Solberg</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Introduction</td>
<td>363</td>
</tr>
<tr>
<td></td>
<td>Which Constructs Should Be Measured?</td>
<td>364</td>
</tr>
<tr>
<td></td>
<td>How Should the Constructs Be Measured?</td>
<td>366</td>
</tr>
<tr>
<td></td>
<td>How Should Validity Evidence Be Gathered?</td>
<td>371</td>
</tr>
<tr>
<td></td>
<td>How Should Scores Be Used?</td>
<td>373</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>375</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>375</td>
</tr>
<tr>
<td>18</td>
<td>Administering Assessments and Decision-Making</td>
<td>377</td>
</tr>
<tr>
<td></td>
<td><em>R. Stephen Wunder, Lisa L. Thomas, and Zupei Luo</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Unproctored Internet Testing</td>
<td>377</td>
</tr>
<tr>
<td></td>
<td>Combination of Candidate Information: The Role of the Decision-Maker</td>
<td>390</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>396</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>396</td>
</tr>
<tr>
<td>19</td>
<td>Evaluation of Measures: Sources of Error, Sufficiency, and Contamination</td>
<td>399</td>
</tr>
<tr>
<td></td>
<td><em>Michael J. Zickar, Jose M. Cortina, and Nathan T. Carter</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reliability</td>
<td>400</td>
</tr>
<tr>
<td></td>
<td>Validity</td>
<td>402</td>
</tr>
<tr>
<td></td>
<td>Sources of Invariance: A Handful of Hypotheses</td>
<td>407</td>
</tr>
<tr>
<td></td>
<td>CFA MI/E in Selection Research</td>
<td>411</td>
</tr>
<tr>
<td></td>
<td>Conclusions</td>
<td>413</td>
</tr>
<tr>
<td></td>
<td>References</td>
<td>414</td>
</tr>
<tr>
<td>20</td>
<td>Assessment Feedback</td>
<td>417</td>
</tr>
<tr>
<td></td>
<td><em>Manuel London and Lynn A. McFarland</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Some Case Examples</td>
<td>417</td>
</tr>
</tbody>
</table>

PART 5  Criterion Constructs in Employee Selection

Kevin R. Murphy and Elaine D. Pulakos, Coeditors

Chapter 21  The Measurement of Task Performance as Criteria in Selection Research .......................... 439
Walter C. Borman, Rebecca H. Bryant, and Jay Dorio

Objective Criteria ................................................................................................................. 439
Subjective Criteria ................................................................................................................. 441
Dimensionality of Job Performance .................................................................................... 447
Predictors of Task Performance Dimensions ...................................................................... 454
References .......................................................................................................................... 458

Chapter 22  Adaptive and Citizenship-Related Behaviors at Work .............................................. 463
David W. Dorsey, Jose M. Cortina, and Joseph Luchman

Conceptualization ................................................................................................................. 463
Adaptive Behavior Defined .................................................................................................. 463
Citizenship Defined .............................................................................................................. 465
Individual Difference Predictors .......................................................................................... 468
Measurement ......................................................................................................................... 474
Moderators and Mediators—Adaptability ............................................................................ 476
Moderators and Mediators—Citizenship ............................................................................. 477
Impact on Organizational Outcomes .................................................................................... 479
Too Much of a Good Thing? ............................................................................................... 480
Conclusions ......................................................................................................................... 481
References .......................................................................................................................... 482

Chapter 23  Counterproductive Work Behavior and Withdrawal .............................................. 489
Maria Rotundo and Paul E. Spector

Nature of CWB ..................................................................................................................... 489
Assessment of CWB ............................................................................................................. 491
Potential Antecedents of CWB ............................................................................................ 492
Environment .......................................................................................................................... 499
Person and Environment ...................................................................................................... 502
Consequences ......................................................................................................................... 503
Future Directions .................................................................................................................. 505
Conclusions ......................................................................................................................... 506
References .......................................................................................................................... 506

Contents xiii

Chapter 24 Defining and Measuring Results of Workplace Behavior ........................................ 513
   Elaine D. Pulakos and Ryan S. O’Leary
   Measuring Workplace Behavior Versus Results ......................................................... 514
   Defining Individual Performance Objectives ............................................................ 516
   Challenges Associated With Developing Individual Objectives and Mitigation Strategies ........................................................................................................ 519
   Measuring Results of Performance Objectives ....................................................... 522
   Challenges Associated With Measuring Results and Mitigation Strategies .............. 524
   Individual Difference Predictors of Results ............................................................. 526
   Conclusions .............................................................................................................. 527
   References .............................................................................................................. 528

Chapter 25 Employee Work-Related Health, Stress, and Safety ........................................ 531
   Lois E. Tetrick, Pamela L. Perrewé, and Mark Griffin
   Healthy Workers ....................................................................................................... 531
   Work Stress ............................................................................................................... 535
   Occupational Safety .................................................................................................. 541
   Conclusions .............................................................................................................. 545
   References .............................................................................................................. 546

Chapter 26 Criterion Validity and Criterion Deficiency: What We Measure Well and What We Ignore ......................................................... 551
   Jeanette N. Cleveland and Adrienne Colella
   Work and Workforce in the 21st Century: Outmoded Assumptions and Bases for Change .......................................................................................................... 552
   Criterion Problem in I-O Psychology ....................................................................... 552
   Multilevel Issues in Defining Performance and Success ......................................... 558
   “Closing In” on Criterion Deficiency: One Approach to Bridging HR Systems With Business Unit Strategy ................................................................. 562
   Conclusions .............................................................................................................. 563
   References .............................................................................................................. 564

PART 6 Legal and Ethical Issues in Employee Selection
   P. Richard Jeanneret and Paul R. Sackett, Coeditors

Chapter 27 Ethics of Employee Selection ........................................................................ 571
   Joel Lefkowitz and Rodney L. Lowman
   Some Meta-Issues ..................................................................................................... 571
   Ethical Principles and Dilemmas ............................................................................. 576
   Role of Ethical Codes in Professional Practice: Historical and Current Perspectives .......................................................................................................... 580
   Some Specific Issues and Sources of Ethical Problems ............................................ 582
   Conclusions .............................................................................................................. 587
   References .............................................................................................................. 589

## Chapter 28 Professional Guidelines/Standards

*P. Richard Jeanneret and Sheldon Zedeck*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>593</td>
</tr>
<tr>
<td>Standards for Educational and Psychological Testing</td>
<td>595</td>
</tr>
<tr>
<td>Principles for the Validation and Use of Personnel Selection Procedures</td>
<td>604</td>
</tr>
<tr>
<td>Uniform Guidelines on Employee Selection Procedures</td>
<td>610</td>
</tr>
<tr>
<td>Comparisons Among the Three Authorities</td>
<td>615</td>
</tr>
<tr>
<td>Final Thoughts</td>
<td>621</td>
</tr>
<tr>
<td>Conclusions</td>
<td>621</td>
</tr>
<tr>
<td>References</td>
<td>623</td>
</tr>
</tbody>
</table>

## Chapter 29 A Sampler of Legal Principles in Employment Selection

*Frank J. Landy, Arthur Gutman, and James L. Outtz*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>627</td>
</tr>
<tr>
<td>Principles and Exemplar Case Law</td>
<td>629</td>
</tr>
<tr>
<td>Conclusions</td>
<td>648</td>
</tr>
<tr>
<td>References</td>
<td>648</td>
</tr>
</tbody>
</table>

## Chapter 30 Perspectives From Twenty-Two Countries on the Legal Environment for Selection

*Paul R. Sackett, Winny Shen, Brett Myors, and Colleagues*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data Collection Methodology</td>
<td>652</td>
</tr>
<tr>
<td>Discussion</td>
<td>673</td>
</tr>
<tr>
<td>Authors’ Note</td>
<td>675</td>
</tr>
<tr>
<td>References</td>
<td>675</td>
</tr>
</tbody>
</table>

## PART 7 Employee Selection in Specific Organizational Contexts

*Rick Jacobs and Ann Howard, Coeditors*

## Chapter 31 Selection and Classification in the U.S. Military

*Wayne S. Sellman, Dana H. Born, William J. Strickland, and Jason J. Ross*

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Military Personnel System</td>
<td>679</td>
</tr>
<tr>
<td>Indicators of Recruit Quality</td>
<td>680</td>
</tr>
<tr>
<td>Need for Military Selection</td>
<td>685</td>
</tr>
<tr>
<td>Short History of Military Personnel Testing (Pre-All Volunteer Force)</td>
<td>686</td>
</tr>
<tr>
<td>Moving to an All-Volunteer Force</td>
<td>686</td>
</tr>
<tr>
<td>ASVAB Misnornning and Job Performance Measurement Project</td>
<td>688</td>
</tr>
<tr>
<td>Enlisted Selection and Classification in Today’s Military</td>
<td>691</td>
</tr>
<tr>
<td>Enlistment Process</td>
<td>692</td>
</tr>
<tr>
<td>Recruit Quality Benchmarks and Enlistment Standards</td>
<td>693</td>
</tr>
<tr>
<td>Selection for Officer Commissioning Programs</td>
<td>693</td>
</tr>
<tr>
<td>Officer Retention and Attrition</td>
<td>695</td>
</tr>
<tr>
<td>Officer Executive Development</td>
<td>696</td>
</tr>
</tbody>
</table>

Contents

Command Selection and Career Broadening Experiences ................................................. 696
Defense Transformation in Military Selection ................................................................. 697
Conclusions ..................................................................................................................... 698
References ....................................................................................................................... 699

Chapter 32 Public Sector Employment........................................................................ 705

Rick Jacobs and Donna L. Denning

Position Classification in the Public Sector ................................................................. 706
Civil Service Examinations .......................................................................................... 706
Linking Tests to Jobs: Validation and Its Many Forms .................................................... 708
Creating a Talent Pipeline: Recruiting Candidates ....................................................... 711
Promotional Processes: Using What We Know About People and Their Capabilities to Our Advantage ................................................................. 712
Personnel Decision-Making and Legal Jeopardy ......................................................... 714
Conclusions ..................................................................................................................... 718
References ....................................................................................................................... 719

Chapter 33 Selection Methods and Desired Outcomes: Integrating Assessment Content and Technology to Improve Entry- and Mid-Level Leadership Performance ................................................................. 721

Scott C. Erker, Charles J. Cosentino, and Kevin B. Tamanini

Current Business Trends Affecting Leadership Selection and Development .......... 722
Changing Behavioral Requirements for Leaders ............................................................ 725
Assessment Tools and Techniques .............................................................................. 728
Case Studies of Leadership Selection ........................................................................... 733
Conclusions ..................................................................................................................... 737
References ....................................................................................................................... 738

Chapter 34 Blue-Collar Selection in Private Sector Organizations ............................... 741

Wanda J. Campbell and Robert A. Ramos

Definition of Blue-Collar Jobs ...................................................................................... 741
Environment ..................................................................................................................... 741
Planning and Developing the Selection Process: Psychometric and Practical Considerations ......................................................................................................................... 744
Planning and Developing the Selection Process: The Constituents, Their Issues, and Preferences ................................................................................................................. 749
Implementing a Selection Procedure .............................................................................. 754
Maintaining a Selection Procedure .............................................................................. 756
Recruitment and Employee Development .................................................................... 759
References ....................................................................................................................... 762

Chapter 35 Selection for Service and Sales Jobs ........................................................... 765

John P. Hausknecht and Angela M. Langevin

Nature of Service and Sales Work ..................................................................................... 765

Contents

Research on Selection for Service and Sales Workers ............................................. 768
Implications for Practice and Future Research ...................................................... 774
Conclusions .............................................................................................................. 776
References ................................................................................................................ 776

Chapter 36 Selection in Multinational Organizations ............................................. 781

Paula Caligiuri and Karen B. Paul

Employee Selection in Multinational Organizations ............................................. 781
Challenges of Cross-National Selection and Assessment ...................................... 785
International Assignments: Implications for Employee Selection ...................... 789
Conclusions .............................................................................................................. 795
References ................................................................................................................ 796

Chapter 37 Selection for Team Membership: A Contingency and Multilevel Perspective ............. 801

Susan Mohammed, Jan Cannon-Bowers, and Su Chuen Foo

Conceptual Framework for Understanding Selection for Team Membership......... 802
Individual-Level Considerations ............................................................................ 804
Team-Level Considerations .................................................................................... 809
Discussion ................................................................................................................. 815
Conclusions .............................................................................................................. 817
References ................................................................................................................ 818

Chapter 38 Selecting Leaders: Executives and High Potentials ................................ 823

George C. Thornton III, George P. Hollenbeck, and Stefanie K. Johnson

Executives and High Potentials: Who Are They? What Do They Do? .................... 823
Executive Competencies and Attributes ................................................................. 824
Assessment Techniques ......................................................................................... 826
Executive Selection in an HRM System ................................................................. 830
Performance Reviews: The Role of Top Executives and Boards ......................... 830
Fit: Individuals, Team, Outcome, Followers, and Culture ..................................... 831
Does It Work? ......................................................................................................... 833
Conclusions .............................................................................................................. 834
References ................................................................................................................ 835

PART 8 Milestones in Employee Selection

Walter C. Borman and Benjamin Schneider, Coeditors

Chapter 39 The Management Progress Study and Its Legacy for Selection .......... 843

Ann Howard

The Foundations of the Assessment Center ........................................................... 843
The Beginnings of Managerial Assessment ............................................................ 845
Predicting Managerial Success ............................................................................. 847
Successful Managers’ Development Over Time .................................................... 852
AC Advantages and Disadvantages ...................................................................... 855

Contents

The MPS Selection Legacy ................................................................. 858
References ....................................................................................... 863

Chapter 40  Project A: 12 Years of R & D ............................................. 865

John P. Campbell and Deirdre J. Knapp

Origins of Project A ........................................................................... 865
Enabling of Project A ......................................................................... 866
Specific Research Objectives ............................................................. 867
Overall Research Design .................................................................... 867
Research Instrument Development: Predictors .................................... 869
Job Analyses and Criterion Development ........................................... 873
Modeling the Latent Structure of Performance ................................. 875
Correlations of Past Performance With Future Performance .............. 878
Criterion-Related Validation ............................................................... 879
Some Broader Implications ............................................................... 882
Conclusions ...................................................................................... 884
References ....................................................................................... 885

Chapter 41  The Dictionary of Occupational Titles and the Occupational Information Network ........................................ 887

Norman Peterson and Christopher E. Sager

Milestones ......................................................................................... 887
Organizing Framework ...................................................................... 887
DOT ...................................................................................................... 889
DOT to O*NET™ Transition ............................................................... 894
Development and Testing of a Prototype O*NET ................................ 895
O*NET Becomes Operational ............................................................. 899
Conclusions ...................................................................................... 905
References ....................................................................................... 906

Chapter 42  Situational Specificity and Validity Generalization ............ 909

Lawrence R. James and Heather H. McIntyre

Situational Specificity ........................................................................ 910
VG ...................................................................................................... 912
Situational Specificity Response to VG .............................................. 914
Substantive Tests of Situational Specificity ........................................ 917
Conclusions ...................................................................................... 918
References ....................................................................................... 919

Chapter 43  Employee Selection in Europe: Psychotechnics and the Forgotten History of Modern Scientific Employee Selection .... 921

Jesús F. Salgado, Neil R. Anderson, and Ute R. Hülsheger

Introduction ........................................................................................ 921
The European Origin of Scientific Employee Selection: The Years of Success (1900–1945) ................................................................. 921

# Contents

The Fall of the European Employee Selection Enterprise After World War II: The Years of Decline and Stagnation (1945–1975) .......................... 931

The Resurgence of European Employee Selection: The Years of Optimism and Growth (Since 1975)................................................................. 934

Conclusions .............................................................................................................. 936

Acknowledgments .................................................................................................... 937

References ................................................................................................................ 937

## Chapter 44 Employee Selection: Musings About Its Past, Present, and Future................. 943

*Robert M. Guion*

Psychometric Musings.............................................................................................. 943

Musings on the Perpetual Disconnect: Research and Practice ....................... 950

Is Education Really Broadening? ........................................................................ 954

A Final Musing ......................................................................................................... 956

References ................................................................................................................ 956

Author Index ............................................................................................................. 959

Subject Index ............................................................................................................ 983
Validation Strategies for Primary Studies

Neal W. Schmitt, John D. Arnold, and Levi Nieminen

NATURE OF VALIDITY

Most early applications of the use of tests as decision-making tools in the selection of personnel in work organizations involved a validation model in which the scores on tests were correlated with some measure or rating of job performance, such as the studies of salespersons by Scott (1915) and streetcar motormen by Thorndike (1911). This view of validity was reinforced in books by Hull (1928) and Viteles (1932). Subsequent reviews by Ghiselli (1966, 1973) were similarly focused on what was by then known as criterion-related validity.

During this time, there was a recognition that tests could and should be based on other logical arguments as well. Standards for Educational and Psychological Tests (American Psychological Association [APA], American Educational Research Association [AERA], and National Council on Measurement in Education [NCME], 1954) identified four aspects of validity evidence: content, predictive, concurrent, and construct. With time, the predictive and concurrent aspects of validity became seen as simply different research designs, the purpose of which was to establish a predictor-criterion relationship; hence, they became known as criterion-related validity. Content and construct validation were seen as alternate methods by which one could validate and defend the use of test scores in decision-making. A much broader view of the nature of validity is accepted today, and in general it is seen as the degree to which the inferences we draw from a set of test scores about job performance are accurate.

TRINITARIAN VERSUS UNITARIAN IDEAS OF VALIDITY

The “trinitarian” approach to validity was popularized and incorporated in the 1974 revision of the APA Standards. This conceptualization of validity holds that there are three approaches to the validation of tests. Content validation involves a demonstration that test items are a representative sample of the behaviors to be exhibited in some performance domain. Content validation typically depends heavily on a job analysis that specifies the tasks performed on a job and how those tasks (or very similar tasks) are reflected in the tests used to make decisions. Criterion-related validation involves the demonstration that scores on a test are related to job performance measures. If job performance measures and test scores are obtained from job incumbents at approximately the same time, the study involves what has become known as concurrent criterion-related validity. If test scores are collected from job applicants and performance measures are collected some time after these individuals are hired, the study represents a test of predictive criterion-related validity. Thousands of criterion-related studies have been summarized in meta-analyses over the last 30 years, and Schmidt and Hunter (1998) have summarized these meta-analyses. Construct validation often includes a series of studies or collections of evidence that a psychological concept or construct explains test performance.
addition, there should be support for the notion that the construct is central to the performance of job tasks.

This separation of approaches to validity produced numerous problems, not the least of which was the notion that there were times when one approach was to be preferred over another or that there were different acceptable standards by which different aspects of validity were to be judged. Most important, however, was the realization on the part of measurement scholars that all were aspects of construct validity—the theoretical reasonableness of our explanations of job behavior. There was a realization that the inferences we derive from test scores was central to all validation work. Content validity and the practices usually associated with it were recognized as desirable practices in the development of any test. Careful consideration of the “theory” and hypotheses that underlie our conceptualization of performance and how the constructs central to job performance are represented in our tests is always important and unifying insofar as our validation efforts are concerned. Traditional criterion-related research represents one type of evidence that can be collected to confirm/disconfirm these hypotheses. This “unitarian” approach to validity was strongly argued in the 1985 Standards and has been incorporated most completely in the 1999 version of the Standards (AERA, APA, & NCME, 1999).

**DIFFERENT APPROACHES TO THE COLLECTION OF DATA ABOUT VALIDITY**

Validity, as defined in the most recent version of the Standards (1999), is the degree to which evidence and theory support the interpretation of test scores proposed by the test user. The user must state explicitly what interpretations are to be derived from a set of test scores including the nature of the construct thought to be measured. The document goes on to describe a variety of evidence that can support such an interpretation. An evaluation of test themes, wording, item format, tasks, and administrative guidelines all comprise the “content” of a test, and a careful logical or empirical analysis of the relationship of this content to the construct measured as well as expert judgments about the representativeness of the items to the construct measured supports validity.

Validity evidence can also take the form of an examination of the response processes involved in responding to an item. For example, in evaluating the capabilities of an applicant for a mechanical job, we might ask the person to read a set of instructions on how to operate a piece of equipment and then ask the applicant to demonstrate the use of the equipment. Because the equipment is used on the job, it would seem valid, but suppose we also find that test scores are highly related to examinees' vocabulary level. We would then want to know if vocabulary is necessary to learn how to use this equipment on the job and depending on the answer to that question, we may want to revise the test.

Yet a third piece of evidence might be to collect data regarding the internal structure of a test. We would examine the degree to which different items in a test (or responses to an interview) yield correlated results and whether items designed to measure different constructs can be differentiated from items written to assess a different construct. Researchers interested in these questions use item means, standard deviations, and intercorrelations as well as exploratory and confirmatory analyses to evaluate hypotheses about the nature of the constructs measured by a test.

Similar to looking at the internal structure of a test, researchers can also examine its external validity by correlating the test with measures of theoretically similar or dissimilar constructs as well as job performance measures that are hypothesized correlates of test scores. Validity in the personnel selection area has been almost synonymous with the examination of the relationship between test scores and job performance measures, most often referred to as criterion-related validity. Because there is a large body of primary studies of many job performance-test relationships, one can also examine the extent to which tests of similar constructs are related to job performance and generalize in a way that supports the validity of a new measure or an existing measure in a new context. These are studies of validity generalization.
Finally, and somewhat controversial among industrial-organizational (I-O) psychologists, the Standards (1999) also suggest that researchers examine the intended and unintended consequences of test use to make decisions. This evidence is referred to as consequential validity (Messick, 1998). The consequences of most concern are the degree to which use of test scores results in disproportionate hiring of one or more subgroups (e.g., gender, race, disabled). The 1999 version of the Standards clearly reflects a unitarian view of validation; the old tripartite notions and terms are not mentioned.

The Society for Industrial and Organizational Psychology’s Principles (2003) described the sources of validity evidence in a way that more reflects the trinitarian approach to validation in that it includes a relatively lengthy discussion of criterion-related evidence, evidence based on content, and evidence based on the internal structure of tests and their relationships to other variables. The Principles, like the Standards, also recognized the central role of psychological constructs in all validation research.

Finally, some I-O psychologists have also noted that the traditional separation of reliability and validity concepts may be inadequate (Lance, Foster, Gentry, & Thoresen, 2004; Lance, Foster, Nemeth, Gentry, & Drollinger, 2007; Murphy & DeShon, 2000). Their ideas are addressed in Chapter 2 of this volume. It is also the case that technology affords the opportunity to make the traditional one-time criterion-related validity study an ongoing effort in which the accumulation of predictor and criterion data can be collected and aggregated across time and organizations.

VALIDATION IN DIFFERENT CONTEXTS

This chapter will discuss validation largely within the context of personnel selection. This is the most common application of the various approaches to validation. It is also the most straightforward example of how validation approaches can be applied.

There is a wide range of contexts in which the validation of measures is desirable; however, organizations should, for example, ensure they are using “validated” tools and processes in their performance management systems, in their assessments of training and development outcomes, in their promotion and succession planning processes, etc.

Each of these circumstances is associated with its own set of challenges as the researcher designs an appropriate validation study. However, the design of the well-constructed study by necessity will follow the same logic as will be discussed for the personnel selection context. Following this logic, the studies should be structured to include the following three elements:

1. **Job analysis**: The foundation of validation in employment settings always involves the development of a clear understanding of job and organizational requirements. For example, for promotion purposes these would be the requirements of the target job(s) into which a person might be promoted. For training and development purposes, these would be the meaningful outcomes in terms of on-the-job performance that are the focus of the training/development efforts.

2. **Systematic development**: As measures are developed, they need to follow an architecture that is firmly grounded in the results of the job analysis. As the development of the measures is planned and as the tools are being constructed, activities need to be focused on ensuring that the measures are carefully targeted to address the intended constructs.

3. **Independent verification**: Once the measures are developed, they need to be subjected to independent verification that they measure the intended constructs. At times, this can involve statistical studies to determine whether the measures exhibit expected relationships with other independent measures (e.g., Does the 360-degree assessment of leadership behavior correlate with an independent interview panel’s judgment of a leader’s...
behavior?). Often, the independent verification is derived from structured expert reviews of the measures that are conducted prior to implementation. Regardless of the method, this “independent verification” is a necessary aspect of verifying the validity of a measure.

**STRONG VERSUS WEAK INFERENCES ABOUT VALIDITY**

The field’s evolving conceptualization of validity has important implications for I-O researchers concerned with designing and conducting primary validation studies. Given that validation is a process of collecting evidence to support inferences derived from test scores (e.g., that a person will perform effectively on a job), the confidence with which inferences are made is a function of the strength of the evidence collected. In this way, the strength of the validity evidence refers to the probability that the inference is correct, with “stronger” evidence connoting higher probabilities. Consistent with the unitarian view, validation can be viewed as a form of hypothesis testing (Landy, 1986; Messick, 1975), and judgments of validity are to be based on the same host of considerations applicable to judgments concerning the veracity with which a null hypothesis would be rejected in any psychological research (e.g., the extent to which threats to validity are ruled out; see Cook & Campbell, 1979). Thus, it is critical for researchers designing and conducting local validation studies to concentrate their efforts on ensuring that studies result in strong evidence for the inferences they wish to make in much the same way that they would otherwise “defend” their conclusions in a hypothesis testing situation.

Gathering stronger evidence of validity almost always necessitates increased effort, resources, and/or costs (e.g., to gain larger sample sizes or expand the breadth of the criterion measures). Thus, a key decision for researchers designing primary validation studies involves determining how to optimize the strength of the study (assurance that inferences are correct) within the bounds of certain practical limitations and organizational realities. Situations may vary in terms of the extent to which feasibility drives the researcher’s choice among validation strategies. In some cases, situational limitations may be the primary determinant of the validation strategies available to researchers. For example, for situations in which adequately powered sample sizes cannot be achieved, validation efforts may require use of synthetic validity strategies (Scherbaum, 2005), transporting validity evidence from another context that is judged to be sufficiently similar (Gibson & Caplinger, 2007), generalizing validity across jobs or job families on the basis of meta-analytic findings (McDaniel, 2007; Rothstein, 1992), or relying on evidence and judgments that the content of the selection procedures is sufficiently similar to job tasks to support their use in decision-making. Other factors noted by the Principles that may serve to limit the feasibility of certain validation strategies include unavailability of criterion data, inaccessibility to subject matter experts (SMEs) as might be the case when consulting SMEs would compromise test security, dynamic working conditions such that the target job is changing or does not yet exist, and time and/or money. Clearly then, validation strategy needs to account for feasibility-driven considerations and researchers’ judgment about the strength of evidence required. Further, because these demands are often competing, researchers are frequently required to make the best of a less than optimal situation.

Given the need to balance several competing demands (e.g., issues of feasibility limiting the approach that can be taken vs. upholding high standards of professionalism and providing strong evidence to support key inferences), it is essential that researchers understand the various factors that have potentially important implications for the strength of evidence that is required in a given validation scenario. In other words, part of the decision process, with regard to planning and implementing validation strategy, is a consideration of how strong the evidence in support of key inferences ought to be. The basic assumption here is that different situations warrant different strategies along several dimensions (Sussman & Robertson, 1986), one of which has to do with the strength of evidence needed in support of inferences. Consideration of how certain situational factors inform the adequacy of a validation effort does not imply that the validation researcher adopt a minimalist

Validation Strategies for Primary Studies

approach. Rather, all validation studies and selection practices should aspire to the ethical and professional guidelines offered in the Principles, which means using sound methods rooted in scientific evidence and exhibiting high standards of quality. However, the Principles’ guidelines are not formulaic to the exclusion of professional judgment, nor are their applications invariant across circumstances. In the following paragraphs, several factors are identified that have potential implications for the strength of the evidence needed by a local validation effort. In general, these factors represent situations in which conclusions regarding the validity of a selection practice need to be made with a higher degree of confidence than usual. In turn, these situations tend to warrant intensified research strategies.

SITUATIONAL FACTORS INFLUENCING THE STRENGTH OF EVIDENCE NEEDED

Although it is beyond the scope of this chapter to describe in full detail the legal issues surrounding validation research and selection practice (see Chapters 29 and 30, this volume, for further discussions of legal issues), it would be difficult if not impossible to describe applied validation strategy without underscoring the influence of litigation or the prospect of litigation. It is becoming almost cliché to state that, in circumstances in which there is a relatively high probability of litigation regarding selection practices, validation evidence is likely to function as a central part of defending selection practices in the courtroom. Indeed, much validation research is stimulated by litigation, whether post facto or in anticipation of lawsuits. Within this context, researchers make judgments regarding the potential for litigation and adapt their validation strategies accordingly. Numerous contextual factors contribute to the probability that litigation will occur. A primary example has to do with the type of selection instrument being validated and the potential for adverse impact, or the disproportionate rejection of identifiable subgroups. Tests that have historically resulted in adverse impact, such as tests of cognitive ability (Schmitt, Rogers, Chan, Sheppard, & Jennings, 1997) or physical ability (Arvey, Nutting, & Landon, 1992; Hogan & Quigley, 1986) tend to promote more litigation, and researchers validating these instruments in a local context should anticipate this possibility. Similarly, selection instruments with low face validity (i.e., the test’s job relevance is not easily discerned by test-takers) are more likely to engender negative applicant reactions (Shotland, Alliger, & Sales, 1998), and decisions based on such tests may lead to applicant perceptions of unfairness (Cropanzano & Wright, 2003). In their recent review of the antecedents and consequences of employment discrimination, Goldman, Gutek, Stein, & Lewis (2006) identified employee perceptions of organizational and procedural justice as important antecedents of discrimination lawsuits. In addition to considering characteristics of the selection instrument(s) being validated, lawsuits over selection practice are more frequent in some industry (e.g., government) and job types (Terpstra & Kethley, 2002).

Researchers should also consider the implications and relative seriousness of selection systems resulting in hiring decisions that are false positives or false-negative errors. A false positive is made by selecting an unqualified individual whose performance on the job will be low, whereas a false-negative error is made by rejecting a qualified individual whose performance on the job would have been high. Making an error of either type can be potentially costly to the organization. However, the relative impact of such errors can differ by occupation type and organizational context. For example, the negative impact of a false positive in high-risk occupations (e.g., nuclear power plant operator or air-traffic controller) or high visibility occupations (e.g., Director of the Federal Emergency Management Agency [FEMA]) can be catastrophic, threaten the organization’s existence, and so on (Picano, Williams, & Roland, 2006). Alternatively, for occupations that are associated with less risk, such that failure on the job does not have catastrophic consequences for the organization or larger society, or when organizations use probationary programs or other trial periods, the cost of false-positive errors may be relatively low. Although validation efforts in both situations would be concerned with selection errors and demonstrating that use of tests can reduce the occurrence and negative consequences of such errors, clearly there are some situations in which this would be more
of a central focus of the validation effort. It is our contention that validating selection systems for high-risk occupations are a special circumstance warranting particularly “watertight” validation strategies in which strong evidence should be sought to support the inferences made. In these circumstances, a test with low validity (e.g., less than $r = .10$) might be used to make hiring decisions if that relationship is found to be statistically significant.

In some circumstances, the cost of false negatives is more salient. For example, strong evidence of a test’s validity may be warranted when an organization needs to fill a position or several positions, but applicants’ test scores are below some acceptable standard, indicating that they are not fit to hire (i.e., predicted on-the-job performance is low or very low). In this case, the organization’s decision to reject an applicant on the basis of their test scores would leave a position or several positions within the organization vacant, a costly mistake in the event that false-negative errors are present. Demonstrable evidence to support the test’s validity would be needed to justify such a decision, and in essence, convince the organization that they are better off with a vacant position than putting the wrong person in the job. In these instances, one might want evidence of a larger test-criterion relationship (perhaps greater than $r = .30$) to warrant use of this test and the possible rejection of competent applicants.

The possibility of false negatives becomes a special concern when members of some subgroup(s) are selected less frequently than members of another subgroup. When unequal ratios of various subgroups are selected, the organization must be prepared to show that false negatives are not primarily of one group as opposed to another. When this is impossible, the legal and social costs can be very high. Concern about these costs is another reason to be concerned about the false negatives aside from the concerns associated with possible vacant positions.

Personnel psychologists have long been aware of the fact that the utility of selection systems increase as a function of selectivity, such that selection instruments even modestly related to important outcomes can have large payoffs when there are many applicants from which only a few are to be selected (Brogden, 1951, 1959). On the other hand, as selection ratios become extremely liberal, such that nearly all applicants are accepted, even selection instruments highly related to performance have less positive implications for utility. From a purely utilitarian perspective, it would seem logical that demonstrating test validity is less of an impetus when selection ratios are liberal (because even the best tests will have little effect) and more of an impetus when selection ratios are low. The consequences of large-scale applicant rejections would also seem to justify more rigorous validation methods from societal and legal perspectives. However, these extremes likely occur less frequently in practice, as indicated by Schmidt and Hunter (1998), who cite typical selection ratios that are moderate, ranging on average from .30 to .70.

In licensing examinations, this utility perspective takes a different form because the major purpose of these examinations is to protect the public from “injuries” related to incompetent practice. In this case, the license-no license decision point using test scores is usually set at a point that is judged to indicate “minimal competence.” Depending on the service provided (e.g., hairdresser vs. surgeon), the cost of inappropriately licensing a person could be very different. On the other hand, certification examinations are usually oriented toward the identification of some special expertise in an area (e.g., pediatric dentistry or forensic photography), hence a decision as to a score that would warrant certification might result in the rejection of larger numbers or proportions of examinees. The cost-benefit balance in this situation (assuming all are minimally competent) might accrue mostly to the individual receiving the certification in the form of greater earning power.

Another factor that can affect the extent of the local validation effort that is required is the availability of existing validation research. The Principles describes three related validation strategies that can be used as alternatives to conducting traditional local validation studies or to support the conclusions drawn at the primary study level. First, “transportability” of validity evidence involves applying validity evidence from one selection scenario to another, on the basis that the two contexts are judged to be sufficiently similar. Specifically, the Principles note that researchers should be concerned with assessing similarity in terms of job characteristics [e.g., the knowledge, skills,
Validation Strategies for Primary Studies

and abilities (or KSAs) needed to perform the job in each context, job tasks and content, applicant pool characteristics, or other factors that would limit generalizability across the two contexts (e.g., cultural differences). Assessing similarity in this manner usually requires that researchers conduct a job analysis or rely on existing job analysis materials combined with their own professional expertise and sound judgment.

Second, synthetic validity is a process in which validity for a test battery is “synthesized” from evidence of multiple predictor-job component relationships (Peterson, Wise, Arabian, & Hoffman, 2001; Scherbaum, 2005). Job analysis is used to understand the various components that comprise a particular job, and then predictor-job-component relationships are collected for all available jobs with shared components. Because evidence can be drawn from other jobs besides the focal job, synthetic validity may be a particularly useful strategy for organizations that have too few incumbents performing the focal job to reach adequate sample sizes for a traditional criterion-related study (Scherbaum, 2005).

Third, validity generalization involves using meta-analytic findings to support the conclusion that predictor-criterion validity evidence can be generalized across situations. Like transportability strategies, meta-analytic findings provide researchers with outside evidence to support inferences in a local context. The argument for validity generalization on the basis of meta-analyses is that some selection tests, such as cognitive ability tests (Ones, Viswesvaran, & Dilchert, 2005), are valid across selection contexts. Thus, the implication is that with validity generalization strategies, unlike transportability, in-depth job analyses or qualitative studies of the local organizational context are unnecessary. In support of this assertion, the program of research initiated by Schmidt and Hunter and colleagues (for review, see Schmidt & Hunter, 2003) has argued that between-study variability in validity coefficients can be largely attributed to statistical artifacts, such as range restriction, unreliability, or sampling error. However, caution is warranted to the extent that meta-analyses have identified substantive moderators, or in the presence of strong theory indicating that some variable may moderate the magnitude of validity. Further, with regard to generalization across contexts, inferences drawn from meta-analytic findings are limited to the contexts of those studies included in the meta-analysis. At minimum, meta-analytic findings should be referenced in test development and can be used to supplement evidence at the local level, either via theoretical or statistical means (Newman, Jacobs, & Bartram, 2007). The argument for more direct use of validity generalization strategies is dependent on the strength of the meta-analytic findings.

DESIGN CONSIDERATIONS WHEN STRONG EVIDENCE IS NEEDED

On the basis of the preceding discussion, situational factors can affect the feasibility and appropriateness of the validation models applied to a given selection context. Moreover, researchers should be particularly attuned to contextual variables that warrant an increased concern for demonstrating the strength of evidence collected and high levels of confidence in the inferences to be made. The validity strategies used reflect consideration of these contextual factors and others. For instance, in apparent response to increasing litigation, Boehm (1982) found longitudinal trends suggesting that published validation studies were using larger sample sizes, relying on supervisory rating criteria less frequently, and utilizing more predictive (as opposed to concurrent) criterion-related designs. Similarly, the discussion that follows is focused on identifying a handful of actionable validation strategies to be considered by researchers when particularly strong evidence is needed.

Importance of the Nomological Net

Binning and Barrett (1989) offered a thorough conceptualization of the nomological network implicit in validity models. Their model identifies multiple inferential pathways interrelating psychological constructs and their respective operational measures. Inferential pathways in the model are empirically testable using observed variables (e.g., linkages between operationalized measures of constructs and linkages between constructs and their operationalized measures).
Others may be theoretically or rationally justified (e.g., construct to construct linkages) or tested using latent variable models, although these applications are relatively rare in personnel selection research (see Campbell, McHenry, & Wise, 1990, for an attempt to model job performance). Consistent with the unitarian conceptualization of validity, all validity efforts in a selection context are ultimately concerned with demonstrating that test scores predict future job performance, and each of the various inferential pathways represents sources or types of evidence to support this common inference. Binning and Barrett (1989, p. 482) described how “truncated” validation strategies often concentrate exclusively on demonstrating evidence for a single inferential pathway and as a result provide only partial support for conclusions regarding test validity. A more cogent argument for validity is built upon demonstration of strong evidence for several inferential pathways within the nomological network. For example, in addition to demonstrating a statistical relationship between operational measures from the predictor and performance domain, as is commonly the main objective in criterion-related validity studies, researchers should provide evidence of the psychological constructs underlying job performance (as well as the predictor measures) and demonstrate that the criterion measure adequately samples constructs from the performance domain.

**Criterion Concerns**

In practice, criterion-related validity studies are often criticized for failing to adequately address validity issues surrounding the criterion measure(s) used. The relative lack of scientific scrutiny focused on criteria, termed the “criterion problem” (Austin & Villanova, 1992), has been a topic of discussion among personnel psychologists for years (Dunnette, 1963; Fiske, 1951; Guion, 1961). Universal to these discussions is the call for more rigorous validation evidence with respect to the criteria that are used. Binning and Barrett (1989) outlined this task, underscoring two interrelated goals for the validation researcher. First, they suggested that the selection of criteria should be rooted in job analysis to the same extent that selection of predictors traditionally are (i.e., more attention to rigorous “criterion development”). Other considerations relevant to the task of criterion development and validation include the use of “hard” or objective criteria versus more proximal behaviors that lead to these outcomes (Thayer, 1992), use of multiple relevant criteria as opposed to a single overall criterion (Dunnette, 1963), and the possibility that criteria are dynamic (Barrett, Caldwell, & Alexander, 1985). Second, researchers should be concerned with demonstrating evidence of construct-related validity for the criterion. Investigators must specify the latent dimensions that underlay the content of their criterion measures. This involves expansion of the nomological network to include inferences that link the criterion measure(s) to constructs in the performance domain (e.g., by demonstrating that criterion measures are neither contaminated nor deficient with respect to their coverage of the intended constructs in the performance domain) and link constructs in the performance domain to job demands that require specific ability or motivational constructs (e.g., by demonstrating through job analysis that constructs in the performance domain are organizationally meaningful). Campbell and his colleagues (e.g., Campbell, McCloy, Oppler, & Sager, 1993) have repeatedly emphasized this emphasis on the nature of criteria or performance constructs. These authors make the somewhat obvious, although often overlooked, point that performance should be defined as behavior (“what people actually do and can be observed”); the products of one’s behavior, or what are often called “hard criteria,” are only indirectly the result of one’s behavior and other factors that are not attributable to an individual job incumbent. Further, we may consider relatively short term or proximal criteria or distal criteria, such as the impact of one’s career on some field of interest. Any specification of a performance or criterion domain must also consider the impact of time (Ackerman, 1989; Henry & Hulin, 1989). In any study of performance, these various factors must be carefully considered when one decides on the nature of the performance constructs and actual operationalizations of the underlying constructs and how those measures might or might not be related to measures of other constructs in the domain of interest.

Multiple Inferences in Validation Research

Gathering evidence to support multiple inferences within a theoretically specified nomological network resembles a pattern-matching approach. The advantage of pattern-matching research strategies is that stronger support for a theory can be gained when complex patterns of observed results match those that are theoretically expected (Davis, 1989). Logically, it would be less likely that a complex pattern of results would be observed simply because of chance. In addition, when experimental control of potentially confounding variables is not possible, pattern matching can be used to preempt alternative explanations for the observed relationships (i.e., threats to validity; Cook & Campbell, 1979).

Campbell and Fiske (1959) described a specific application of pattern matching that can be used to support inferences regarding the appropriateness with which a construct is operationalized in measurement. A slight variation of the convergence/divergence idea is included in Binning and Barrett’s (1989) elaborated nomological network model. The elaborated model includes inferences regarding the proximal versus distal nature of relationships between constructs. A sound argument for validity can be made on the basis of results that indicate a reliable pattern, in which strong statistical relationships are obtained for constructs that are theoretically proximal to one another and weaker statistical relationships are obtained for constructs that are theoretically more distal. This, of course, is the rationale underlying mediation (Baron & Kenny, 1986) and the idea of causal chains. Even without testing mediation directly, a strong argument can be made by demonstrating that a predictor correlates highly with proximal operational measures of the criterion construct, and that this relationship is attenuated as criterion measures become more distally related to the construct space. For example, Thayer (1992) noted that in many cases objective outcome criteria, although important at the organizational level, are distal indicators of the job performance construct because they are contaminated by factors outside of the employee’s control. Also, low reliability was found for the theoretically removed, “hard” criteria measures discussed by Thayer. In contrast, criteria that assess employee behavior directly as defined above are more proximal to the job performance construct. To the extent that these proximal criteria can be identified and measured, stronger support for validity is gained in the event that test scores correlate more highly with proximal as compared to distal criteria.

A more extensive form of pattern matching involves the use of multiple studies, or research programs, to corroborate evidence of validity. Again, the logic is straightforward; stronger evidence is gained when a constellation of findings all lead to the same conclusion. Sussman and Robertson (1986) suggested that programs of research could be undertaken, “composed of multiple studies each utilizing a different design and aimed at collecting different types of evidence” (p. 467). Extending the rationale of the multi-trait multi-method (MTMM) (Campbell & Fiske, 1959), convergent evidence across studies may indeed be stronger if gained through different research designs and methods. Landy’s (1986) assertion that test validation is a form of hypothesis testing, and that judgments of validity are to be based on a “preponderance of evidence” (p. 1191; Guion, as cited in Landy, 1986), provides the context for consideration of research strategies such as quasi-experimental designs (Cook & Campbell, 1979) and program evaluation research (Strickland, 1979). Binning and Barrett (1989) presented a similar rationale by calling for “experimenting organizations” (p. 490) in which local validation research is treated as an ongoing and iterative process. Published research on use of these research methods in a selection-validation context remains sparse to date.

SPECIAL CONSIDERATIONS IN CRITERION-RELATED STUDIES

In the event that a criterion-related strategy is part of the validation effort undertaken, a special set of considerations is relevant. Power analysis is a useful framework for interrelating the concepts of statistical significance, effect size, sample size, and reliability (Cohen, 1988) and has design and evaluation implications for the statistical relationships sought in criterion-related studies. For
instance, the sample size needed to demonstrate a statistically significant predictor-criterion relationship decreases as the magnitude of the relationship that exists between predictor and criterion (i.e., effect size) increases. Sussman and Robertson (1986), in their assessment of various predictive and concurrent validation designs, found that those strategies that allowed larger sample sizes gained a trivial increment in power. This suggests that, as long as sample sizes can support the use of a criterion-related design, further attention toward increasing \( N \) may not reap large benefits. Other factors affecting power include the interrelatedness and number of predictors used, such that the addition of nonredundant predictors increases power (Cascio, Valenzi, & Silbey, 1978). The reliability of the predictors and criteria and the decision criteria used for inferring that a relationship is nonzero also impact power.

By incorporating power analysis in validation design, researchers can increase the likelihood that relationships relevant to key inferences will be tested with sufficient sample size upon which to have confidence in the results. However, from a scientific standpoint, the importance of demonstrating that predictor-criterion relationships are statistically significant may be overstated, given that relationships, which may not be practically meaningful, can reach statistical significance with large enough sample sizes. For instance, a statistically significant relationship, in which a test accounts for less than 5% of the variance in job performance, is not unequivocal support for the test’s use. This is especially evident when there is reason to suggest that other available tests could do a better job predicting performance. Nonetheless, we agree with Cortina and Dunlap’s (1997) overall contention that statistical significance testing remains a useful tool to researchers when decision criteria are needed. For this reason, strong evidence in support of the predictor-criterion relationship should be derived based on the size and significance of the validity coefficients.

Operationally, there are several other important considerations in criterion-related research (e.g., job analyses that support the relevance of predictor and criterion constructs and the quality of the measures of each set of constructs). However, those concerns are addressed repeatedly in textbooks (e.g., Guion, 1998; Ployhart, Schneider, & Schmitt, 2006). In the next section of this chapter, we address a very important concern that is rarely discussed.

**CONCERNS ABOUT THE QUALITY OF THE DATA: CLEANING THE DATA**

Once data have been collected, quality control techniques should be applied to ensure that the data are clean before proceeding to statistical analysis. Some basic quality control techniques include double-checking data for entry errors, spot checks for discrepancies between the electronic data and original data forms, inspecting data for out-of-range values and statistical outliers, and visual examination of the data using graphical interfaces (e.g., scatter plots, histograms, stem-and-leaf plots). Special concern is warranted in scenarios with multiple persons accessing and entering data or when data sets from multiple researchers are to be merged. Although these recommendations may appear trite, they are often overlooked and the consequence of erroneous data can be profound for the results of analyses and their interpretations.

A study by Maier (1988) illustrated, in stepwise fashion, the effects of data cleaning procedures on validity coefficients. Three stages of data cleaning were conducted, and the effects on correlations between the Armed Services Vocational Aptitude Battery (ASVAB) and subsequent performance on a work sample test for two military jobs (radio repairers and automotive mechanics) were observed. Selection was based on the experimental instrument (the ASVAB) and the work sample criterion tests were administered to incumbents in both occupations after some time had passed. In Phase 1 of the data cleaning process, the sample was made more homogenous for the radio repairers group by removing the data of some employees who received different or incomplete training before criterion data collection. In comparison to the total sample, the validity coefficient for the remaining, more representative group that had received complete training before criterion collection was decreased (from .28 to .09). The initial estimate had been inflated because of the partially trained group having scored low on the predictor and criterion.
In Phase 2, scores on the criterion measure (i.e., ratings from a single rater on a work sample) were standardized across raters. There were significant differences between raters that were attributed to different rating standards and not to group differences in ratees, such as experience, rank, or supervisor performance ratings. The raters were noncommissioned officers and did not receive extensive training in the rating task, so that differences between raters in judgmental standards were not unexpected. As a result, the validity coefficients for both jobs increased (radio repairers, from .09 to .18; automotive mechanics, from .17 to .24). In Phase 3, validity coefficients were corrected for range restriction, which again resulted in an increase in the observed validity coefficients (radio repairers, from .18 to .49; automotive mechanics, from .24 to .37). Maier noted that the final validity coefficients were within the expected range on the basis of previous studies.

The Maier (1988) study is illustrative of the large effect that data cleaning can have for attaining more accurate estimates of validity coefficients in a predictive design scenario. Several caveats are also evident, so that researchers can ensure that data cleaning procedures conducted on sound professional judgment are not perceived as data “fudging.” First, the cleaning procedures need to have a theoretical or rational basis. Researchers should document any decision criteria used and the substantive changes that are made. For example, researchers should record methods used for detecting and dealing with outliers. In addition, a strong case should be built in support of any decisions made. The researcher bears the burden of defending each alteration made to the data. For example, in the Maier study, the decision to standardize criterion data across raters (because raters were relatively untrained and used different rating standards) was supported by empirical evidence that ruled out several alternative explanations for the mean differences observed between raters.

**MODES OF DECISION-MAKING AND THE IMPACT ON UTILITY AND ADVERSE IMPACT**

If we have good quality data, it still matters how we use those data in making decisions as to whether or not use of the test produces aggregated performance improvements. In this section, we will discuss the impact of various modes of decision-making on two outcomes that are of concern in most organizations: Overall performance improvement or utility and adverse impact on some protected group defined as unequal proportions of selection across subgroups. Advancing both outcomes is often in conflict, especially when one uses cognitive ability tests to evaluate the ability of members of different racial groups or physical ability when evaluating male and female applicants for a position. Measures of some other constructs (e.g., mechanical ability) produce gender or race effects, but the subgroup differences that are largest and affect the most people are those associated with cognitive and physical ability constructs.

**TOP-DOWN SELECTION USING TEST SCORES**

If a test has a demonstrable relationship to performance on a job, it is the case that the optimal utility in terms of expected employee performance will occur when the organization selects the top scoring persons on the test to fill its positions (Brown & Ghiselli, 1953). Expected performance is a direct linear function of the test score-performance relationship in the situation in which the top scoring individuals are selected. However, use of tests in this fashion when it is possible will mean that lower scoring subgroups will be less likely to be selected (Murphy, 1986). This conflict between maximization of expected organizational productivity and adverse impact is well known and has been quantified for different levels of subgroup mean differences in ability and selection ratios (Sackett, Schmitt, Ellingson, & Kabin, 2001; Sackett & Wilk, 1994; Schmidt, Mack, & Hunter, 1984). For social, legal, and political reasons as well as long-term organizational viability in some contexts, the adverse impact of a strict top-down strategy of test use often cannot be tolerated. For these reasons as well as others, researchers and practitioners have often experimented with and used other ways of using test scores.

BANDING AND CUT SCORES

One method of reducing the consequences of subgroup differences in test scores and top-down selection is to form bands of test scores that are not considered different usually using a statistical criterion known as the standard error of the difference, which is based on the reliability of the test. Most of us are familiar with a form of banding commonly used in academic situations. Scores on tests are usually grouped into grades (e.g., A, B, C, etc.) that are reported without specific test score information. So persons with scores of 99 and 93 might both receive an A in a course just as two with scores of 88 and 85 would receive a B. The theory in employment selection use of banding is that the unreliability inherent in most tests makes the people within a band indistinguishable from each other just as occurs when grades are assigned to students.

Because minorities tend to score lower on cognitive ability tests, creating these bands of indistinguishable scores helps increase the chances that minority applicants will fall in a top band and be hired. There are two ways in which banding can increase minority hiring. One is to make the bands very wide so that a greater number of minority test scorers will be included in the top bands. Of course, a cynic may correctly point out that a test of zero reliability will include everyone in the top band and that this approach supports the use of tests with low reliability. A second way in which to impact the selection of minority individuals is the manner in which individuals are chosen within a band. The best way to increase the selection of minority individuals is to choose these persons first within each band before proceeding to consider other individuals in the band, but this has proven difficult to legally justify in U.S. courts (Campion et al., 2001). Other approaches to selection within a band include random selection or selection on secondary criteria unrelated to subgroup status, but these procedures typically do not affect minority hiring rates in practically significant ways (Sackett & Roth, 1991). A discussion of various issues and debates regarding the appropriateness of banding is contained in an edited volume by Aguinis (2004).

An extreme departure from top-down selection occurs when an organization sets a minimum cutoff test score such that individuals above some score are selected whereas those below that score are rejected. In essence, there are two bands of test scores—those judged to represent a passable level of competence and those representing a failing level of performance. Perhaps the most common use of cutoff scores is in licensing and credentialing, in which the effort is usually to identify a level of expertise and knowledge of the practice of a profession below which a licensure candidate is likely to bring harm to clients. In organizational settings a cutoff is often required when selection of personnel is done sequentially over time rather than from among a large number of candidates at a single point in time. In this case, hire-reject decisions are made about individuals and a pass score is essential.

Use of a single cutoff score will certainly reduce the potential utility inherent in a valid test because it ignores the individual differences in ability above the test score cutoff. There exists a great deal of evidence (e.g., Coward & Sackett, 1990) that test score-job performance relationships are linear throughout the range of test scores. However, using a minimum cut score on a cognitive ability test on which we usually see the largest minority-majority differences to select employees and selecting above that cutoff on a random basis or on the basis of some other valid procedure that does not display subgroup differences may very much reduce the adverse impact that usually occurs with top-down selection using a cognitive ability test.

Perhaps the biggest problem with the use of cutoff scores is deriving a justifiable cut score. Setting a cutoff is always judgmental. Livingston (1980) and Cascio, Alexander, and Barrett (1988) among others have usually specified the following as important considerations in setting cutoffs: the qualifications of the experts who set the cutoff, the purpose for which the test is being used, and the consideration of the various types of decision errors that can be made (i.e., denying a qualified person and accepting an unqualified individual). One frequently used approach is the so-called Angoff method, in which a representative sample of experts examines each test item and determines the probability that a minimally competent person (the definition and experts’ understanding
Validation Strategies for Primary Studies

of minimally competent is critical) would answer the question correctly. These probabilities are summed across experts and across items. The result is the cutoff score. A second approach to the setting of cut scores is to set them by reference to some acceptable level of performance on a criterion variable. In this case, one could end up saying that an individual with a score of 75 on some test has a 10% (or any percent) chance of achieving success on some job. However, this “benchmarking” of scores against criteria does not resolve the problem because someone will be asked to make sometimes equally difficult decisions about what constitutes acceptable performance. Cizek (2001) provided a comprehensive treatment of methods of setting performance standards.

The use of score cutoffs to establish minimum qualifications or competency is common in licensing exams. Credentialing exams may require evidence of a higher level of skill or performance capability in some domain, but they too usually require only a “pass-fail” decision. Validation of these cutoffs almost always relies solely on the judgments of experts in the performance area of interest. In these cases, careful explication of the behaviors required to perform a set of tasks and the level of “acceptable” performance is essential and likely the only possible form of validation.

USING PROFILES OF SCORES

Another possibility when scores on multiple measures of different constructs are available is that a profile of measured abilities is constructed and that this profile is matched to a profile of the abilities thought to be required in a job. In this instance, we might measure and quantify the type of job experiences possessed by a job candidate along with their scores on various personality tests, and their oral communications and social skills as measured in an interview and scores on ability tests. If this profile of scores matches that required in the job, the person would be selected. This contrasts with the traditional approach described in textbooks in which the person’s scores on these tests would be linearly related to performance and combined using a regression model so that each score was optimally linearly related to job performance. In using profiles, one is interested in patterns of scores rather than an optimally weighted composite. Use of profiles of scores presents various complex measurement and statistical problems of which the user should be aware (Edwards, 2002). Instances in which selection decisions are made in this fashion include individual assessments (Jeanneret & Silzer, 1998), which involve the use of multiple techniques using multiple methods of assessment and a clinical judgment by the assessor that a person is qualified for some position (Ryan & Sackett, 1987; 1992; 1998). Another venue in which profiles of test scores are considered is in assessment centers in which candidates for positions (usually managerial) are evaluated in various exercises on different constructs and assessors make overall judgments that are then used in decision-making. Overall judgments based on these procedures have shown criterion-related validity [see Ryan & Sackett (1998) for a summary of data relevant to individual assessment and Gaugler, Rosenthal, Thornton, & Bentson (1987) or Arthur, Day, McNelly, & Edens (2003) on assessment center validity], but we are aware of no evidence that validates a profile or configural use of scores.

Perhaps the best description of the research results on the use of profiles to make high-stakes decisions is that we know very little. The following would be some of the issues that should receive research attention: (a) Is a profile of scores actually used, implicitly or explicitly, in combining information about job applicants and what is it? (b) What is the validity of such use and its incremental validity over the use of individual components of the profile or linear composites of the scores in the profile? and (c) What is the adverse impact on various subgroups using profile judgments?

CLINICAL VERSUS STATISTICAL JUDGMENT

Clinical judgment refers to the use and combination of different types of information to make a decision or recommendation about some person. In psychology, clinical judgment may be most often discussed in terms of diagnoses regarding clinical patients (Meehl, 1954). These judgments
are likely quite similar to those made in the individual assessments discussed in the previous section of this chapter but also may occur when judgments are made about job applicants in employment interviews, assessment centers, and various other instances in which human resource specialists or psychologists make employment decisions. Clinical judgment is often compared with statistical judgment in which test scores are combined on the basis of an arithmetic formula that reflects the desired weighting of each element of information. The weights may be determined rationally by a group of job experts or by using weights derived from a regression of a measure of overall job success on scores on various dimensions using different methods of measurement. Meehl's original research (1954) showed that the accuracy of the worst regression estimate was equal to the judgments made by human decision-makers. A more recent treatment and review of this literature by Hastie and Dawes (2001) has reaffirmed the general conclusion that predictions made by human experts are inferior to those based on a linear regression model. However, human experts are required to identify the types of information used in the prediction task. The predictions themselves are likely best left to some mechanical combination rule if one is interested in maximizing a performance outcome. The overall clinical judgment when used to make decisions should be the focus of the validation effort, but unless it is clear how information is combined by the decision-maker, it is unclear what constructs are playing a role in their decisions. The fact that these clinical judgments are often not as highly correlated with externally relevant and important outcomes suggests that the constructs these decision-makers use are not relevant.

In clinical judgment, the presence or absence of adverse impact can be the result of a combination of information that does not display sizable subgroup differences or a bias on the part of the person making the judgment. Psychologists making clinical judgments may mentally adjust scores on the basis of their knowledge of subgroup differences on various measures. There are again no studies of which we are aware that address the use or appropriateness of such adjustments.

**SCIENTIFIC OR LONG-TERM PERSPECTIVE: LIMITATIONS OF EXISTING PRIMARY VALIDATION STUDIES, INCLUDING THE CURRENT META-ANALYTIC DATABASE**

There are a great many meta-analyses of the criterion-related validity of various constructs in the prediction of job performance and many thousands of primary studies. Secondary analyses of meta-analyses have also been undertaken (e.g., Schmidt & Hunter, 1998). The studies that provided these data were nearly all conducted more than 30 years ago. Although it is not necessarily the case that the relationships between ability and performance documented in these studies have changed in the last half-century or so, this database has some limitations. In this section, we describe these limitations and make the case that researchers continue their efforts to evaluate test-performance relationships and improve the quality of the data that are collected.

**CONCURRENT VALIDATION DESIGNS**

In criterion-related validation research, concurrent validation studies in which predictor and criterion data are simultaneously collected from job incumbents are distinguished from predictive designs. In the latter, predictor data are collected before hiring from job applicants and criterion data are collected from those hired presumably on the basis of criteria that are uncorrelated with the predictor data after some appropriate period of time when job performance is thought to have stabilized. Defects in the concurrent design (i.e., restriction of range and a different motivational set on the part of incumbents versus applicants) have been described frequently (Barrett, Phillips, & Alexander, 1981). Most comparisons of predictive and concurrent designs indicate that they provide similar estimates of validity. However, it is probably the case that tests more susceptible to motivational differences between job incumbents and applicants, as might be the case for many noncognitive measures which would display differences in validity when the participants in the research
were actually being evaluated for employment versus a situation in which they were responding “for research purposes.” To our knowledge, this comparison has not been made frequently, and when it has been done in meta-analyses cognitive and noncognitive test validities have not been separated (Schmitt, Gooding, Noe, & Kirsch, 1984). It is the case that practical considerations have made the use of concurrent designs much more frequent than predictive designs (Schmitt et al., 1984).

Meta-analytic data suggest that there are not large differences in the validity coefficients resulting from these two designs. Further, range restriction corrections can be applied to correct for the fact that data for lower-scoring persons are absent from concurrent studies, but these data are often absent in reports of criterion-related research. Nor can we estimate any effects on test scores that might result from the fact that much more is at stake in a testing situation that may result in employment as opposed to one that is being done for research purposes. Moreover, as Sussman and Robertson (1986) maintained, the manner in which some predictive studies are designed and conducted make them little different than concurrent studies.

**Unidimensional Criterion Versus Multidimensional Perspectives**

Over the last two decades, the view that job performance is multidimensional has become much more widely accepted by I-O psychologists (Borman & Motowidlo, 1997; Campbell, Gasser, & Oswald, 1996). Early validation researchers often used a single rating of what is now called task performance as a criterion, or they combined a set of ratings into an overall performance measure. In many cases a measure of training success was used as the criterion. The Project A research showed that performance was comprised of clearly identifiable dimensions (Campbell et al., 1990) and subsequent research has very often included the use of measures of contextual and task performance (Motowidlo, 2003). Some researchers also argue that the nature of what constitutes performance has changed because jobs have changed (Ilgen & Pulakos, 1999). In all cases, the underlying performance constructs should be specified as carefully as possible, perhaps particularly so when performance includes contextual dimensions, which, as is true of any developing literature, have included everything that does not include “core” aspects of a job. Validation studies (and meta-analyses) that include this multidimensional view of performance are very likely to yield information that updates earlier validation results.

**Small Sample Sizes**

The limitations of small sample sizes in validity research have become painfully obvious with the development of meta-analyses and validity generalization research (Schmidt & Hunter, 1977) as well as the recognition that the power to reject a null hypothesis that there is no test score-performance relationship is very low in much early validation work (Schmidt, Hunter, & Urry, 1976). Although methods to correct for the variability in observed validity coefficients are available and routinely used in meta-analytic and validity generalization research, the use of small samples does not provide for confidence in the results of that research and can be misleading in the short term as enough small sample studies are conducted and reported to discern generalizable findings. This may not be a problem if we are satisfied that the relationships studied in the past are the only ones in which our field is interested, but it is a problem when we want to evaluate new performance models (e.g., models that include a distinction between task, contextual dimensions, or others), new predictor constructs (e.g., some noncognitive constructs or even spatial or perceptual measures), or when we want to assess cross- or multilevel hypotheses.

**Inadequate Data Reporting**

The impact of some well-known deficiencies in primary validation studies is well known. Corrections for range restriction and criterion unreliability (in the mean and variance of validity coefficients

and for the variability due to small sample size are also well known and routinely applied in validity generalization work. However, most primary studies do not report information that allows for sample-based corrections for criterion unreliability or range restriction. Schmidt and Hunter (1977) in their original meta-analytic effort used estimates of the sample size of the validity coefficients they aggregated because not even sample size was available in early reports. Consequently, in estimating population validity coefficients, meta-analysts have been forced to use assumed artifact distributions based on the small amount of data that are available. There is some evidence that these assumptions are approximately correct (e.g., Alexander, Carson, Alliger, & Cronshaw, 1989; Sackett & Ostgaard, 1994) for range restriction corrections, but the use of such assumed artifact distributions would not be necessary with adequate reporting of primary data. Unfortunately, such information for most of our primary database is lost. In addition, researchers disagree regarding the appropriate operationalization of criterion reliability (Murphy & DeShon, 2000; Schmidt, Viswesvaran, & Ones, 2000).

CONSIDERATION OF MULTILEVEL ISSUES

As described in the section above on the utility and adverse impact associated with selection procedures, selection researchers have made attempts to estimate the organizational outcomes associated with the use of valid tests (Boudreau & Ramstad, 2003). Utility is linearly related to validity minus the cost of recruiting and assessing personnel. When multiplied by the number of people and the standard deviation of performance in dollar terms, the estimates of utility for most selection instruments are very large (e.g., see Schmidt, Hunter, Outerbridge, & Trattner, 1986).

Another body of research has focused on the relationship between organizational human resource practices such as the use of tests and measures of organizational success. The organizational-level research has documented the usefulness of various human resource practices including test use. Terpstra and Rozell (1993) reported correlational data that supported the conclusion that organizations that used various selection procedures such as interviews, cognitive ability tests, and biodata had higher annual levels of profit, growth in profit, and overall performance.

Various other authors have called for multilevel (individuals, work groups, organizations) or cross-level research on the relationship between knowledge, skills, abilities, and other characteristics (KSAOs) and organizational differences (Schneider, Smith, & Sipe, 2000). Ployhart and Schmitt (2007) and Schneider et al. (2000) have proposed a series of multilevel questions that include considerations of the relationships between the variance of KSAOs and measures of group and organizational effectiveness. In the context of the attraction-selection-attrition model (Schneider, 1987), there are many issues of a multilevel and longitudinal nature that researchers are only beginning to address and about which we have very little or no data. These questions should be addressed if we are to fully understand the relationships between KSAOs and individual and organizational performance.

VALIDATION AND LONG-TERM OR SCIENTIFIC PERSPECTIVE

Given the various limitations of our primary database noted in the previous sections of this chapter, we believe selection researchers should aim to conduct additional large-scale or consortium studies like Project A (Campbell, 1990; Campbell & Knapp, 2001). These studies should include the following characteristics:

1. They should be predictive (i.e., longitudinal with data collection at multiple points), concurrent, and of sufficient sample size to allow for adequate power in the tests of hypotheses. Large-scale studies in which organizations continue data collection over time on an ever-expanding group of participants should be initiated.

2. Multiple criteria should be collected to allow for evaluation of various KASO-performance relationships.

3. Data should be collected to allow for artifact corrections such as unreliability in the criteria and range restriction.
4. Unit-level data should be collected to allow for evaluation of multilevel hypotheses. These data should include basic unit characteristics and outcome data.
5. Demographic data should be collected to allow for evaluation of subgroup differences in the level of performance and differences in KASO-performance relationships across subgroups.
6. Data on constructs thought to be related (and unrelated) to the target constructs of interest should be collected to allow for evaluation of broader construct validity issues.

Obviously, these studies would necessitate a level of cooperation and planning not characteristic of multiple researchers, much less multiple organizations. However, real advancement in our understanding of individual differences in KSAOs and performance will probably not come from additional small-scale studies or meta-analyses of primary studies that address traditional questions with sample sizes, research designs, and measurement characteristics that are not adequate.

CONCLUSIONS

It is certainly true that meta-analyses have provided our discipline with strong evidence that many of the relationships between individual differences and performance are relatively strong and generalizable. However, many situations where validation is necessary do not lend themselves to validity generalization or the use of meta-analytic databases. As a result, practitioners frequently find themselves in situations where well-designed primary studies are required. A focus on the appropriate designs for these studies is therefore important.

Additionally, without primary studies of the relationships between individual differences and performance, there can be no meta-analyses, validity transport, or validity generalization. The quality and nature of the original studies that are the source of our meta-analytic database determine to a great extent the currency and quality of the conclusions derived from the meta-analyses, statistical corrections notwithstanding.

We argue that the field would be greatly served by large-scale primary studies of the type conducted as part of Project A (see Sackett, 1990 or Campbell & Knapp, 2001). These studies should begin with a clear articulation of the performance and predictor constructs of interest. They should involve the collection of concurrent and predictive data and improve upon research design and reporting issues that have bedeviled meta-analytic efforts for the past three decades. Demographic data should be collected and reported. Data should be collected across multiple organizational units and organizations (and perhaps globally), and data describing the organizational context should be collected and recorded. We know much more about the complexities of organizational behavior, research design, measurement, and individual differences than we did 80–100 years ago, and this should be reflected in how we collect our data and make them available to other professionals.

REFERENCES


Validation Strategies for Primary Studies


Validation Strategies for Primary Studies


In the United States, the legal context plays a major role in how psychologists approach selection system development. Psychologists know well the set of protected groups, the approaches to making an a priori case of discrimination (e.g., differential treatment vs. adverse impact), the key court cases influencing selection, and the prohibitions against preferential treatment (e.g., the 1991 ban on score adjustment or within-group norming). Selection texts (e.g., Guion, 1998) and human resource management texts (e.g., Cascio & Aguinis, 2008) give prominent treatment to the legal context. In recent years, there has been a growing internationalization of industrial-organizational (I-O) psychology such that psychologists from all over the world work with clients in other countries and contribute to our journals and to our conferences. Test publishers and consulting firms establish offices all over the world. As this internationalization continues to increase, it becomes increasingly useful to take a broader look at the legal environment for selection, examining similarities and differences in various countries. For example consider a U.S firm with operations in several other countries. Although U. S. fair employment law applies only to those overseas employees who are U.S. citizens,

1 All authors contributed equally to this chapter. Paul R. Sackett and Winny Shen integrated the text materials provided by each author. Portions of this chapter were drawn from an article by the same set of authors: Myors, B., Lievens, F., Schollaert, E., Van Hoye, G., Cronshaw, S. F., Mladinic, A., et al. (2008). International perspectives on the legal environment for selection. *Industrial and Organizational Psychology: Perspectives on Science and Practice, 1*, 200–256. Used by permission of the Society for Industrial and Organizational Psychology and Wiley Blackwell.

the employment by U.S. firms of host country nationals or third country nationals is subject to the legal environment of the host country.

**DATA COLLECTION METHODOLOGY**

To compare and contrast the legal environment for selection in various countries, the senior author prepared a set of questions about the legal environment for selection, prepared model answers describing the legal environment in the United States, and contacted psychologists in various countries, asking them to prepare a document responding to each question and describing the legal environment in their country. They were also invited to suggest additional project participants in other countries. Some invitees declined; some initially agreed, but subsequently did not participate. The goal was to obtain a range of perspectives by sampling about 20 countries, thus, this is by no means a complete catalog of the legal environment around the world. Researchers and practitioners who are experts on the topic of selection participated from the following 22 countries: Australia, Belgium, Canada, Chile, France, Germany, Greece, India, Israel, Italy, Japan, Kenya, Korea, The Netherlands, New Zealand, South Africa, Spain, Switzerland, Taiwan, Turkey, the United Kingdom, and the United States. As the list indicates, the countries covered do broadly sample the world. Because of space constraints, the write-up for each country was subsequently summarized and organized by issue (e.g., what groups are protected; is preferential treatment of minority group members permitted) rather than by country to create this chapter. For more context on the legal, social, and political environment of the countries surveyed, see Myors et al. (2008). Contributing authors from each of the 22 countries responded to several questions, nine of which are addressed in turn in this chapter.

**Question 1: Are there racial/ethnic/religious subgroups such that some are viewed as “advantaged” and others as “disadvantaged”?**

Table 30.1 identifies the major groups viewed as “disadvantaged” in each country (note that gender is treated separately in the next section, and specific legal protections for disadvantaged groups are treated under Question 4). As Table 30.1 indicates, the disadvantaged groups differ on several dimensions. First, the basis for disadvantaged status varies: (a) native/aboriginal people in a setting where colonizers became the majority group (e.g., Native Americans in the United States, Maori in New Zealand, First Nations Peoples in Canada), (b) recent immigrants (e.g., many European countries), (c) racial groups either native to or with long histories in the country (e.g., African Americans in the United States; Blacks, colored individuals, and Indians in South Africa), (d) religious groups (e.g., India), and (e) language groups (e.g., Francophones in Canada, Rhaeto-Romanic speakers in Switzerland). Second, the size of the minority population varies, from a very small percentage of the population in some countries to the South African extreme of a previously disadvantaged Black majority. These findings illustrate that there is considerable variability from country to country in what constitutes a disadvantaged group.

**Question 2: What is the general picture regarding women in the workplace (e.g., historical trends regarding employment for women; current data on percentage of women in the workforce; and current status regarding occupational segregation, such as gender representation in various job classes and at various organizational levels)?**

Among the countries surveyed, women make up a substantial portion of the workforce. In general, women make up from over one quarter to slightly less than one half of the working population (see Table 30.2). Great strides have been made such that women are being increasingly involved in the workforce across all countries surveyed, as evidenced by reports of the increased rate of women’s participation in the workforce, with the exception of Turkey, who reports a slight decline in the recent years (34% in the early 1990s down to 25.4% in 2004; State Institute of Statistics, 2006). There is substantial variability among countries in terms of the percentage of women who
TABLE 30.1
Disadvantaged Groups Within Each Country

<table>
<thead>
<tr>
<th>Country</th>
<th>Group</th>
<th>Percentage of Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>Indigenous Australians</td>
<td>2.5</td>
</tr>
<tr>
<td>Belgium</td>
<td>Non-Western immigrants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Moroccan</td>
<td>0.8</td>
</tr>
<tr>
<td></td>
<td>Turkish</td>
<td>0.4</td>
</tr>
<tr>
<td>Canada</td>
<td>Immigrants</td>
<td>18.4</td>
</tr>
<tr>
<td></td>
<td>Visible minorities</td>
<td>13.4</td>
</tr>
<tr>
<td></td>
<td>First Nations peoples</td>
<td>2.1</td>
</tr>
<tr>
<td></td>
<td>Francophones</td>
<td>15.7</td>
</tr>
<tr>
<td>Chile</td>
<td>Recent immigrants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Argentinean</td>
<td>1.2</td>
</tr>
<tr>
<td></td>
<td>Peruvian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bolivian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ecuadorian</td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>Immigrant groups</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>European</td>
<td>3.33</td>
</tr>
<tr>
<td></td>
<td>North African</td>
<td>2.22</td>
</tr>
<tr>
<td></td>
<td>Other African</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>Asian</td>
<td>0.96</td>
</tr>
<tr>
<td>Germany</td>
<td>Migrant workers/immigrants</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Turkish</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>Southern European countries</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Reimmigrants (Volga-Germans)</td>
<td>2.8</td>
</tr>
<tr>
<td>Greece</td>
<td>Immigrants</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>Albanian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Bulgarian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Georgian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Romanians</td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>Within Hindu Castes(^{a})</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Scheduled castes</td>
<td>15.06</td>
</tr>
<tr>
<td></td>
<td>Scheduled tribes</td>
<td>7.51</td>
</tr>
<tr>
<td></td>
<td>Other backward classes</td>
<td>43.70</td>
</tr>
<tr>
<td></td>
<td>Muslims</td>
<td>13.0</td>
</tr>
<tr>
<td>Israel</td>
<td>Palestinian Arabs</td>
<td>22.0</td>
</tr>
<tr>
<td></td>
<td>Druze</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Sephardic Jews</td>
<td>31.0</td>
</tr>
<tr>
<td></td>
<td>Iraq</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Iran</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Morocco</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Ethiopia</td>
<td></td>
</tr>
<tr>
<td>Italy</td>
<td>Albanian</td>
<td>1.0</td>
</tr>
<tr>
<td></td>
<td>Rumanian</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Moroccan</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td>Ukrainian</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>North and South Korean</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>0.4</td>
</tr>
<tr>
<td></td>
<td>Brazilians</td>
<td>0.2</td>
</tr>
<tr>
<td></td>
<td>Philippines</td>
<td>0.1</td>
</tr>
</tbody>
</table>

continued
### TABLE 30.1 (continued)
Disadvantaged Groups Within Each Country

<table>
<thead>
<tr>
<th>Country</th>
<th>Group</th>
<th>Percentage of Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>Foreigners</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Asians</td>
<td>7.0</td>
</tr>
<tr>
<td></td>
<td>Europeans</td>
<td>51.5</td>
</tr>
<tr>
<td></td>
<td>Less populous Kenyan tribes</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(Swahili, Kalenjin, Kamba, Kisii, Ameru, Embu, Maasai, Somali, Turkana, Taita, and Samburu)</td>
<td></td>
</tr>
<tr>
<td>Korea</td>
<td>Foreigners</td>
<td>0.8</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>Non-Western immigrants</td>
<td>10.5</td>
</tr>
<tr>
<td></td>
<td>Turkish</td>
<td>2.2</td>
</tr>
<tr>
<td></td>
<td>Moroccan</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Surinamese</td>
<td>2.0</td>
</tr>
<tr>
<td></td>
<td>Antillean/Aruban</td>
<td>0.8</td>
</tr>
<tr>
<td>New Zealand</td>
<td>Pacific peoples</td>
<td>6.4</td>
</tr>
<tr>
<td></td>
<td>Maori</td>
<td>13.5</td>
</tr>
<tr>
<td>South Africa</td>
<td>Black (disadvantaged majority)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>African</td>
<td>79.5</td>
</tr>
<tr>
<td></td>
<td>Colored (mixed race)</td>
<td>8.9</td>
</tr>
<tr>
<td></td>
<td>Indian</td>
<td>2.5</td>
</tr>
<tr>
<td>Spain</td>
<td>Immigrant groups</td>
<td>9.25</td>
</tr>
<tr>
<td></td>
<td>Moroccan</td>
<td>1.16</td>
</tr>
<tr>
<td></td>
<td>Ecuadorian</td>
<td>1.01</td>
</tr>
<tr>
<td></td>
<td>Rumanian</td>
<td>0.89</td>
</tr>
<tr>
<td></td>
<td>Colombian</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>Argentinean</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>Bolivian</td>
<td>0.31</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>Peruvian</td>
<td>0.21</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Immigrant groups</td>
<td>21.9</td>
</tr>
<tr>
<td></td>
<td>Ex-Yugoslavia</td>
<td>4.7</td>
</tr>
<tr>
<td></td>
<td>Italians</td>
<td>4.1</td>
</tr>
<tr>
<td></td>
<td>Portuguese</td>
<td>2.5</td>
</tr>
<tr>
<td></td>
<td>Germans</td>
<td>2.4</td>
</tr>
<tr>
<td></td>
<td>Rhaeto-Romanic-speaking</td>
<td>0.5</td>
</tr>
<tr>
<td></td>
<td>Swiss</td>
<td></td>
</tr>
<tr>
<td>Taiwan</td>
<td>Taiwanese aborigines</td>
<td>2.0</td>
</tr>
<tr>
<td>Turkey</td>
<td>Religious minorities</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Alevi</td>
<td>20.0</td>
</tr>
<tr>
<td></td>
<td>Christian and Jewish</td>
<td>0.3</td>
</tr>
<tr>
<td></td>
<td>Kurdish</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>Arabic</td>
<td>1.5</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>Armenian</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greek</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Jewish</td>
<td></td>
</tr>
</tbody>
</table>
participate in the workforce, ranging from approximately one quarter of women in Turkey (State Institute of Statistics, 2006) to between 60 and 70% in France (Attal-Toubert & Lavergne, 2006), Kenya (primarily due to the high involvement of women in small-scale farming and pastoralism), New Zealand, and the United Kingdom. These differences are undoubtedly at least partially due to the multitude of differences among countries including those in history, culture and values, economic conditions, and political conditions. It is interesting to note that in no instance is the female participation rate higher than the male participation rate; this may partially reflect the traditional division of labor between men and women, such that women are more likely to have household and childcare duties.

Although women are less likely to participate in the workforce than their male counterparts, it appears that there tend to be no or small differences in the unemployment rate for men and women (usually within 1 or 2 percentage points). In fact, in recent years in Taiwan, the unemployment rate for women has been lower than that for men. Exceptions to this trend include Greece (where the unemployment rate of women is often 2 to 3-fold that of men), Kenya, and Switzerland, where women are still substantially more likely to be unemployed than male workers. However, it must be noted that even small changes in the unemployment rate may have strong repercussions for the economic, political, and social situation of a country.

Among all nations surveyed, there is still gender disparity in pay, and this disparity continues to be substantial in magnitude. Among all countries where gender disparity information was available, women earned between 11 and 34% less than men. However, this figure may be lower or higher among countries where we currently do not have the information available. Although it is unclear as to whether these estimates take into account factors such as differences in occupations, differences in full- versus part-time work, differences in educational attainment, etc., other research has shown that even taking into account some of these extraneous factors, women still earn less than their male counterparts (although the magnitude does decrease slightly). The U.S. General Accounting Office (2003) reported that women still only earn 80% of what men earn (compared to 75% when not taking into account differences) in 2000 after taking into account occupation, industry, race, marital status, and job tenure. Currently, the most positive outlook for women’s earning are in Belgium, France, Israel, New Zealand, Switzerland, and the United Kingdom, where women earn 80 cents or more for every dollar earned by men (Equal Opportunities Commission, 2004).

There continues to be occupational segregation to some extent in all 22 countries. Across the board, women are still more likely to be found in clerical or secretarial, retail or sales, healthcare

---

**TABLE 30.1 (continued)**

Disadvantaged Groups Within Each Country

<table>
<thead>
<tr>
<th>Country</th>
<th>Group</th>
<th>Percentage of Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td>Indian</td>
<td>1.78</td>
</tr>
<tr>
<td></td>
<td>Pakistani</td>
<td>1.26</td>
</tr>
<tr>
<td></td>
<td>Black Caribbean</td>
<td>0.95</td>
</tr>
<tr>
<td></td>
<td>Black African</td>
<td>0.82</td>
</tr>
<tr>
<td></td>
<td>Bangladeshi</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>Chinese</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>Other</td>
<td>2.1</td>
</tr>
<tr>
<td>United States</td>
<td>Black/African American</td>
<td>12.3</td>
</tr>
<tr>
<td></td>
<td>Hispanic/Hispanic American</td>
<td>12.5</td>
</tr>
<tr>
<td></td>
<td>Native American and Alaskan Native</td>
<td>0.9</td>
</tr>
</tbody>
</table>

*a The Hindu caste system differentiates between “forward” (advantaged) and “backward” (disadvantaged) groups. A national “schedule” or classification of castes differentiates between scheduled castes (previously “untouchable” castes), scheduled tribal groups, and other backward castes.*

TABLE 30.2
Women’s Status in the Workplace Within Each Country

<table>
<thead>
<tr>
<th>Country</th>
<th>Percentage of Workforce Population</th>
<th>Percentage of Men Participation in Workforce</th>
<th>Percentage of Women Participation in Workforce</th>
<th>Male Unemployment Rate</th>
<th>Female Unemployment Rate</th>
<th>Wage Differential</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>72.0</td>
<td>57.0</td>
<td></td>
<td></td>
<td></td>
<td>66.0</td>
</tr>
<tr>
<td>Belgium</td>
<td>73.6</td>
<td>58.3</td>
<td>7.6</td>
<td>9.6</td>
<td></td>
<td>85.0</td>
</tr>
<tr>
<td>Canada</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>64.0</td>
</tr>
<tr>
<td>Chile</td>
<td></td>
<td>38.2</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>France</td>
<td>46.4</td>
<td>74.5</td>
<td></td>
<td></td>
<td></td>
<td>81.0</td>
</tr>
<tr>
<td>Germany</td>
<td>47.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Greece</td>
<td>64.1–65.0</td>
<td>38.9–42.7</td>
<td>5.1–8.2</td>
<td>13.0–18.6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>India</td>
<td>30.0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Israel</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>81.6</td>
</tr>
<tr>
<td>Italy</td>
<td></td>
<td>69.7</td>
<td>45.3</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Japan</td>
<td>30.0b</td>
<td>73.2–79.4</td>
<td>48.0–50.0</td>
<td></td>
<td></td>
<td>67.1</td>
</tr>
<tr>
<td>Kenya</td>
<td>29.0c</td>
<td>74.7</td>
<td>72.6</td>
<td>25.0c</td>
<td>38.0c</td>
<td></td>
</tr>
<tr>
<td>Korea</td>
<td>41.7</td>
<td>74.4</td>
<td>50.0</td>
<td>3.8</td>
<td>3.1</td>
<td>66.2</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>70.0–77.0</td>
<td>54.0</td>
<td>4.5</td>
<td>6.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>New Zealand</td>
<td></td>
<td>61.2</td>
<td></td>
<td></td>
<td></td>
<td>81.0–87.0</td>
</tr>
<tr>
<td>Spain</td>
<td></td>
<td>42.2</td>
<td>3.5</td>
<td>4.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>South Africa</td>
<td>45.7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td></td>
<td>72.8</td>
<td>56.9</td>
<td></td>
<td></td>
<td>79.0–89.0</td>
</tr>
<tr>
<td>Taiwan</td>
<td>67.6</td>
<td>48.1</td>
<td>4.3</td>
<td>3.9</td>
<td></td>
<td>76.9</td>
</tr>
<tr>
<td>Turkey</td>
<td>36.3</td>
<td>72.3</td>
<td>25.4</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>78.0</td>
<td>69.0</td>
<td></td>
<td></td>
<td></td>
<td>83.0</td>
</tr>
<tr>
<td>United States</td>
<td>46.4</td>
<td>74.0</td>
<td>59.0</td>
<td>4.7</td>
<td>4.5</td>
<td>77.0</td>
</tr>
</tbody>
</table>

The authors representing the various countries have undertaken to report the most recent data available from their country; there may be slight discrepancies between the years reported for each country.

a Percent of women’s salary compared to men’s salary (men’s salary = 100%).
b Percent of full-time workforce.
c In urban areas.
d Within the modern wage sector.
(e.g., nursing, childcare services), education (e.g., elementary school teachers), public services, or small-scale agricultural farming occupations (e.g., Kenya and Turkey) than their male counterparts. Furthermore, the occupations that women are most heavily concentrated in tend to be in the lower income segment. Women remain underrepresented in business and management positions, particularly higher levels of management. In most countries, women continue to lag behind in representation for technical and scientific positions, professional jobs, higher-level governmental positions (e.g., judges, cabinet members, etc.), and most higher-level jobs across sectors.

Authors for several countries note that women are more likely to join the workforce as part-time workers (e.g., Belgium, France, Germany, Japan, Switzerland, and the United Kingdom) to better balance work and family demands or leave the workforce because of childcare demands (e.g., Japan, Korea, and the United States). The latter trend is particularly pronounced in Japan, where the participation ratio by age groups shows an M-shaped curve, because labor force participation rate declines in women’s early 30s because of childcare responsibilities. During the period of 1970–2004, the valley section of this M-curve has shifted northeastward due in part to the trend of late marriage and late childbirth. In addition, both peaks of this M-curve have become higher, indicating that women’s workforce participation has substantially increased in their 20s and late-30s or older (Japan Institute of Labor Policy and Training, 2007). However, some countries also indicated that the wage gap between men and women may be even more pronounced among part-time workers. For example, in the United Kingdom, women are paid 17% less than men in full-time work and 38% less in part-time work (Equal Opportunities Commission, 2004).

Question 3: Is there research documenting mean differences between groups identified above on individual difference measures relevant to job performance?

Mean differences on ability and personality measures are commonly examined in the United States, with enough data for large-scale meta-analytic summaries. Mean differences on tests of developed abilities of roughly 1.00 standard deviation (SD) between Whites and African Americans and roughly 0.67 SD between Whites and Hispanics have been consistently reported. The largest-scale summary of this literature is a meta-analysis by Roth, Bevier, Bobko, Switzer, and Tyler (2001). This abundance of data proves to be in marked contrast to the pattern of findings in the countries examined here. In fact, for most countries, the authors reported finding either no research or research with samples so small that they refrained from drawing conclusions (i.e., Chile, France, Greece, Italy, Japan, Korea, Spain, Switzerland, Turkey, and the United Kingdom). Although limited, there are some data on group differences in some countries.

Two countries (Australia and Taiwan) report research on cognitive ability differences between aborigines and the advantaged group. The lower cognitive ability scores for Australian aborigines may reflect differences in language and culture. Aborigines in Taiwan, who typically have lower educational attainment (Council of Indigenous Peoples, 2002), also score lower than nonaborigines on several cognitive ability tests. Data from the United Arrangement Commission for college entrance examinations in Taiwan in 2006 showed $d$ values between 0.44 and 0.68 in favor of nonaborigines, depending on the particular test subject (A. Chuang, personal communication, May 1, 2007).

In South Africa, mean score differences on cognitive tests between Black and White groups are normally larger than U.S. studies and have $d$ values of approximately 1.00–1.50, with Whites obtaining the higher mean scores. In a study performed in a South African financial services organization,
Handbook of Employee Selection

d values of 0.99 for verbal ability, 1.03 for a numerical ability, and 1.14 for a diagrammatic ability test were found (SHL, 2006). In South Africa, these differences are largely ascribed to the differences in the educational level of the racial groups. In the 2001 census, it was determined that 22.3% of Africans, 8.3% of Colored (mixed race), 5.3% of Indians, and 1.4% of Whites had no schooling (Statistics South Africa, 2001).

Limited data report lower scores for Arabs than Jews in Israel (Zeidner, 1986), for Canadian First Nations people than for Whites, for New Zealand Maori than for Whites (Chernysnhenko, 2005; Guenole, Englert, & Taylor, 2003), and differences between individuals in various provinces in Kenya (Kinyungu, 2006). Data on personality measures are even more limited than for cognitive ability, with authors reporting personality data from only two countries: a large-scale study of Black-White differences in South Africa (Kriek, 2006) showing small differences and several studies of Dutch-immigrant differences in the Netherlands showing much larger differences (van Leest, 1997; te Nijenhuis, van der Flier, & van Leeuwen, 1997, 2003).

Overall, several findings of interest emerge. First, it is clear that gathering data and reporting mean differences by group is generally far more common in the United States than in virtually all of the countries contributing to this report. This is likely the result of the legal scrutiny to which tests are held in the United States. The Uniform Guidelines on Employee Selection Procedures (U.S. Equal Employment Opportunity Commission, 1978) use adverse impact computations as the basis for a prima facie case of discrimination, and thus, adverse impact resulting from test use is routinely examined, with mean differences between groups and the method of test use (e.g., a high or a low cutoff) functioning as key determinants of adverse impact. Second, although data tend to be more sparse than in the United States, group differences are studied and observed in various settings involving different types of disadvantaged groups (e.g., immigrant groups in Belgium and The Netherlands; native peoples in Australia, New Zealand, and Canada; tribal and provincial differences in Kenya; the native Black population in South Africa; and Arab groups in Israel). Third, as in the United States, there is interest not only in whether there are group differences, but also in understanding the basis for these differences. Language, culture, and differences in educational access and attainment are seen as key concerns in understanding differences in test scores across groups.

In the United States, disparate impact is the basis for a prima facie case of discrimination. The implicit assumption is that various groups are expected to obtain similar mean scores absent bias in the measure. Our data suggest that many European countries target certain groups as immigrants to meet specific labor shortages. Thus, immigrants might have higher or lower abilities, depending whether a country tried to attract highly skilled people (e.g., recent immigrants into Switzerland from northern and western Europe) or tried to attract people with low skills (e.g., Turkish immigrants to Germany). In other words, even if one has a general expectation of no group differences at the population level, a finding of differences between locals and immigrants would be expected given this targeted immigration.

Question 4: Are there laws prohibiting discrimination against specific groups and/or mandating fair treatment of such groups? Which groups are protected? Which employers are covered? Which employment practices are covered (e.g., selection, promotion, dismissal)?

Table 30.3 presents summary information addressing the above questions for each country. Several findings emerge. First, there is some basis for legal protections for members of specified groups in all countries. The bases for these protections vary widely. In many cases the national constitution provides general, or at times specific, protections. This may be seen as analogous to the 5th and 14th Amendments to the U.S. Constitution, which respectively state that “no person shall … be deprived of life, liberty, or property without due process of law,” and that “no state shall … deny to any person within its protection the equal protection of the laws.” However, in virtually all cases there are also specific laws defining specified protected classes, specifying which employment practices are covered and which employers are required to comply. The intent here is to identify the

### TABLE 30.3

**International Laws and Practices**

<table>
<thead>
<tr>
<th>Country</th>
<th>Law</th>
<th>Employers Covered</th>
<th>Employment Practices Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chile</td>
<td>Constitution, Chapter 3 (Rights and Duties), article 19 Nª 16 (Freedom of Work and its protection) and Work Code, Article 2º (2002)</td>
<td>All employers</td>
<td>The Constitution establishes the general nondiscrimination principle on the basis of race, color, sex, age, marital status, union membership status, religion, political opinions, nationality, and national or social origin. In March 2008, a new law went into take effect (law # 20.087). This new law defines discrimination as any action that is against the equal opportunity for all workers. A new regulation will specify the practices that are covered by the law.</td>
</tr>
</tbody>
</table>

*continued*
<table>
<thead>
<tr>
<th>Country</th>
<th>Law</th>
<th>Employers Covered</th>
<th>Employment Practices Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>France</td>
<td>French Constitution of 1958</td>
<td>All employers</td>
<td>Many employment practices including selection, access to training, pay, layoffs, transfers, and job classification.</td>
</tr>
<tr>
<td></td>
<td>International convention of the United Nations (1965) ratified in 1971</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>International convention of the International Labor Organization (1958) ratified in 1981</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“The law concerning the fight against racism” of 1972</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>“The law concerning worker’s liberties in organizations” of 1982</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Treaty of Amsterdam of 1997</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>L. 122-45 from Labor Law</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>225-1 and 225-2 from the Penal Code</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Germany</td>
<td>Allgemeines Gleichbehandlungsgesetz: General Equal Opportunity Law</td>
<td>All employers, except tendency organizations (e.g. religious organizations)</td>
<td>All stages of the employment relationship including placing a job ad, hiring and selection, definition of payment, performance appraisal and promotion, job-related training and job counseling, corporate health services, design of working conditions, social services, and dismissal.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Conditions for access to employment, to self-employment, or to occupation, including selection criteria and recruitment conditions; promotion; access to all types and to all levels of vocational guidance, vocational training, advanced vocational training and retraining, including practical work experience, employment and working conditions; dismissals, pay, membership, and involvement in an organization of workers or employers, or any organization whose members carry on a particular profession, including the benefits provided for by such organizations; social protection, including social insurance and sanitary relief; social provisions; education; and access to disposal and to provision of benefits, which are provided to the public, including housing.</td>
</tr>
<tr>
<td>Greece</td>
<td>Greek Law 3304 of 2005, equal treatment</td>
<td>All employers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Greek Law 3488 of 2006, on equal treatment between people in the labor market</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Legal Entity/Constitution</td>
<td>All Employers/Specifics</td>
<td>Key Practices/Provisions</td>
</tr>
<tr>
<td>----------</td>
<td>---------------------------</td>
<td>-------------------------</td>
<td>--------------------------</td>
</tr>
<tr>
<td>India</td>
<td>Government entities, public sector organizations, and organizations receiving government funding</td>
<td>Selection; previously promotion.</td>
<td>Compensation, staffing, conditions of employment, promotion, training and development, dismissal, severance pay, retirement benefits.</td>
</tr>
<tr>
<td>Israel</td>
<td>All employers 6+</td>
<td>Compensation, staffing, conditions of employment, promotion, training and development, dismissal, severance pay, retirement benefits.</td>
<td>Recruitment, selection, promotion, employment agencies, outplacement procedures, training, working conditions.</td>
</tr>
<tr>
<td>Italy</td>
<td>All employers</td>
<td>Recruitment and hiring, assignment, promotion, demotion, training, fringe benefits, change in job type and employment status, encouragement of retirement, mandatory retirement age, dismissal and renewal of employment contract.</td>
<td>Wages, working hours, other working conditions.</td>
</tr>
<tr>
<td>Japan</td>
<td>All employers</td>
<td>Recruitment and hiring, assignment, promotion, demotion, training, fringe benefits, change in job type and employment status, encouragement of retirement, mandatory retirement age, dismissal and renewal of employment contract.</td>
<td>Recruitment and hiring.</td>
</tr>
<tr>
<td>Kenya</td>
<td>All employers</td>
<td>Recruitment and hiring, assignment, promotion, demotion, training, fringe benefits, change in job type and employment status, encouragement of retirement, mandatory retirement age, dismissal and renewal of employment contract.</td>
<td>Mandatory retirement.</td>
</tr>
<tr>
<td>Country</td>
<td>Law</td>
<td>Employers Covered</td>
<td>Employment Practices Covered</td>
</tr>
<tr>
<td>------------</td>
<td>----------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td></td>
<td>Equal Employment Act of 1987</td>
<td>All employers</td>
<td>Recruitment, selection, compensation, education, training, job placement, promotions, setting a mandatory retirement age, retirement, and dismissal.</td>
</tr>
<tr>
<td></td>
<td>The Act of Employment Promotion and Vocational Rehabilitation for the Disabled of 1990</td>
<td>Employers of 500+ workers for affirmative action clause</td>
<td>Hiring, promotion, transfer, education, and training.</td>
</tr>
<tr>
<td></td>
<td>The Aged Employment Promotion Act of 1991</td>
<td>Employers with 50+ workers</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The Basic Employment Policy Act of 1993</td>
<td>Government employees</td>
<td></td>
</tr>
<tr>
<td></td>
<td>The Netherlands Constitution, Article 1 of 2003</td>
<td>Employers with 300+ employers</td>
<td>Recruitment, hiring, and dismissal.</td>
</tr>
<tr>
<td>New Zealand</td>
<td>Human Rights Act of 1993</td>
<td>All employers</td>
<td>Recruitment, selection, employment agencies, dismissal, labor agreements, education before and during employment, promotion, and working conditions.</td>
</tr>
<tr>
<td></td>
<td>Labour Relations Act, Act 66, of 1995</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Employment Equity Act, No. 55, of 1998</td>
<td></td>
<td>Includes, but is not limited to, recruitment procedures, advertising, selection criteria, appointment and appointment process, job classification and grading, remuneration, employment benefits, terms and conditions of employment, job assignments, working environment and facilities, training and development, performance evaluation systems, promotion, transfer, demotion, disciplinary measure other than dismissal, and dismissal.</td>
</tr>
<tr>
<td>Country</td>
<td>Legal Documents</td>
<td>Employer Categories</td>
<td>Key Areas</td>
</tr>
<tr>
<td>-----------</td>
<td>---------------------------------------------------------------------------------</td>
<td>----------------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Bundesverfassung of 1999 (Swiss Federal Constitution), Bundesgesetz über die Beseitigung von Benachteiligungen von Menschen mit Behinderungen of 2002 (Federal Law for the Equal Treatment of People with Disabilities), Bundesgesetz über die Gleichstellung von Mann und Frau of 1995 (Federal Law for the Equal Treatment of Men and Women), Schweizerisches Zivilgesetzbuch of 1907 (Swiss Civil Code), Bundesgesetz betreffend die Ergänzung des Schweizerischen Zivilgesetzbuches — Obligationenrecht of 1912 (Swiss Code of Obligations)</td>
<td>Public employers, All employers</td>
<td>Includes pre- (particularly), during, and postemployment practices. Includes pre-, during, and postemployment practices (i.e., recruitment, sexual harassment, earnings, promotions, etc.). Protection of employee personality and personal data throughout all stages of the employment process.</td>
</tr>
</tbody>
</table>

continued
<table>
<thead>
<tr>
<th>Country</th>
<th>Law</th>
<th>Employers Covered</th>
<th>Employment Practices Covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>Turkey</td>
<td>Republic of Turkey Constitution of 1982</td>
<td>All employers</td>
<td>Article 70 specifically covers selection for public institutions; other practices are implicitly covered including pay, promotion, and dismissal in other articles.</td>
</tr>
<tr>
<td></td>
<td>Article 10</td>
<td></td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Article 49</td>
<td>All employers (except sea transportation, air transport, agricultural and forestry with less than 50 employees, home services, internships, professional athletes, rehabilitation workers, businesses with 3 workers, handmade art jobs done at home, journalists)</td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Article 50</td>
<td></td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Labor Law, Article 5 of 2003</td>
<td>All employers</td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>UN’s Convention on the Elimination of All Sorts of Discrimination Against Women Article 11</td>
<td>All employers</td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Prime Minister’s office circular of 2004</td>
<td>Public employers</td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Race Relations Act of 1976</td>
<td>All employers, trade unions, professional bodies, and employment agencies</td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Sex Discrimination Act of 1975</td>
<td>All employers, trade unions, professional bodies, and employment agencies</td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Employment Equality (Age) Regulations 2006</td>
<td>All ages, young and old</td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Equal Pay Act of 1970</td>
<td></td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>Disability Discrimination Act 1995</td>
<td></td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td></td>
<td>European Community Directives</td>
<td></td>
<td>Performance appraisal, pay, promotion, and termination practices are implicitly covered; selection is not covered because the law only covers private sector employees who are already employed.</td>
</tr>
<tr>
<td>United States</td>
<td>Civil Rights Act of 1964, Title VII (amended 1972, 1991)</td>
<td>All public employers and private employers with 15+ employees</td>
<td>Range of employment decisions including hiring, compensation, terms, conditions, and privileges of employment.</td>
</tr>
<tr>
<td></td>
<td>Age Discrimination Act 1967</td>
<td>Private employers with 20+ employees, state and local governments</td>
<td>Prohibits discrimination against individuals age 40 or older.</td>
</tr>
<tr>
<td></td>
<td>Americans with Disabilities Act 1990 and Rehabilitation Act 1973</td>
<td>ADA covers private employers, state and local governments; Rehabilitation Act covers federal government; Virtually all employers</td>
<td>Prohibits discrimination against individuals with disabilities in the full range of employment decisions.</td>
</tr>
<tr>
<td></td>
<td>Equal Pay Act 1963</td>
<td>Virtually all employers</td>
<td>Prohibits discrimination against women in pay decisions.</td>
</tr>
</tbody>
</table>

TABLE 30.3 (continued)  
International Laws and Practices
major contemporary federal laws and government decrees, and as such it is not a complete record of all historical employment regulations. For example, in the United States a specialist can rightly note that the Civil Rights Acts of 1866 and 1871 are still relied upon on occasion, although these are not listed in the table. Also, several states and cities have additional statutes, offering protection to groups beyond those covered by federal law.

Second, the protections offered are generally quite sweeping in terms of the types of employers covered. In most cases all employers are covered. Some laws are restricted to government employees, and in some cases, coverage is restricted to larger employers, with the coverage threshold varying quite widely for some statutes (e.g., more than 6 employees in Israel, 15 in the U.S., 100 in Taiwan, and 300 in Korea).

Third, it is typical for a broad range of employment practices to be included. Employee selection is specifically included in all countries except Chile, which has the least developed set of employment rights regulations of the countries examined here, and which has yet to specify a set of covered employment practices. However, Chile does prohibit discrimination based on race, color, sex, age, marital status, union membership, status, religion, political opinions, nationality, and national or social origin in its Constitution but does not specify which specific employment practices are covered.

Fourth, there is considerable commonality and variation in the classes that receive protection in each country. Table 30.4 identifies the most common protected classes and indicates whether those classes are covered in each of the contributing countries. The classes covered in U.S. Civil Rights law emerge as widely commonly covered across countries: race, color, religion, gender, national origin, age, and disability status. Three categories not protected by federal statute in the United States are protected in most countries: political opinion, sexual orientation, and marital/family status. Several protected classes are covered in only a few counties or are unique to a few countries; Table 30.5 identifies these less common protected classes. Examples include language, physical appearance, union membership, socioeconomic status, and HIV status.

Question 5: What is required as prima facie evidence of discrimination? What is required to refute a claim of discrimination?

In most countries, direct (e.g., differential treatment) and indirect (e.g., disparate impact) prima facie evidence of discrimination are acknowledged. In India, disparate impact is necessary but not sufficient to prove a case of discrimination; underrepresentation must be shown to be due to historical social or religious discrimination toward a particular group. Only two countries require evidence of the intent to discriminate, Taiwan and Turkey, thus ruling out a disparate impact theory of discrimination.

However, although disparate impact evidence can be used as evidence in most countries, highly specific evidentiary rules used in the United States (e.g., the four-fifths rule and tests of the statistical significance of the difference between passing rates for various groups) are generally not in use (Canada, is an exception, because cases using the four-fifths rule in the United States have been used to make a case for a similar standard). Commentators note that in most cases there are few or no cases involving disparate treatment challenges to predictors commonly used by psychologists, and thus, there is not the extensive case law that has developed in the United States. Recall that the four-fifths rule in the United States derives from guidelines issued by enforcement agencies, and the use of significance testing derives from case law; neither the concept of disparate impact nor the mechanisms for identifying its presence are contained in statute. Absent a history of challenges resulting in case law, it is not surprising to see the lack of specificity as to evidentiary standards.

A similar lack of specificity applies to the question of what is required to refute a claim of discrimination. Table 30.6 summarizes information across countries. In general, there is some version of the shifting burden of proof model in countries where disparate impact evidence is permissible. After a prima facie showing, the burden to justify the use of the employment practice shifts to the employer in all countries except Switzerland, where the burden of showing that the practice is not
### TABLE 30.4
Most Common Protected Classes

<table>
<thead>
<tr>
<th>Country</th>
<th>Race</th>
<th>Sex</th>
<th>National/Ethnic Origin</th>
<th>Color</th>
<th>Age</th>
<th>Religion</th>
<th>Disability</th>
<th>Political Opinion</th>
<th>Sexual Orientation</th>
<th>Marital/Family Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Belgium</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Canada</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Chile</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>France</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Germany</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Greece</td>
<td>X</td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>India</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Israel</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Italy</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Japan</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Kenya</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Korea</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>New Zealand</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>South Africa</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Spain</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Switzerland</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Taiwan</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Turkey</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>United States</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>
job-related is only partially reduced or remains with the plaintiff. There is a general notion that the employer should present evidence to support the job relatedness of the employment practice in question, but rarely is the required form of such evidence specified. The identification of validity evidence as a mechanism for establishing job relatedness is rare.

Question 6: What are the consequences of violation of the laws?

Table 30.6 also summarizes possible consequences of violation in each participating country. There is considerable variation in the array of possible remedies. As a point of reference, note that in the United States the focus is on compensatory or “make-whole” remedies, with punitive damages reserved for instances of intentional discrimination. Similarly, make-whole remedies are part of the landscape in all countries for which information could be obtained. Several countries also provide fines and punitive damages (e.g., Switzerland and Turkey), and several include imprisonment as a possible consequence (e.g., Belgium, France, and Greece).

Question 7: Are particular selection methods limited or banned as a result of legislation or court rulings?

There are relatively few restrictions on specific selection methods. As a point of reference, U.S. law regulates the use of the polygraph, prohibiting its use for most private employers; several other countries restrict polygraph use as well (e.g., Germany, Israel, and Turkey). The only selection method specifically mentioned in U.S. law is the reference in the Tower amendment to Title VII of the Civil Rights Act of 1964 (U.S. Code, 1964) to the permissibility of professionally developed

TABLE 30.6  
Evidence Needed to Refute a Discrimination Claim, Consequences of Violation, and Permissibility of Preferential Treatment by Country

<table>
<thead>
<tr>
<th>Country</th>
<th>Evidence Needed to Refute a Claim</th>
<th>Consequences of Violation</th>
<th>Permissibility of Preferential Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australia</td>
<td>Inherent requirements of the job, existence of special measures to eliminate discrimination, occupational requirements, actions required by law, employment within small organizations, consistent beliefs (e.g. religious organizations or educational institutes). The statutes make no reference to the psychological concept of validity nor has it arisen in case law.</td>
<td>Injunction to stop the act, award of damages, order to the organization to redress the situation, variation, or cancellation of a contract or agreement that violates the law.</td>
<td>Within-group norming is not banned and is used by some psychological testers as a means of complying with legislation (Myors, 2003). Targets may be used in some EEO plans, but explicit quotas are avoided.</td>
</tr>
<tr>
<td>Belgium</td>
<td>Statistical data or practical tests can be used as evidence.</td>
<td>Mediation or binding judgment from civil court. Imprisonment and/or fines.</td>
<td>Preferential treatment is permitted to remedy a historical discrimination against a group. Quotas are permitted, but seldom utilized. Some organizations also utilize target numbers.</td>
</tr>
<tr>
<td>Canada</td>
<td>The employer must demonstrate that the employment policy, practice, or procedure that is challenged is a bona fide occupational requirement. Tribunals and courts are quite liberal in the evidence that they will accept from employers in defense of their employment practices. Empirical and statistical evidence generated by I-O psychologists (e.g., local validation studies) may be useful in defending employment practices, but courts and tribunals often lack the sophistication to make full use of such detailed and complex technical information.</td>
<td>Fines, payment for lost wages, reinstatement, and ordering of special programs.</td>
<td>Preferential treatment permitted (mainly in the public sector).</td>
</tr>
<tr>
<td>Chile</td>
<td>Unclear, unless for sexual harassment or unionization suits. Empirical evidence not required.</td>
<td>Unknown. Currently, sexual harassment suits may result in monetary compensation and up to 3 years imprisonment.</td>
<td>Government has enacted an informal quota for women in minister positions; however, this has not crossed over into the private sector.</td>
</tr>
<tr>
<td>Country</td>
<td>Requirements for Legal Environment</td>
<td>Sanctions for Discrimination</td>
<td>Considerations</td>
</tr>
<tr>
<td>----------</td>
<td>------------------------------------</td>
<td>-----------------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>France</td>
<td>Vague. Employer should present any information showing the decision is legitimate, nondiscriminatory, and based on objective information.</td>
<td>Three years imprisonment and/or a fine for conviction in a criminal court. Discriminatory act is annulled in a civil court and possibly financial compensation.</td>
<td>Considerable discussion about this; politically, preferential treatment is seen as undesirable. However, there are settings where it is used. When parties present lists of candidates for regional and senatorial elections they are required to have an equal number of men and women. Also, there are quotas in one setting: at least 6% of workforce needs to be handicapped for organizations with more than 20 employees.</td>
</tr>
<tr>
<td>Germany</td>
<td>Needs to be based on job requirements.</td>
<td>Employee has right to refuse to work while on payroll and sue employers for damages.</td>
<td>No formalization, but public authorities are to give preference to women and handicapped persons. Preferential treatment to prevent or compensate for disadvantages linked to any of the protected classes.</td>
</tr>
<tr>
<td>Greece</td>
<td>Employer must show that there has been no breach of the principle of equal treatment.</td>
<td>The employer who infringes the laws about equal treatment on the grounds of racial or ethnic origin, religion or belief, disability, age or sex is punished by imprisonment of 6 months up to 3 years and together with a penalty of 1,000 up to 5,000 euros.</td>
<td>Preferential treatment in the form of a relaxation of qualifying scores for protected groups in external recruitment is permitted; however, a common standard is required for promotion. Not all members of protected groups are equally eligible, also dependent on social/economic status. Government positions also use quotas.</td>
</tr>
<tr>
<td>India</td>
<td>At the discretion of the judge.</td>
<td>Small fines. Hiring, reinstatement, or career advancement of plaintiff, payment of back wages.</td>
<td>Preferential treatment is required by public organizations and state-owned enterprises for women and minorities. Preferential treatment is permitted in the private sector.</td>
</tr>
<tr>
<td>Israel</td>
<td>Evidence of test reliability and validity, which can be based on validity generalization. In addition, the National Labor Court recently ruled that employers seeking to prove their innocence will be subject to less severe tests of selection validity to the extent that they are accused of discriminating against internal as opposed to external candidates; the logic being that employers typically have far greater information upon which to base a selection decision when choosing among internal candidates.</td>
<td>Unknown.</td>
<td>Preferential treatment permitted for women.</td>
</tr>
<tr>
<td>Italy</td>
<td>Validity evidence not requested. Evidence to refute a claim is currently unclear.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Evidence Needed to Refute a Claim</td>
<td>Consequences of Violation</td>
<td>Permissibility of Preferential Treatment</td>
</tr>
<tr>
<td>---------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------------------------------------</td>
</tr>
<tr>
<td>Japan</td>
<td>Administrative advice.</td>
<td>Preferential treatment permitted and supported by the government. Quotas required for disabled.</td>
<td></td>
</tr>
<tr>
<td>Kenya</td>
<td>Must show that decisions were based on applicant aptitudes and abilities. Empirical validity evidence not required.</td>
<td>Remedy by following recommendations of Ministry of Health, Labour, &amp; Welfare. Possible public announcement of violation. Civil fine of maximum 200,000 yen ($2,400 U.S.).</td>
<td>Different cut-off scores are set for members from different ethnic groups to ensure that some members from each group will be selected. There are required quotas of 5% in the private and public sector for disabled individuals.</td>
</tr>
<tr>
<td>Korea</td>
<td>Show job relatedness, but specific method unclear.</td>
<td>National Humans Right Commission will make a binding conciliation resolution. Fines.</td>
<td>Quotas required for disabled. Preferential treatment for women, although firms with over 50% women in workforce are exempt.</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>Generally no validity evidence is requested because the validity of common psychological tests, such as tests for cognitive abilities, personality inventories and assessment center exercises, is taken for granted. Most claims concern direct discrimination or treatment discrimination (Commissie Gelijke Behandeling, 2006). Exceptions are clear-cut cases of indirect discrimination in which inappropriate job requirements were set.</td>
<td>Nonbinding judgment by the Commission of Equal Treatment and possibly judgment referral to a civil court.</td>
<td>Preferential treatment is permitted for women and ethnic minorities (does not have to be equally qualified).</td>
</tr>
<tr>
<td>New Zealand</td>
<td>Unclear, because few cases make it to court. Genuine Occupational characteristics (GOQ).</td>
<td>Apology, payment or compensation, assurance that the discriminatory act will not be repeated, or referral to a Human Rights Tribunal for further judgment.</td>
<td>This is currently being explored. Preferential treatment appears to be permitted (and may be soon applied to the Maori population).</td>
</tr>
<tr>
<td>South Africa</td>
<td>Qualitative and empirical data can be brought to bear to support validity.</td>
<td>Fines. Possible cancellation of government contracts.</td>
<td>Preferential treatment is permitted and applied. Racial quotas are legal and practiced by many large employers. The practical implication for this is that it is legal in the South African context to use race norming, or within-group top-down selection strategies, to address affirmative action needs of organizations.</td>
</tr>
<tr>
<td>Country</td>
<td>Legal Requirements</td>
<td>Remedies Provided</td>
<td>Treatment Provided</td>
</tr>
<tr>
<td>-------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------</td>
<td>---------------------------------------------------------</td>
</tr>
<tr>
<td>Spain</td>
<td>Recent laws may lead to greater focus on empirical evidence; up until now, validity of tests was taken for granted.</td>
<td>Compensation, rejection of the decision, and subsequent application of the court decision, repetition of the selection process with new procedures.</td>
<td>Preferential treatment for women in some cases.</td>
</tr>
<tr>
<td>Switzerland</td>
<td>Empirical evidence not generally presented or required.</td>
<td>Courts can award damages including payment of owed earnings and payment of compensation and satisfaction.</td>
<td>Preference is permitted but not required.</td>
</tr>
<tr>
<td>Taiwan</td>
<td>Provide evidence of job relatedness.</td>
<td>Fines.</td>
<td>Quotas required for aborigines (at least 1% of private organizations’ workforce).</td>
</tr>
<tr>
<td>Turkey</td>
<td>Show that requirement is justified. The employer can show that they took all “reasonable” steps to prevent discrimination. No impact cases involving tests have reached the stage of a court decision, so there is as yet no requirement of validity evidence.</td>
<td>Reinstatement, back pay, and/or monetary damages.</td>
<td>Preference is not required or permitted and is actually forbidden.</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>Show that requirement is justified. The employer can show that they took all “reasonable” steps to prevent discrimination. No impact cases involving tests have reached the stage of a court decision, so there is as yet no requirement of validity evidence.</td>
<td>Court has discretion. Compensation to the plaintiff. Formal investigation by governing bodies that can recommend changes in procedures.</td>
<td>Preference is not permitted, but “positive action” such as encouraging certain groups to apply or offering training to these groups.</td>
</tr>
<tr>
<td>United States</td>
<td>Evidence that the challenged practice is job-related for the position in question and consistent with business necessity (largely through validity studies).</td>
<td>Upon a finding of discrimination, a judge can specify “make whole” remedies, such as back pay, hiring, or reinstatement. There are no punitive damages absent a finding of intentional discrimination.</td>
<td>1991 amendments to Title VII of Civil Rights Act prohibit preferential treatment, specifically in the form of adjusting scores or using separate norms for minority group members. Preferential treatment is permitted after a finding of discrimination as part of a judicially ordered remedy.</td>
</tr>
</tbody>
</table>
ability tests, provided that such tests are not designed, intended, or used to discriminate. Additional instances reported of restrictions on specific selection methods in participating countries include a prohibition against comprehensive personality assessment in Switzerland and a restriction on the use of certain Minnesota Multiphasic Personality Inventory (MMPI) and California Psychological Inventory (CPI) items in Spain.

The most strikingly different approach to regulating selection practices is found in South Africa. Rather than the common approach of a presumptive right of an employer to use a particular method absent a successful challenge by a plaintiff, South African law puts the burden immediately on the employer. According to the Employment Equity Act of 1998 (Government Gazette, 1999), psychological testing and other similar assessments are prohibited unless the test is proven to be scientifically valid and reliable, can be applied fairly to all employees, and is not biased against any employee or group. The Society for Industrial and Organizational Psychology (SIOP) in South Africa published “Guidelines for the Validation and Use of Assessment Procedures for the Workplace” during 2005 to provide guidelines for practitioners in the field of I-O psychology to ensure that their assessment instruments and practices comply with the scientific requirements and international best practices. These guidelines were largely based on the American SIOP guidelines.

Question 8: What is the legal status of preferential treatment of members of minority groups (e.g., quotas or softer forms of preference)?

To set the stage, note that the term “affirmative action” is used in various contexts, only some of which involve preferential treatment for protected groups. Some forms of affirmative action involve outreach efforts to publicize openings and to encourage applications from members of protected groups. However, there is no preferential treatment given once an individual is in the applicant pool. Approaches involving preferential treatment fall into two main classes: (a) those that set differing standards for protected and nonprotected groups without setting aside a specified number or proportion of openings for members of protected groups (e.g., using different cut-off scores, using within-group norming) and (b) quota approaches that set aside a fixed number or proportion of openings for members of protected groups.

Table 30.6 summarizes the status of preferential treatment in the participating countries. Preferential treatment is a domain in which the United States emerges as a clear outlier. Preferential treatment in terms of differing score cutoffs or separate norming of tests within group is prohibited by the U.S. Civil Rights Act of 1991 (U.S. Code, 1991), and the use of quotas is restricted to very limited settings, such as a court-ordered remedy following a finding of discrimination. In contrast, in only two countries do commentators report a prohibition against minority preference (Turkey and the United Kingdom). The types of preference permitted and the settings in which it is used vary widely. The status of quotas varies, from prohibited (Australia), to permitted but rarely used (Belgium), to permitted and widely used (South Africa), to used in government sectors (backward classes in India and women in Chile), to required for certain groups (e.g., aborigines in Taiwan, individuals with disabilities in France, Japan, Kenya, and Korea). Several commentators note that applying lower standards to protected groups (e.g., different cutoffs or within-group norming) is used (Australia, India, and South Africa). In India, lower qualifying scores for protected groups are permitted for external selection, but not for promotion.

Question 9: How have laws and the legal environment affected the practice of science-based employee selection in this country?

In only a few countries (Canada, South Africa, and the United States) is the legal environment seen as having a large effect on science-based employee selection. In general, this can partially be attributed to the much more amorphous legal standards and consequences with regards to employment discrimination in most countries surveyed. The reciprocal relationship between science-based selection and the legal environment will need to be continually monitored because many countries...
are still in the process of developing legal statutes and requirements or establishing guidelines for the prosecution and rulings on employment discrimination.

Overall, most employers in the countries surveyed have great latitude in choosing what selection procedures to utilize. However, most employers are aware of the social and political nature of selection procedures and seem to err on the side of mainstream, popular, and usually well-validated selection methods. The most common type of selection procedures do vary by country. It is common to see reports of increased use of the tools and techniques of science-based selection, but the driving forces are more commonly the presence of multinational firms and consulting firms that import these techniques into the country.

**DISCUSSION**

In this section we offer 35 broad summary statements about the patterns emerging from the narratives from the various countries.

**DISADVANTAGED GROUPS**

1. Disadvantaged groups could be divided into four main groups: immigrants or foreign residents, religious minorities, racial/ethnic minorities, and language group minorities (speak different primary language).
2. Many European (especially European Union) nations have disadvantaged groups who are immigrants or foreign workers. The groups that are disadvantaged are usually Eastern European or African.
3. Many Asian countries also have disadvantaged groups who are immigrants or foreign workers.
4. Many of the racial/ethnic minorities are indigenous people (e.g., Australia, Canada, New Zealand, Taiwan, and the United States).
5. Most disadvantaged groups are a relatively small proportion of the population, most below the 20% “breaking point” specified in research on tokenism (Kanter, 1977).
6. Disadvantaged groups can constitute the majority of the population (e.g., South Africa).

**WOMEN IN THE WORKPLACE**

7. Women are now well represented in the workforce, and between one quarter to approximately one half of the workforce are women in most countries.
8. Women have generally substantially increased their participation rate in the workforce in the last decade. However, men’s rate of participation in the workforce continues to greatly outstrip that of women.
9. Women are still underrepresented in management and professional positions. However, European nations and the United States have a sizeable representation of women in lower and middle-management positions. However, all countries have very few women in top- and senior-management positions.
10. Wage differentials are still sizeable between men and women; women generally earn 60–80 cents to the dollar compared with men.
11. Considerable occupational segregation remains for women, such that women tend to be heavily concentrated in lower-income-segment occupations. These include clerical/sec- retarial jobs, service jobs, nursing and childcare services, and primary education.
12. Women tend to engage in more part-time work (partly because of childcare responsibilities).
**SUBGROUP MEAN DIFFERENCES**

13. Very few countries have research exploring potential mean differences in cognitive ability, personality, or job performance. In terms of cognitive ability, findings usually favor the advantaged group and/or men.
14. Mean differences between local and immigrant populations are affected by immigration policies. Targeting either high- or low-skill immigrants can affect the magnitude and direction of mean differences.

**DISCRIMINATION LAWS**

15. Every country has a law or directive that prevents discrimination on the basis of sex or race/ethnic origin and many other personal characteristics and beliefs.
16. Most discrimination cases seem to be settled by special commissions and/or courts rather than by juries (which do not exist in several countries).
17. In many countries, few actual cases are actually filed and/or brought to trial, not because discrimination does not occur, but because workers do not understand their rights, are not used to protecting these rights (e.g., collectivistic orientation, etc.), or do not see much benefit in going to court.
18. Punishment is generally usually rather light (e.g., minimal to moderate fine or reinstatement, payment of back wages).
19. Concerns about privacy are very prominent in Europe. Many European countries are so concerned that data on race or gender are not collected.

**MAKING AND REFUTING A CLAIM OF DISCRIMINATION**

20. For many countries, although there are laws in place, there is very little clarity about how to establish discrimination and/or what kind of evidence required.
21. Intent to discriminate is not required in most countries (exceptions are Taiwan and Turkey).
22. Most discrimination cases are handled on a case-by-case basis and are based on treating people differently on the basis of group membership (direct discrimination) rather than a procedure or test that systematically disadvantages a group (indirect discrimination). In most countries surveyed, both are illegal.
23. Few actual cases outside of the United States challenging the adverse impact or discriminatory nature of formal tests (cognitive ability or personality) exist, and therefore most countries do not really use validity evidence to refute discrimination.
24. Most countries do not require validity evidence. In many places the empirical validity of formal tests (e.g., cognitive ability, personality) is implicitly assumed.
25. Most countries do not use relevant workforce comparisons as a basis for discrimination although this information is sometimes taken under consideration in certain countries.
26. The evidence to refute a claim of discrimination is usually some qualitative evidence of job-relatedness or bona fide occupational requirement.

**MINORITY PREFERENCE**

27. Minority preference is permitted (and even recommended) in most countries. This is more likely to be true for women or those with disabilities than for racial groups.
28. It is more common for government entities than for private-sector firms to engage in practices involving preferential treatment.
29. Forms of affirmative action vary, ranging from active recruitment and training of women or racial groups that have been traditionally disadvantaged to lower standards for these groups.
30. Quotas are relatively rare but are present in several countries; for example, India (lower castes), Taiwan (aborigines), Korea and France (disabled), and South Africa (race and gender).
31. Explicitly forbidding preferential treatment is rare (e.g., Turkey).

**SPECIFIC SCIENCE-BASED SELECTION TOOLS**

32. Generally, science-based tools are not explicitly referenced in laws or in common legal practices (exceptions include South Africa, Switzerland, and the United Kingdom).
33. Generally, although firms are free to use whatever selection methods they desire, large firms tend to be aware of social and business pressures for effective selection.
34. The selection method that is most limited/banned is the polygraph.
35. Selection practice tends to be influenced more by the presence of multinational corporations and consulting firms than by legal pressures (with the exception of the United States, Canada, and South Africa).

We fully anticipate that some readers may question the value of knowing the legal environment of countries other than their own, because they are inevitably bound by the legal constraints of the country they operate in. We have several responses. First, in today’s global world, more and more firms engage in business that extends across national boundaries. Second, there is value in extending one’s framework beyond the national setting with which one is most familiar. Discovering that the same issue is treated differently elsewhere breaks the mold of viewing a certain set of circumstances as inevitable. Third, documenting these differences sets the stages for comparative research asking questions about why certain variations are found. For example, why is preferential treatment not generally permitted and held in such negative popular opinion in the United States and not in many other countries? Why are some groups protected in some countries but not others?

In conclusion, we hope this compilation of information about perspectives from a wide range of countries is useful to students, researchers, and practitioners around the globe. We encourage international collaborations on other workplace issues, and hope this project provides a useful model.

**AUTHORS’ NOTE**

This research was conducted while Antonio Mladinic was on leave from the Pontificia Universidad Católica de Chile and holding a visiting appointment at the University of Texas at El Paso and while Herman Aguinis was on sabbatical leave from the University of Colorado Denver and holding a visiting appointment at the University of Salamanca (Spain). Oleksandr Chernyshenko is now at the Nanyang Business School, Nanyang Technological University (Singapore).

**REFERENCES**


