Computer-Based Testing and the Internet
Issues and Advances

Edited by

Dave Bartram
SHL Group plc, Thames Ditton, Surrey, UK

Ronald K. Hambleton
University of Massachusetts at Amherst, USA
Computer-Based Testing
and the Internet
Contents

About the Editors vii
List of Contributors ix

Introduction: The International Test Commission and its Role in Advancing Measurement Practices and International Guidelines 1
Thomas Oakland

1 Testing on the Internet: Issues, Challenges and Opportunities in the Field of Occupational Assessment 13
Dave Bartram

2 Model-Based Innovations in Computer-Based Testing 39
Wim J. van der Linden

3 New Tests and New Items: Opportunities and Issues 59
Fritz Drasgow and Krista Mattern

4 Psychometric Models, Test Designs and Item Types for the Next Generation of Educational and Psychological Tests 77
Ronald K. Hambleton

5 Operational Issues in Computer-Based Testing 91
Richard M. Luecht

6 Internet Testing: The Examinee Perspective 115
Michael M. Harris

7 The Impact of Technology on Test Manufacture, Delivery and Use and on the Test Taker 135
Dave Bartram

8 Optimizing Quality in the Use of Web-Based and Computer-Based Testing for Personnel Selection 149
Lutz F. Hornke and Martin Kersting
<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>9</td>
<td>Computer-Based Testing for Professional Licensing and Certification of Health Professionals</td>
<td>163</td>
</tr>
<tr>
<td></td>
<td>Donald E. Melnick and Brian E. Clauser</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>Issues that Simulations Face as Assessment Tools</td>
<td>187</td>
</tr>
<tr>
<td></td>
<td>Charles Johnson</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>Inexorable and Inevitable: The Continuing Story of Technology and Assessment</td>
<td>201</td>
</tr>
<tr>
<td></td>
<td>Randy Elliot Bennett</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>Facing the Opportunities of the Future</td>
<td>219</td>
</tr>
<tr>
<td></td>
<td>Krista J. Breithaupt, Craig N. Mills and Gerald J. Melican</td>
<td></td>
</tr>
</tbody>
</table>

Index 253
About the Editors

Dave Bartram is Past President of the International Test Commission and is heading ITC projects on international guidelines for standards in test use and standards for computer-based testing and the Internet. He is Chair of the British Psychological Society’s Steering Committee on Test Standards and Convenor of the European Federation of Psychologists’ Associations Standing Committee on Tests and Testing. He is President-Elect of the IAAP’s Division 2.

Professor Bartram is Research Director for SHL Group plc. Prior to his appointment with SHL in 1998, he was Dean of the Faculty of Science and the Environment, and Professor of Psychology in the Department of Psychology at the University of Hull. He is a Chartered Occupational Psychologist, a Fellow of the British Psychological Society (BPS) and a Fellow of the Ergonomics Society. In 2004 he received the BPS award for Distinguished Contributions to Professional Psychology. His specialist area is computer-based testing and Internet assessment systems. Within SHL he is leading the development of their next generation of Internet-based delivery systems and the development of a multi-dimensional generic Competency Framework.

He has published large numbers of popular, professional and academic articles and book chapters, and has been the Senior Editor of the BPS Test Reviews. He has been an editor or co-author of several works including the 1992, 1995 and 1997 BPS Reviews of Psychometric Tests; Organisational Effectiveness: the Role of Psychology (with Ivan Robertson and Militza Callinan, published in 2002 by Wiley) and the BPS Open Learning Programme for Level A (Occupational) Test Use (with Pat Lindley, published by BPS Blackwell in 1994).

Ronald K. Hambleton holds the title of Distinguished University Professor and is Chairperson of the Research and Evaluation Methods Program and Executive Director of the Center for Educational Assessment at the University of Massachusetts, Amherst, in the United States. He earned a B.A. in 1966 from the University of Waterloo in Canada with majors in mathematics and psychology, and an M.A. in 1967 and Ph.D. in 1969 from the University of Toronto with specialties in psychometric methods and statistics. Professor Hambleton teaches graduate-level courses in educational and psychological testing, item response theory and applications, and classical test theory.
models and methods, and offers seminar courses on applied measurement topics. He is co-author of several textbooks including (with H. Swaminathan and H. Jane Rogers) *Fundamentals of Item Response Theory* (published by Sage in 1991) and *Item Response Theory: Principles and Applications* (published by Kluwer in 1985), and co-editor of several books including *International Perspectives on Academic Assessment* (with Thomas Oakland, published by Kluwer in 1995), *Handbook of Modern Item Response Theory* (with Wim van der Linden, published by Springer in 1997) and *Adaptation of Educational and Psychological Tests for Cross-Cultural Assessment* (with Peter Merenda and Charles Spielberger, published by Erlbaum in 2005). His research interests are in the areas of item response model applications to educational achievement and credentialing exams, standard-setting, test adaptation methodology, score reporting and computer-based testing. He has received several honors and awards for his more than 35 years of measurement research including honorary doctorates from Umea University in Sweden and the University of Oviedo in Spain, the 1994 National Council on Measurement in Education Career Award, the 2003 Association of Test Publisher National Award for Contributions to Computer-Based Testing, and the 2005 E. F. Lindquist Award for Contributions to Assessment. Professor Hambleton is a frequent consultant to state departments of education, national government agencies and credentialing organizations.
Contributors

Dave Bartram, SHL Group, The Pavilion, 1 Atwell Place, Thames Ditton, Surrey, KT7 0NE, UK

Randy Bennett, Educational Testing Service, Rosedale Road, MS 17-R, Princeton, NJ 08451, USA

Krista J. Breithaupt, American Institute of Certified Public Accountants, Harborside Financial Center, 201 Plaza Three, Jersey City, NJ 07311-3881, USA

Brian E. Clauser, National Board of Medical Examiners, 3750 Market Street, Philadelphia, PA 19104-3190, USA

Fritz Drasgow, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, USA

Ronald K. Hambleton, University of Massachusetts at Amherst, Center for Educational Assessment, Hills South, Room 152, Amherst, MA 01003, USA

Michael Harris, College of Business Administration, University of Missouri, 8001 Natural Bridge Road, St Louis, MO 63121-4499, USA

Lutz Hornke, University of Aachen, Department of Psychology, Aachen Technical University, D-52056 Aachen, Germany

Charles Johnson, Competence Assurance Solutions Ltd, 9 Lawford Road, Rugby, CV21 2DZ, UK

Martin Kersting, University of Aachen, Department of Psychology, Aachen Technical University, D-52056 Aachen, Germany

Richard Luecht, Curry Building 209/ERM, University of North Carolina at Greensboro, P.O. Box 26170, Greensboro, NC 27402-6170, USA

Krista Mattern, Department of Psychology, University of Illinois, 603 East Daniel Street, Champaign, IL 61820, USA
Gerald J. Melican, American Institute of Certified Public Accountants, Harborside Financial Center, 201 Plaza Three, Jersey City, NJ 07311-3881, USA

Donald Melnick, National Board of Medical Examiners, 3750 Market Street, Philadelphia, PA 19104-3190, USA

Craig Mills, American Institute of Certified Public Accountants, Harborside Financial Center, 201 Plaza Three, Jersey City, NJ 07311-3881, USA

Tom Oakland, University of Florida at Gainesville, 1403 Norman Hall, Gainesville, FL 32611-7047, USA

Wim J. van der Linden, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands
INTRODUCTION

The International Test Commission and its Role in Advancing Measurement Practices and International Guidelines

Thomas Oakland
University of Florida, USA

We reside in various communities and live in one world. Although we may reside in Beijing, China, on a kibbutz in Israel, or in Muleshoe, Texas, we are aware of the impact of world events on our lives. This has not always been true. Throughout most of history, life generally was impacted by dominant qualities in one’s community or a region reachable within one day. Important events that impacted the lives of those living thousands of miles away generally were unknown and had little impact outside that immediate area. Now, however, knowledge of events that occur throughout the world can be received in real time and can impact others’ lives greatly.

Electronic technology accounts for much of this change. The use of telephones, radios, televisions, and computers has drawn people who live hundreds, even thousands of miles away into a common neighborhood by enabling them to have improved access to information, establish and maintain relationships, and in other ways engage in activities beyond their communities and nations.


People increasingly are never away from their phones. Cameras transmit pictures showing the flow of cars and people. Credit card use requires digitized information be sent to satellites and then returned to the earth, perhaps to a location thousands of miles from where the card was used. TVs convey pictures of warfare in real time. The use of technology is pervasive.

Computer use may account for more change during the last two decades than any other form of technology. Our use of computers has shaped the ways we work, attend school, bank, invest, pay taxes, acquire and disseminate information, plan vacations, order food, and date—to name only some of the more obvious areas.

The development, use, and availability of any technology, to be acceptable, must serve the public good and should assist professionals, as needed, in their efforts to serve the public.

Change that follows the introduction of new technology may be beneficial or harmful to society. Technology generally is intended to enhance the quality of life for a large number of people by providing needed services at lower costs. However, if unregulated and used inappropriately, technology can adversely impact lives. For example, the viewing of violence and sex by children and youths contributes to behaviors and attitudes that often are narcissistic and do not serve the public. Unregulated technology can be harmful. Thus, efforts are needed to help insure technology serves people well.

Technology also should serve professionals well. The introduction of new technology impacts the manner in which professionals can be expected to conduct their work. Professionals incorporate various forms of electronic technology into their practices when they increase effectiveness and efficiency. Their use of technology is preceded by study that informs them of the best ways to shape and apply this technology.

Tests constitute some of psychology’s most important technology. Their use is universal, often starting with newborns and extending through the elderly. Their use is intended to serve the public by improving the ability to describe important behaviors, identify personal strengths, diagnose disabling conditions, assist in making administrative decisions, and help predict future behaviors. However, test use, if unregulated and used inappropriately, can adversely impact lives.

Computer use is shaping the ways tests are constructed, normed, validated, administered, and scored. The reach of this technology on test practices is broad and international. Thus, involvement of organizations that transcend one community, even one nation, is needed to envision, create, promote, regulate, and in other ways assist in forming and reforming test-related services that serve the professions responsible for test development and use as well as the public.

Many issues pertaining to test development and use are international in scope and thus need to be addressed at this level. The International Test Commission, International Association for the Evaluation of Educational Achievement, International Association of Applied Psychology, International Union of Psychological Sciences, Organization for Economic Cooperation and
Development, World Bank, and World Health Organization are among those providing leadership internationally. The work of the International Test Commission is becoming particularly prominent in this endeavor and is summarized below.

This introductory chapter reviews some historical features of test development and use and highlights the role of the International Test Commission in furthering these efforts. Thus, a goal of this chapter is to describe the context for forming the International Test Commission, initiating the 2002 Winchester ITC Conference on Computer-Based Testing and the Internet, as well as the contents of this book.

**SOWING THE SEEDS OF TEST DEVELOPMENT AND USE**

The first widespread use of tests occurred in China more than 3000 years ago. Measures of problem solving, visual spatial perception, divergent thinking, and creativity were used somewhat commonly. Later, under the Sui dynasty (581–618), a national civil service examination system was established that assessed three broad and important areas: cultural knowledge, planning and administration, and martial arts (Wang, 1993). Forms of this examination system continued in China to the end of the 19th century. However, few if any other countries seemingly duplicated these assessment practices or developed others on a national scale until the 20th century. A discipline of psychology, devoted to the study of individual differences in human behavior, was needed to initiate and sustain advocacy for these developments.

Psychological science first emerged from laboratories established by Fechner, Weber, and Wundt in Germany, by Galton in England, and by other pioneers who helped establish the scientific foundation for this fledging discipline. These early efforts to develop and use various tests and other measures, largely for research purposes, set the stage for later efforts that lead to the creation of psychometrics and other test-related specializations within psychology.

Before World War II, leadership in psychology rested among those who saw psychology as an academic discipline, one that was not sufficiently prepared to offer professional services at the same level as well established professions. Tests were developed and used mainly to conduct research and assess educational attainment. Thus, most psychology departments initially resisted pressure to offer programs that prepared practitioners to develop and use tests.

World War I provided one of the first large scale opportunities to evaluate the use of tests in applied settings. Within the United States, group tests of mental abilities were developed, found to be useful in selecting soldiers, and found to be psychometrically sound. Later efforts to develop other group and individually administered tests relevant to issues in psychology and education also proved to be effective. World War II expended the scope of test use, thus supporting its use in the selection, training, and placement of military recruits. By the 1950s, the viability of applied uses of tests was becoming widely accepted.