

# ASSESSMENT IN HEALTH PROFESSIONS EDUCATION

Edited by

---

Steven M. Downing  
Rachel Yudkowsky



# Assessment in Health Professions Education

The health professions, i.e., persons engaged in teaching, research, administration, and/or testing of students and professionals in medicine, dentistry, nursing, pharmacy, and other allied health fields, have never had a comprehensive text devoted specifically to their assessment needs.

*Assessment in Health Professions Education* is the first comprehensive text written specifically for this audience. It presents assessment fundamentals and their theoretical underpinnings, and covers specific assessment methods. Although scholarly and evidence-based, the book is accessible to non-specialists.

- This is the first text to provide comprehensive coverage of assessment in the health professions. It can serve as a basic textbook in introductory and intermediate assessment and testing courses, and as a reference book for graduate students and professionals.
- Although evidence-based, the writing is keyed to consumers of measurement topics and data rather than to specialists. Principles are presented at the intuitive level without statistical derivations.
- Validity evidence is used as an organizing theme. It is presented early ([Chapter 2](#)) and referred to throughout.

**Steven M. Downing** (PhD, Michigan State University) is Associate Professor of Medical Education at the University of Illinois at Chicago and is the Principal Consultant at Downing & Associates. Formerly he was Director of Health Programs and Deputy Vice President at the National Board of Medical Examiners and Senior Psychometrician at the American Board of Internal Medicine.

**Rachel Yudkowsky** (MD, Northwestern University Medical School, MHPE, University of Illinois at Chicago) is Assistant Professor of Medical Education at the University of Illinois at Chicago. She has been director of the Dr. Allan L. and Mary L. Graham Clinical Performance Center since 2000, where she develops standardized patient and simulation-based programs for the instruction and assessment of students, residents, and staff.

# **Assessment in Health Professions Education**

*Edited by*  
*Steven M. Downing, PhD*  
*Rachel Yudkowsky, MD MHPE*

 **Routledge**  
Taylor & Francis Group  
NEW YORK AND LONDON

First published 2009  
by Routledge  
270 Madison Ave, New York, NY 10016

Simultaneously published in the UK  
by Routledge  
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

This edition published in the Taylor & Francis e-Library, 2009.

To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to [www.eBookstore.tandf.co.uk](http://www.eBookstore.tandf.co.uk).

© 2009 Taylor and Francis

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

**Trademark Notice:** Product or corporate names may be trademarks or registered trademarks, and are used only for identification and explanation without intent to infringe.

*Library of Congress Cataloging-in-Publication Data*

Assessment in health professions education / edited by Steven M. Downing, Rachel Yudkowsky.

p. ; cm.

Includes bibliographical references and index.

1. Medicine—Study and teaching. 2. Educational tests and measurements. I. Downing, Steven M. II. Yudkowsky, Rachel. [DNLM: 1. Health Occupations—education. 2. Educational Measurement— methods. W 18 A83443 2009]

R834.5.A873 2009

610.71—dc22

ISBN 0-203-88013-7 Master e-book ISBN

ISBN 10: 0-8058-6127-0 (hbk)

ISBN 10: 0-8058-6128-9 (pbk)

ISBN 10: 0-203-88013-7 (ebk)

ISBN 13: 978-0-8058-6127-3 (hbk)

ISBN 13: 978-0-8058-6128-0 (pbk)

ISBN 13: 978-0-203-88013-5 (ebk)

*To my husband Moshe Looking forward to the next billion seconds and more*

*To our children Eliezer and Channah Who bring us much pride and joy*

*And in memory of our son Yehuda Nattan May his memory be a blessing to all who loved him*

*~ RY*

# List of Figures

- 1.1 George Miller's Pyramid
- 4.1 G Coefficient for Various Numbers of Raters and Stations
- 6.1 Eight Steps for Standard Setting
- 6.2 Hofstee Method
- 6.3 Hofstee Example
- 6.4 Borderline Group Method
- 6.5 Contrasting Groups
- 6.6 Minimizing Passing Errors
- 8.1 Comparison of the Miller Pyramid and Kirkpatrick Criteria for Learning Outcomes
- 9.1 Miller's Pyramid: Competencies to Assess with Performance Tests
- 9.2 Essential Elements of a Standardized Patient Case
- 9.3 Rating Scale Anchors
- 9.4 Sample Rubric for Scoring a Student's Patient Chart Note
- 9.5 An 8-Station OSCE for an Internal Medicine Clerkship
- 9.6 OSCE Case Example
- 9.7 Construct Underrepresentation
- 9.8 Creating Blueprint Specifications for an OSCE
- 9.9 Some Frameworks to Assist in Blueprinting OSCEs
- 11.1 Miller's Pyramid: Competencies to Assess with Oral Exams
- 12.1 Kolb's Experiential Learning Cycle

12.2 Miller's Pyramid: Sample Elements that can be Included in an Omnibus Assessment Portfolio

12.3 Assessment Portfolio Steps

# List of Tables

- 1.1 ACGME Toolbox Assessment Methods
- 2.1 : Evidence Based on Messick (1989) and AERA, APA, & NCME (1999)
- 2.2 Some Sources of Validity Evidence for Proposed Score Interpretations and Examples of Some Types of Evidence
- 2.3 Threats to Validity of Assessments
- 3.1 Hypothetical Five-Item MC Quiz Results for 10 Students
- 3.2 Hypothetical Communication Skills Ratings for 200 Students by Two Judges
- 3.3 Hypothetical Clinical Performance Ratings of 10 Students by 5 Judges/Raters
- 3.4 Spearman Rho Correlations of Rater Agreement
- 4.1 Data for the Example OSCE Measurement Problem: Ratings from a Piloted Version of the OSCE Examination
- 4.2 G Study Results for Example Problem
- 4.3 D Study Results for Example Problem
- 4.4 ANOVA Table
- 5.1 Types of Scores
- 5.2 Raw Scores, z-Scores, and T-Scores
- 5.3 Item Analysis Example
- 5.4 Correlation of One Test Item Score with Score on the Total Test
- 5.5 Item Classification Guide by Difficulty and Discrimination
- 5.6 Example of Summary Item Statistics for Total Test
- 6.1 Standard Setting Feedback Questionnaire for Judges
- 6.2 Sample Angoff Ratings and Calculation of Passing Score

- 6.3 Sample Simplified/Direct Angoff Ratings and Calculation of Passing Score
- 6.4 Sample Ebel Ratings and Calculation of Passing Score
- 6.5 Sample Hofstee Ratings and Calculation of Passing Score
- 6.6 Comparison of Six Standard Setting Methods
- 7.1 Constructed-Response and Selected-Response Item Formats: Strengths and Limitations
- 7.2 Examples of Constructed-Response and Selected-Response Items
- 7.3 A Revised Taxonomy of Multiple-Choice Item Writing Guidelines
- 7.4 Example of Analytic Scoring Rubric for Short-Answer Essay on the Anatomy of the Inner-Ear
- 8.1 Observational Assessment Goals and Tools
- 8.2 Threats to Validity: In vivo Observational Methods
- 9.1 Threats to the Validity: Performance Tests
- 9.2 Sources of Error in an OSCE
- 10.1 Simulation Based Assessment Goals and Tools
- 10.2 Simulation Based Assessment Planning Steps
- 10.3 Test Blueprint for Clinical Cardiology Using the “Harvey” CPS
- 11.1 Characteristics of a Structured Oral Exam
- 11.2 Blueprinting and Logistical Decisions
- 11.3 Steps in Examiner Training for a Structured Oral Exam
- 11.4 Threats to Validity: Oral Examinations
- 12.1 Threats to Validity: Portfolios
- 12.2 Portfolios in the Health Professions
- 12.3 Practical Guidelines for Assessment Portfolios

# Preface

The purpose of this book is to present a basic yet comprehensive treatment of assessment methods for use by health professions educators. While there are many excellent textbooks in psychometric theory and its application to large-scale standardized testing programs and many educational measurement and assessment books designed for elementary and secondary teachers and graduate students in education and psychology, none of these books is entirely appropriate for the specialized educational and assessment requirements of the health professions. Such books lack essential topics of critical interest to health professions educators and may contain many chapters that are of little or no interest to those engaged in education in the health professions.

*Assessment in Health Professions Education* presents chapters on the fundamentals of testing and assessment together with some of their theoretical and research underpinnings plus chapters devoted to specific assessment methods used widely in health professions education. Although scholarly, evidence-based and current, this book is intended to be readable, understandable, and practically useful for the non-measurement specialist. Validity evidence is an organizing theme and is the conceptual framework used throughout the chapters of this book, because the editors and authors think that all assessment data require some amount of scientific evidence to support or refute the intended interpretations of the assessment data and that validity is the single most important attribute of all assessment data.

## ***The Fundamentals***

Chapters 1 to 6 present some of the theoretical fundamentals of assessment, from the special perspective of the health professions educator. These chapters are basic and fairly non-technical but are intended to provide health professions instructors some of the essential background needed to understand, interpret, develop, and successfully apply many of the specialized assessment methods or techniques discussed in Chapters 7 to 12.

In Chapter 1, Downing and Yudkowsky present a broad overview of assessment in the health professions. This chapter provides the basic concepts and language of assessment and orients the reader to the conceptual framework for this book. The reader who is unfamiliar with the jargon of assessment or is new to health professions education will find this chapter a solid introduction and orientation to the basics of this specialized discipline.

Chapter 2 (Downing & Haladyna) discusses validity and the classic threats to validity for assessment data. Validity encompasses all other topics in assessment and thus this chapter is placed early in the book to emphasize its importance. Validity is the organizing principle of this book, so the intention of this chapter is to provide readers with the interpretive tools needed to apply this concept to all other topics and concepts discussed in later chapters.

Chapters 3 and 4 both concern reliability of assessment data, with Chapter 3 (Axelson & Kreiter) discussing the general principles and common applications of reliability. In Chapter 4, Kreiter presents the fundamentals of an important special type of reliability analysis, Generalizability Theory, and applies this methodology to health professions education.

In Chapter 5, Downing presents some basic information on the statistics of testing, discussing the fundamental score unit, standard scores, item analysis, and some information and examples of practical hand-calculator formulas used to evaluate test and assessment data in typical health professions education settings.

Standard setting or the establishment of passing scores is the topic presented by Yudkowsky, Downing, and Tekian in Chapter 6. Defensibility of absolute passing scores—as opposed to relative or normative passing score methods—is the focus of this chapter, together with many examples provided for some of the most common methods utilized for standard setting and some of the statistics used to evaluate those standards.

## ***The Methods***

The second half of the book—[Chapters 7](#) to [12](#)—cover all the basic methods commonly used in health professions education settings, starting with written tests of cognitive knowledge and achievement and proceeding through chapters on observational assessment, performance examinations, simulations, oral exams and portfolio assessment. Each of these topics represents an important method or technique used to measure knowledge and skills acquisition of students and other learners in the health professions.

In [Chapter 7](#), Downing presents an overview of written tests of cognitive knowledge. Both constructed-response and selected-response formats are discussed, with practical examples and guidance summarized from the research literature. Written tests of all types are prevalent, especially in classroom assessment settings in health professions education. This chapter aims to provide the instructor with the basic knowledge and skills needed to effectively test student learning.

[Chapter 8](#), written by McGaghie and colleagues, overviews observational assessment methods, which may be the most prevalent assessment method utilized, especially in clinical education settings. The fundamentals of sound observational assessment methods are presented and recommendations are made for ways to improve these methods.

Yudkowsky discusses performance examinations in [Chapter 9](#). This chapter provides the reader with guidelines for performance assessment using techniques such as standardized patients and Objective Structured Clinical Exams (OSCEs). These methods are extremely useful in skills testing, which is generally a major objective of clinical education and training at all levels of health professions education.

High-tech simulations used in assessment are the focus of [Chapter 10](#), by McGaghie and Issenberg. Simulation technology is becoming ever more important and useful for teaching and assessment, especially in procedural disciplines such as surgery. This chapter presents the state-of-the-art for simulations and will provide the reader with the tools needed to begin to understand and use these methods effectively.

[Chapters 11](#) and [12](#), written by Tekian and Yudkowsky, provide basic information on the use of oral examinations and portfolios. Oral exams in various forms are used widely in health professions education worldwide. This chapter provides information on the fundamental strengths and limitations of the oral exam, plus some suggestions for improving oral exam methods. Portfolio assessment, discussed in [Chapter 12](#), is both old and new. This method is currently enjoying a resurgence in popularity and is widely applied in all levels of health professions education. This chapter presents basic information that is useful to those who employ this methodology.

## ***Acknowledgments***

As is often the case in specialized books such as this, the genesis and motivation to edit and produce the book grew out of our teaching and faculty mentoring roles. We have learned much from our outstanding students in the Masters of Health Professions Education (MHPE) program at the University of Illinois at Chicago (UIC) and we hope that this book provides some useful information to future students in this program and in the many other health professions education graduate and faculty development programs worldwide.

We are also most grateful to all of our authors, who dedicated time from their over-busy professional lives to make a solid contribution to assessment in health professions education.

We thank Lane Akers, our editor/publisher, at Routledge, for his encouragement of this book and his patience with our much delayed writing schedule. We also wish to acknowledge and thank all our reviewers. Their special expertise, insight, and helpful comments have made this a stronger publication.

Brittany Allen, at UIC, assisted us greatly in the final preparation of this book and we are grateful for her help. We also thank our families, who were most patient with our many distractions over the long timeline required to produce this book.

*Steven M. Downing  
Rachel Yudkowsky  
University of Illinois at Chicago, College of Medicine  
July 2008*

# **Chapter-specific Acknowledgments**

## **Chapter 2 Acknowledgments**

This chapter is a modified and expanded version of two papers which appeared in the journal, *Medical Education*. The full references are:

Downing, S.M. (2003). Validity: On the meaningful interpretation of assessment data. *Medical Education*, 37, 830–837.

Downing, S.M., & Haladyna, T.M. (2004). Validity threats: Overcoming interference with proposed interpretations of assessment data. *Medical Education*, 38, 327–333.

## ***Chapter 5 Acknowledgments***

The author is grateful to Clarence D. Kreiter, PhD for his review of this chapter and helpful suggestions.

## **Chapter 6 Acknowledgments**

This chapter is an updated and expanded version of a paper that appeared in *Teaching and Learning in Medicine* in 2006:

Downing, S., Tekian, A., & Yudkowsky, R. (2006). Procedures for establishing defensible absolute passing scores on performance examinations in health professions education. *Teaching and Learning in Medicine*, 18(1), 50–57.

The authors are grateful to the publishers Taylor and Francis for permission to reproduce here material from the paper. The original paper is available at the journal's website [www.informaworld.com](http://www.informaworld.com).

## ***Chapter 7 Acknowledgments***

The author is most grateful to Thomas M. Haladyna, PhD for his review of and constructive criticisms and suggestions for this chapter.