

Preface

In the years since the first edition of this book was published, interest in the field of psychological testing and assessment has been expanding dramatically. Evidence of this renewed interest can be seen in the form of new assessment-related journals, as well as a growing array of assessment-related books, articles, and computer accessories. We are proud of the part we have played, however large or small, in stimulating student interest in this field through the creation of a user-friendly teaching system that helps make learning about—and teaching—measurement an enjoyable and intellectually stimulating task. With this third edition, the tradition continues in the form of a teaching system that consists of this textbook, an instructor's manual (Swerdlik & Phillips, 1996), a test bank (Swerdlik, Cohen & Phillips, 1996), and a companion student workbook and study guide entitled *101 Exercises in Psychological Testing and Assessment* (Cohen, 1996).

Before the publication of the first edition of this book, a tacit assumption about measurement textbooks seemed to prevail. Probably dating back to the mid-1950s, the assumption was that a comprehensive measurement textbook was of necessity challenging for students to get through. Given such a widely held expectation, it is hardly surprising that psychological testing textbooks have been jokingly characterized as rites of passage. Knowing the reputation of the textbooks, students became more and more reluctant to take psychological testing courses. For that reason, and considering other factors, including academic battles about the existence of psychological traits, many psychology departments changed the status of the testing course from mandatory to optional for psychology and education majors. In deference to the complaints of students, and perhaps even the concerns of some faculty members about teaching psychological testing from the available textbooks, the testing course was offered only infrequently at some universities, if not dropped altogether. Thus, even though psychological testing was unquestionably a core subject within psychology in the eyes of many, school after school did not find it desirable to offer the course on a regular basis.

Even as enrollments in psychological testing courses dwindled, most instructors through the 1970s and much of the 1980s continued to rely on 1950s-vintage psychological testing texts. Clearly, such textbooks continued to offer the only available authoritative and comprehensive treatment of the subject; whether they were user-friendly was a secondary or tertiary consideration. Another (more emotional than rational) reason to use such books was that they were the very same books from which the instructors themselves had learned about psychological testing.

By the late 1980s, change was long overdue. Many instructors were using the same psychological testing material for undergraduate and graduate

courses, and not finding it entirely satisfactory for either audience. Also, as I reflected in the first edition of the present work, existing texts had other problems:

As I look back on my own experience with psychological testing textbooks through undergraduate, master's-, and doctoral level training, I remember being keenly aware of lack of a hands-on "feel" in the writing; while comprehensive and technically sound, these textbooks were written, to put it charitably, very, very dryly. . . . In fact, some colleagues half-jokingly advised, "If you're going to write a measurement text, make sure it's not very clear." Asked why, they responded, "Because if the writing's clear, few measurement experts and statisticians will think it's very good." (Cohen et al., 1988, p. xx)

Contrary to such well-intentioned advice, the writing in our book was designed to be clear and engaging. Pedagogical aids, including "Close-up" features, photos, and brief biographies of luminaries in testing and assessment, were included. The "feel" of the book with reference to the tests discussed was decidedly hands-on. In my capacity as a clinical psychology intern and then Senior Psychologist at Bellevue Hospital in Manhattan, I had actually administered and supervised the administration of many of the psychological tests discussed. Authors of competing books did not have that same kind of hands-on experience with tests, and that void was apparent in their writing, most noticeably in chapters having to do with clinical and educational testing.

In 1988, the first edition of our book was published. And just as many consumers understandably stuck with the "safe" choice of IBM or AT&T in the face of the introduction of competitors such as Apple or MCI, so many instructors stuck with a more traditional choice of measurement textbook. Still, some instructors were intent on using the best possible textbook for a new generation of students, even if that meant venturing forward with something brand new. For those instructors, time-worn teaching notes (keyed to a text first introduced 34 years before our first edition was published!) had to be traded in. What they were traded-in for, however, was not just a textbook but an integrated teaching system. Beyond the fresh and contemporary perspective of a new and comprehensive textbook, we developed an instructor's manual for the measurement course that was truly an instructor's manual and not just a test-item bank. Another key element of the teaching package was a student's workbook replete with innovative exercises designed to foster active learning of the subject matter, not merely student playback of the textbook's content.

Research conducted by our publisher has since confirmed that few, if any, of the instructors who switched to our book from whatever they were using before have switched again to any other textbook; with every successive year, the base of instructors using our book has grown. Why?

Instructors have found our textbook to be reasonably comprehensive, and yet different from other comprehensive measurement textbooks in several key ways. Perhaps most important, our book strives to create a dialogue with students, speaking to them and encouraging them to think about and cognitively process the material. We genuinely care about communicating with students. Each page has been written and rewritten with the intention of fostering student understanding of the material. We believe that the care and concern in-

vested in the writing of each page comes through to students. Technical material is presented with sensitivity to the student's understanding. New terms are introduced gradually, laying the proper groundwork for the introduction of more new terms. Throughout, we have tried to keep the tone of our work scholarly yet conversational—even when discussing topics so technical that only psychometricians would raise them in conversation.

Over the years, we have also tried to be responsive to the needs of users of this text. We have replied to every letter sent to us by instructors and students. We keep records of letter writers' suggestions and possible modifications in the text for future editions. And even when instructors do not take the initiative to write to us, we have taken the initiative to reach out to them. In fact, one routine mailing seeking response from a sampling of instructors using our second edition yielded rather momentous results. . . .

A letter requesting instructor reaction to our second edition was sent to Professor Peter Peregoy who, according to our publisher's records, taught the measurement course at the University of Pittsburgh, Johnstown. Professor Peregoy forwarded our letter to his colleague, Professor Suzanne M. Phillips, who taught the course more regularly than he did. Professor Phillips wrote us a thoughtful and thought-provoking letter that said, in part, "I think that the text is probably among the best for the students which I teach, but I believe that it has a basic flaw which most texts in this area seem to share. . . . The problem is that parts of the text which focus on the specific tests are too encyclopedic to serve as a teaching tool. . . ."

By way of background, one of our criticisms of many existing texts was that they tended to bear too much resemblance to *Tests in Print* in some sections. In contrast, our own objective in discussing psychometric aspects of various kinds of tests had been to illustrate how those tests meet or fail to meet certain psychometric criteria. Specific tests were discussed for the purpose of integrating facts about their structure, scoring, and interpretation into a more global understanding of general measurement theory, practice, and issues. Still, Professor Phillips's concerns—presumably representative of some segment of the instructor population using our book—had to be addressed. In response to her letter and subsequent correspondence, we began charting a course that would provide for more "anchoring" of discussion of various tests to psychometric principles. This would be accomplished by weaving more psychometric-related material throughout the text and the *Close-ups* and by creating a new feature, *Everyday Psychometrics*, designed to elaborate on such issues in professional and everyday life. Who better to assist us with all of this rewriting than Suzanne M. Phillips? The correspondence between us caused us to take note not only of Professor Phillips's ideas but also of her writing style, which we believed students would find as accessible and "friendly" as our own. Hence, what started as a query addressed to one instructor resulted in a writing collaboration with another. We welcome Suzanne's contribution to this third edition and look forward to working with her in subsequent editions.

In addition to being a bit more focused on the application of psychometric principles than previous editions, this third edition of our text responds to the needs expressed by other instructors who have written to us asking for more discussion of topics such as (1) culture-related issues in test development and

test use; (2) fairness and bias in testing; (3) assumptions inherent in the assessment enterprise; (4) qualitative item analysis; and (5) portfolio and "authentic" assessment. If you are one of the instructors who wrote to us about one of these issues (or any other), we hope that this edition demonstrates to you that we took your suggestions to heart.

This textbook is designed to be a tool for instructors who teach a course, and the students who take a course, variously called "Psychological Testing," "Psychological Assessment," "Psychological Measurement," "Testing and Assessment in Education," and so forth. Two lessons we have learned since the publication of our first edition are that (1) as a group, the instructors who use our book do so with a very wide range of objectives regarding technical aspects of psychological testing and assessment and (2) it is impossible to please all of the instructors all of the time. Consider, for example, that we have on file requests from instructors variously asking for more or less coverage of relatively technical material such as the multitrait-multimethod matrix and the kappa statistic. How could such requests be fielded in the best interest of all concerned?

The second edition of our text ran a total of 886 pages. Although that was only about 70 pages more than the comparable edition of a competitive text, we still believed that the book had to be edited down to a more manageable length for use in a one-semester or one-quarter course. Hence, on the criterion of length alone, the discussion of a technique such as the multitrait-multimethod matrix, which already occupied some half-dozen pages in the second edition, could not be lengthened. In an effort to reduce the overall length of the book, and still be at liberty to add new material and updating where necessary, we shifted some of the material that we believed was not being used by many instructors to the student workbook. Thus, for example, instructors who would like to cover the multitrait-multimethod matrix in detail will find detailed coverage and useful exercises on pages 82–87 in the companion student workbook (Cohen, 1996). Instructors familiar with previous editions of this text will find some features of the main text now transferred to the student workbook.

The seven objectives outlined in the first edition and carried through in the second edition remain unchanged in this third edition. It remains our intention to

1. Provide a thorough, state-of-the-art, and readable description of basic measurement concepts at a level of technical complexity sufficient to equip students to understand technical terms in professional journals, test manuals, and test reports
2. Present up-to-date, reasonably detailed, and well-balanced discussion of various issues in measurement ranging from the issue of heritability in intelligence, to general legal/ethical issues, to administration, scoring, and interpretation concerns with respect to computer-assisted psychological assessment
3. Blend theoretical and applied material in a way so as to provide the student with both a rationale for, and a hands-on feel of, the assessment process
4. Provide ample case illustrations of the wide range of "real world" contexts in which psychological testing and assessment occurs, including clinical,

counseling, neuropsychological, educational, industrial/organizational, and consumer contexts

5. Provide a historical perspective on measurement complete with biographical material on many important contributors such as Alfred Binet, David Wechsler, Hermann Rorschach, Henry Murray, David Rappaport, and Lauretta Bender
6. To excite genuine interest in the field of testing and assessment by writing with warmth, even occasional humor, and liberally illustrating with relevant material
7. To impart a sense of the authors' belief in and respect for the psychological assessment enterprise balanced by a healthy and realistic degree of self-criticism and an eye toward the challenges that still lie ahead.

Previous editions of this text have been praised by many reviewers for the extent to which we have met the objectives we set for ourselves. Every effort has been made in this third edition to maintain that tradition. In the interest of engaging and stimulating student interest, some of the contexts used to illustrate measurement principles are quite varied. Discussion ranges, for example, from psychometric issues in the interpretation of breathalyzer data, to estimating risk of child abuse from test data, to "confessions" of a staff behavior rater in an inpatient institution that maintains a 24-hour behavioral observation and rating program. Closer to home, students will be exposed to some intriguing facts about tests such as the Graduate Record Examination (GRE). In Chapter 7, the *Everyday Psychometrics* feature contains a guide for "translating" student complaints about examination into psychometric parlance. For example, "I spent all last night studying Chapter 3, and there wasn't one item on that test from that chapter!" is translated as "I question the examination's content validity!"

In each chapter of this book, students are invited to think critically and generatively about numerous measurement-related issues ranging from the workplace use of a test that purports to be a measure of personality type (the Myers-Briggs Type Indicator) to the psychometric pros and cons of a widely used neuropsychological test battery (the Luria-Nebraska). Additionally, students are encouraged to understand the complexity of, and form their own opinions about, some of the controversial issues of the day. These range from questions debated in the professional literature (e.g., "Are unisex norms needed for the MMPI?") to assessment-related issues discussed extensively in the popular media, including affirmative action and the assertions of Herrnstein and Murray (1994) in *The Bell Curve*.

As in previous editions, students are provided early on with a firm grounding in areas such as norms, reliability, validity, and test development. We depart from previous editions, however, in the extent to which we demand that students keep applying their knowledge of such areas to specific tests. In this text as well as in its companion study guide and workbook (Cohen, 1996), the student is repeatedly challenged to apply learned templates and models to new instruments. For example, students may be asked to consider issues of reliability in relation to an infant ability measure (the Bayley Scales for Infant Development, Second Edition) or challenged to apply Cronbach's generalizability theory to a measure of field independence (the Group Embedded Figures Test).

Instructors who have used previous editions of this text will find a great deal of new material in the present revision. Chapter 1 has been revised to provide a much broader overview of the field in general, including various psychometric and cultural issues, as well as issues related to the growing use of computers in the assessment enterprise. Discussion of issues related to society at large in the context of assessment has been greatly expanded. We have added material related to the use of tests in self-help publications and the popular media. Perhaps the greatest change in the first chapter is the addition of a discussion concerning assumptions that underlie testing and assessment. Our combined classroom experience suggested that such a discussion of fundamental assumptions was useful in preparing students for and orienting them to the discussion of theory, application, and issues to come.

Some of the other many changes in this edition include an integration of culture-related issues throughout, a presentation of the five-factor model of personality assessment, and greater reliance on and citation of meta-analytic studies reported in the professional literature. When a particular test is focused on, we have generally devoted most or all of the space in a discussion of it to its most current version. Thus for example, the reader will find ample discussion of the MMPI-2, but coverage of the MMPI is somewhat curtailed relative to the second edition. Another change from the second edition entailed the integration of key material from the former "Preschool Assessment" chapter into the third edition's Chapter 10 (Educational Assessment).

We would like to express our appreciation to Louis H. Primavera and Bernard S. Gorman for their joint contribution to this work in the form of the *Close-up* in Chapter 6 on factor analysis, and for their readiness to supply helpful comments and suggestions. Thanks also to Kevin L. Moreland for his assistance in the material in this book on computer-related assessment, and for his co-authorship of a previous version of Chapter 19. Thanks to David W. Stewart for his assistance in the preparation of an earlier version of the chapter on consumer assessment. Thanks to Patricia M. Greene, administrative assistant to John E. Exner, Jr., who prepared for us biographical material on Dr. Exner from which we drew liberally. We would also like to express our appreciation to Stephen A. Karp, George Washington University; Karl Kuhnert, University of Georgia; Chris Piotrowski, University of West Florida; Michael F. Shaughnessy, Eastern New Mexico University; and Terry A. Stinnett, Eastern Illinois University, who served as reviewers for this third edition. Thanks also to Frank Graham at Mayfield Publishing, who pioneered both the use of the term "sponsoring editor" and the best traditions of what it means to be one.

In the pages that follow, one of the concepts to be discussed in the psychometric sense is that of error. But here we make reference to that concept in its nonpsychometric, everyday meaning when we note that the responsibility for any error in this book rests with the authors and not with any of the people who were kind enough to provide us with assistance.

July, 1995

Ronald Jay Cohen
Pawling, New York

Instructors who have used previous editions of this text will find a great deal of new material in the present revision. Chapter 1 has been revised to provide a much broader overview of the field in general, including various psychometric and cultural issues, as well as issues related to the growing use of computers in the assessment enterprise. Discussion of issues related to society at large in the context of assessment has been greatly expanded. We have added material related to the use of tests in self-help publications and the popular media. Perhaps the greatest change in the first chapter is the addition of a discussion concerning assumptions that underlie testing and assessment. Our combined classroom experience suggested that such a discussion of fundamental assumptions was useful in preparing students for and orienting them to the discussion of theory, application, and issues to come.

Some of the other many changes in this edition include an integration of culture-related issues throughout, a presentation of the five-factor model of personality assessment, and greater reliance on and citation of meta-analytic studies reported in the professional literature. When a particular test is focused on, we have generally devoted most or all of the space in a discussion of it to its most current version. Thus for example, the reader will find ample discussion of the MMPI-2, but coverage of the MMPI is somewhat curtailed relative to the second edition. Another change from the second edition entailed the integration of key material from the former "Preschool Assessment" chapter into the third edition's Chapter 10 (Educational Assessment).

We would like to express our appreciation to Louis H. Primavera and Bernard S. Gorman for their joint contribution to this work in the form of the *Close-up* in Chapter 6 on factor analysis, and for their readiness to supply helpful comments and suggestions. Thanks also to Kevin L. Moreland for his assistance in the material in this book on computer-related assessment, and for his co-authorship of a previous version of Chapter 19. Thanks to David W. Stewart for his assistance in the preparation of an earlier version of the chapter on consumer assessment. Thanks to Patricia M. Greene, administrative assistant to John E. Exner, Jr., who prepared for us biographical material on Dr. Exner from which we drew liberally. We would also like to express our appreciation to Stephen A. Karp, George Washington University; Karl Kuhnert, University of Georgia; Chris Piotrowski, University of West Florida; Michael F. Shaughnessy, Eastern New Mexico University; and Terry A. Stinnett, Eastern Illinois University, who served as reviewers for this third edition. Thanks also to Frank Graham at Mayfield Publishing, who pioneered both the use of the term "sponsoring editor" and the best traditions of what it means to be one.

In the pages that follow, one of the concepts to be discussed in the psychometric sense is that of error. But here we make reference to that concept in its nonpsychometric, everyday meaning when we note that the responsibility for any error in this book rests with the authors and not with any of the people who were kind enough to provide us with assistance.

July, 1995

Ronald Jay Cohen
Pawling, New York