Measurement, Design, and Analysis:

An Integrated Approach

Elazar J. Pedhazur
Liora Pedhazur Schmelkin

Measurement, Design, and Analysis

An Integrated Approach

Elazar J. Pedhazur

Liora Pedhazur Schmelkin
Hofstra University


LAWRENCE ERLBAUM ASSOCIATES, PUBLISHERS
1991 Hillsdale, New Jersey Hove and London

Lawrence Erlbaum Associates Inc., Publishers
365 Broadway
Hillsdale, New Jersey 07642

Library of Congress Cataloging-in-Publication Data

Pedhazur, Elazar J.
Measurement, design, and analysis : an integrated approach /
Elazar J. Pedhazur, Liora Pedhazur Schmelkin.
p. cm.
Includes bibliographical references and index.
ISBN 0-89859-555-X/ISBN 0-8058-1063-3 Student Edition
1. Social sciences—Research—Methodology. 2. Social sciences—Statistical methods. 3. Psychology—Research—Methodology.
I. Schmelkin, Liora Pedhazur. II. Title.
H62.P325 1991
300'.1—dc20 91-7184
CIP

Printed in the United States of America
10 9 8 7 6 5 4 3 2 1

Measurement, Design, and Analysis

An Integrated Approach

To Geula
Alan, Danielle, Alex, and Ben Hadar, Lois, and Pearl

Preface

As we begin the writing of this preface, we are reminded of Richards's (1926) observation about his book on the principles of literary criticism: "Few of the separate items are original. One does not expect novel cards when playing so traditional a game; it is the hand which matters" (p. 1). Much has been written on one aspect or another of the topics we address. We hope that we succeeded in dealing a distinctive hand—more important, that the hand we dealt will matter to you.

As indicated by the title of the book, we attempted to present an integrated approach to research in sociobehavioral sciences.[1] An overview of the content, organization, and orientation of the book is given in Chapter 1. Here, we will explain briefly what we hoped to accomplish in the writing of this book.

Generally, the topics we address are dealt with disjointly in textbooks and courses. For example, in books and courses devoted to statistics, substantive and measurement issues tend to be disregarded, unwittingly creating the impression that they have no bearing on the analysis of the data and on the interpretation of the results. In books and courses on measurement, little or no attention is paid to design and analytic questions, and books and courses in research design tend to treat analytic and measurement issues superficially or to ignore them altogether.

This fragmentary approach inevitably fosters a lack of appreciation of the interrelations and interdependencies among the various aspects of the research endeavor. As a result, many students, especially those not planning to engage in research except for the obligatory dissertation or terminal project, view the expectation that they be conversant in "methodology" as inane, if not sadistic.

We are *not* suggesting that it is imperative for all professionals, regardless of their specialties and substantive interests, to be "expert" in measurement, design, and analysis. We do, however, contend that a basic understanding of these areas and how they bear on one

[1]Although some authors distinguish between social and behavioral sciences, most use these terms interchangeably. Yet it has become common practice to refer to "social and behavioral sciences" and "social and behavioral research." For convenience, we adopted Gergen's (1986) nomenclature and speak of sociobehavioral sciences and sociobehavioral research.

another is essential for becoming an intelligent consumer of research, not to mention a competent researcher.

Our aim, then, is to help you become not only proficient in various aspects of research but also to develop the perspective that takes into account their interrelations and interdependencies. Moreover, we hope to help you learn to appreciate the paramount role of theory in guiding the research enterprise.

Although we assume that you have a background in statistics on the level of an introductory course (e.g., variance, covariance, simple ANOVA, correlation), we review these topics before extending the presentation to more advanced ones. A perusal of the table of contents and of Chapter 1 will reveal that we present some advanced topics that until recently were accessible only to people versed in the mathematical language in which they are necessarily couched. However, the widespread availability of computers and software has made it possible even for people lacking in mathematical background to apply the most sophisticated of analytic approaches. The ease with which this can be done has, unfortunately, contributed to a dramatic increase in misapplications of analytic techniques and misinterpretations of results. In the hope of helping you learn to apply meaningfully the analytic approaches we present, we offer extensive commentaries on inputs and outputs of computer programs in the context of the topics being presented.

We doubt whether the wide array of topics we present can be covered, let alone mastered, in a two-semester course. Accordingly, we hope that ours will not be one of those books you rush to sell as soon as the semester is over. We hope that this book will become a companion to which you will keep returning as you broaden and deepen your understanding of research. It is primarily with this in mind that we have provided an extensive bibliography that, we believe, you will find invaluable in your pursuit of knowledge in the broad area of sociobehavioral research.

Finally, as is discussed in Chapter 1, the organization and the mode of presentation afford a great deal of flexibility in the choice and sequencing of topics as well as the level of sophistication with which they are dealt. We hope that this will enable instructors to tailor the book according to their emphases and the level of their students.

Acknowledgments

We are indebted to Jim Gibbons, Office of Epidemiologic Research, New York City Department of Health; Dr. Ellen Koenigsberg; Professor Robert L. Linn, University of Illinois; and Dr. Elizabeth Taleporos, New York City Board of Education, for reviewing portions of the manuscript and for their constructive suggestions for its improvement.

It is with great pride and pleasure that we acknowledge the assistance of Hadar Pedhazur—computer Maven par excellence—who came to our rescue whenever it was necessary to tame our computers, to upgrade them, or to tailor software to our needs.

Our deepest appreciation to Larry Erlbaum for the understanding and patience he exhibited in the face of repeated delays in the completion of the manuscript. His empathy toward authors is in the best tradition of book publishing. We are indebted to Art Lizza for guiding the production of the book with great skill, perspicacity, and responsiveness.

The second author wishes to thank Hofstra University for its support during the writing of this book.

Elazar J. Pedhazur
Liora Pedhazur Schmelkin
Brooklyn and Oceanside, New York

Contents

PART 3: ANALYSIS

Chapter 1
Overview

Unfortunately, many students and professionals in sociobehavioral sciences lack the background necessary to be intelligent consumers of research literature in areas of their interest or employment. Further, in an era of specialization, students and professionals alike frequently resort to the "services" of a "methodologist-consultant" when doing research. Regrettably, expert advice, necessary and beneficial in many situations, is frequently confused with expert prescriptions. Many researchers and doctoral students deem it proper to delegate to the "consultant" the "tasks" of analyzing the data, interpreting the results, and drawing implications from them. Following blindly consultants' prescriptions or entrusting them with the analysis of the data and interpretation of the results is nothing short of bidding them do one's thinking.

A pernicious consequence of this state of affairs is the tendency to not exercise one's critical faculties altogether. Many professionals do not deem it their role to assess research reports but rather to become familiar with their findings, conclusions, and implications. Thus, acceptance or rejection of findings and implications is based not on informed judgment but on extraneous matters (e.g., conventional wisdom, status of the author of the study). We invite you to consider the potentially deleterious effects of this state of affairs on practice and policy decisions in the broad domain of sociobehavioral sciences.

ABOVE ALL: THINK!

The obvious bears, indeed requires, repeating: To be meaningful, any activity, including the reading of research reports, has to emanate, first and foremost, from sound critical thinking. If there is a message we would like to convey from the very beginning, it is that you exercise common sense, that you not let mumbo jumbo and technical jargon get the better of you. An adverse effect of inordinate reliance on methods and quantification is a diminution of critical thinking among researchers and consumers of research alike. There is such an allure, an almost magical quality, in specialized terminologies, in formulas and fancy analyses, particularly when performed by computer, that the likelihood of giving little or no thought to what they mean is high.

Needless to say, knowledge of the methods and analytic approaches employed is essential for a critical evaluation of a research report. Nevertheless, to underscore the importance of using common sense, we would like, at this early stage, to give some examples where exercising it would have sufficed to cast doubt about authors' assertions. Consider the following "justification" offered by Long (1986) for her use of an instrument in her study: "Although the BSRI [Bem Sex-Role Inventory] has been criticized . . . it continues to be widely used" (p. 324. For the same "justification," see Long, 1989, p. 85). Or consider Furnham's (1984) statement that the instruments used in her study "were chosen for their robustness, [and] psychometric satisfactoriness" (p. 283). Referring to one of the instruments, Furnham stated: "It has been found to be a reliable, valid, and economic instrument and used in many studies" (p. 284). Other instruments were characterized as having "a satisfactory psychometric structure" (p. 284) and/or as having "been used extensively" (p. 284).

The preceding quotations constitute *all* one is told about the properties of the measures used. We hope you see the vacuousness of such statements, even if you know little about measurement.

Here is a different kind of example. Reporting on experiments on the "consequences of schematic referencing for social behavior," Sandelands and Calder (1984, p. 755) stated:

> We first tested the possibility that words in the self-referencing condition were more unusual or less common. All of the words given by subjects were coded for their occurrence in everyday language using word frequency norms (Thorndike & Lorge, 1941). (p. 761)

We submit that a modicum of thought would suffice to question, if not reject out of hand, the appropriateness of word frequency norms established in the late 1930's for research carried out in the 1980's. Yet the authors apparently felt that a criterion, no matter how irrelevant, was required lest they be criticized for making unsubstantiated claims. Indeed, so ingrained is the norm of following scientific protocol that the mere inclusion of a reference, no matter how irrelevant, seems to endow the presentation with an aura of scientific rigor and objectivity, even in the eyes of referees and editors.

To reiterate: In order to become an intelligent consumer of research, not to mention a competent researcher, it is essential that you develop knowledge and skills in various aspects of the research endeavor. However, *no amount of technical proficiency will do you any good if you do not think.*

We turn now to an overview of the content of the book, its organization, and orientation.

CONTENT

Presentations of measurement, design, and analysis—the major areas to which this book is devoted—have filled many books and myriad articles. Therefore, it goes without saying that our treatment is not exhaustive. In this section, we delineate the topics we have chosen to present and make some general observations about the reasons for our choices.

MEASUREMENT

Measurement is the Achilles' heel of sociobehavioral research. Although most programs in sociobehavioral sciences, especially doctoral programs, require a modicum of exposure to statistics and research design, few seem to require the same where measurement is con-

cerned. Thus, many students get the impression that no special competencies are necessary for the development and use of measures, and it is, therefore, not surprising that little or no attention is given to properties of measures used in many research studies. Unfortunately, many readers and researchers fail to realize that no matter how profound the theoretical formulations, how sophisticated the design, and how elegant the analytic techniques, they cannot compensate for poor measures.

Many books and numerous articles, whose presentations vary in scope, depth, and sophistication, are available in the broad area of measurement. Some offer general introductory overviews, whereas others deal with more or less specific topics. For example, there are many books and papers on the measurement of achievement, mental abilities, attitudes, and personality, to name but a few areas. Furthermore, there are books and papers devoted to measurement theory, measurement models, psychometric theory, or some such characterization. Certain topics may be unique to a given area, whereas others may require more or less elaboration, depending on the specific context. For example, topics such as multiple-choice versus essay tests, guessing, grading practices, criterion-referenced versus norm-referenced tests, and test equating are dealt with primarily, if not exclusively, in presentations devoted to achievement testing. Projective techniques, response styles, response sets, and the like are generally discussed in presentations dealing with the measurement of personality, attitudes, and the like. Measurement theories also differ to a greater or lesser extent, depending on the specific substantive area addressed (e.g., achievement, abilities, attitudes, personality).

The foregoing remarks should suffice to indicate why choices of topics, scope of coverage, level of presentation, and the like must be made. Our choice of topics was determined primarily on the basis of their role in the research endeavor and in light of their generality or pervasiveness.

Chapter 2 is devoted to a general introduction to the role of measurement in scientific inquiry. Among topics included are definition of measurement, scales of measurement, and the relation between measurement and statistics.

Validity is the single most important topic in sociobehavioral measurement; therefore, we devote two chapters to it. Chapter 3 focuses on criterion-related validation, and Chapter 4 focuses on construct validation.

Among topics presented in Chapter 3 are definition of criterion, nature and types of criteria, prediction, predictive efficiency, and differential prediction.

Chapter 4 begins with a consideration of the meaning of construct and its relation to indicators. Construct validation approaches are then presented under three main headings: (a) logical analysis, in the context of which are discussed such aspects as construct definition, item content, measurement and scoring procedures; (b) internal-structure analysis, in the context of which intuitive introductions to exploratory and confirmatory factor analysis are given; and (c) cross-structure analysis, in the context of which notions of convergent and discriminant validation, and the use of the multitrait-multimethod matrix approach for assessing them, are presented. The chapter concludes with a comment on content validity.

Chapter 5 is devoted to theoretical and practical considerations in the estimation of reliability. Among topics presented are status of reliability in measurement and research, conceptions of reliability, classical test theory and some variations on and extensions of it, approaches to the estimation of reliability with an emphasis on internal consistency, relations between validity and reliability, and adverse effects of unreliability.

Chapter 6 provides an introduction to selected approaches to measurement in sociobehavioral research, organized under the following categories: (a) Rating Scales, (b) Semantic Differential, (c) Interviewing, and (d) Observation. Issues addressed in connection with each of the preceding include construction, application, analysis, interpretation, and sources of bias.

DESIGN

Part 2 begins with a general introduction to science and scientific inquiry (Chapter 7). Among topics addressed are basic versus applied research, differences and similarities between natural and sociobehavioral sciences, sociobehavioral research findings and policy advocacy, and substance and methods in sociobehavioral sciences.

Chapter 8 is addressed to definitions and variables. Under the former, we discuss the role of definitions in scientific inquiry, criteria for good definitions, theoretical definitions in general and in sociobehavioral research in particular, and empirical definitions. Variables are then defined and discussed from measurement and design perspectives.

The three interrelated topics of theory, problems, and hypotheses are the subject of Chapter 9. After some observations regarding definitions of theory, we concentrate on its paramount role in scientific research. Issues considered include theory and facts, theory as frame of reference, and the biasing effects of theory. This is followed by a consideration of confirmation and falsification in scientific inquiry, and the progress of science. The state of theory in sociobehavioral sciences is then examined. Observations about the volatility of sociobehavioral research conclude this section.

The section on problems begins with a discussion of what constitute problems in scientific research. This is followed by a presentation of different formats for problem statement. Relations of problem formats to given theoretical formulations and their implications for the type of design and analysis to be used are discussed and illustrated. We then take up the complex questions of the substantive meaningfulness of problems, researchable and nonresearchable problems, and the role of past research in problem formulation.

The section on hypotheses begins with a presentation and discussion of hypotheses whose formats parallel those of problem statements presented earlier. This is followed by a discussion of hypotheses as both guides and misguides—the role of disconfirmation and of testing alternative hypotheses derived from different theoretical perspectives. A special section is then devoted to statistical tests of hypotheses. Among topics dealt with are controversies surrounding such tests, interpretation and misinterpretation of P values, and distinction between statistical significance and substantive importance. The chapter concludes with a presentation of the components of a decision-based approach to statistical tests of hypotheses.

Chapter 10 is devoted to basic principles and concepts of research design. Two interrelated themes are presented: control and validity. Following a discussion of the critical role of control in scientific inquiry, different forms of control are presented and their relations to the type of research design considered.

The broad topic of validity is then taken up. After a brief overview, the remainder of this section is devoted to internal and external validity. Various threats to internal and external validity are discussed in varying detail, depending on their complexity.

Chapter 11 supplements Chapter 10 by focusing on pervasive artifacts and pitfalls in sociobehavioral research and the threats they pose to the validity of conclusions from such research. The chapter is organized around two major sources of artifacts and pitfalls: the subjects and the researcher. As in Chapter 10, topics are dealt with in greater or lesser detail, depending on their pervasiveness and/or complexity.

Each of the next three chapters is devoted to a different class of designs: Chapter 12 is addressed to experimental designs, Chapter 13 to quasi-experimental designs, and Chapter 14 to nonexperimental designs. Broadly, these chapters consist of definitions and elaborations of elements of the class of designs under consideration; their unique features, strengths, and weaknesses, with special emphasis on their implications for validity; and research examples.