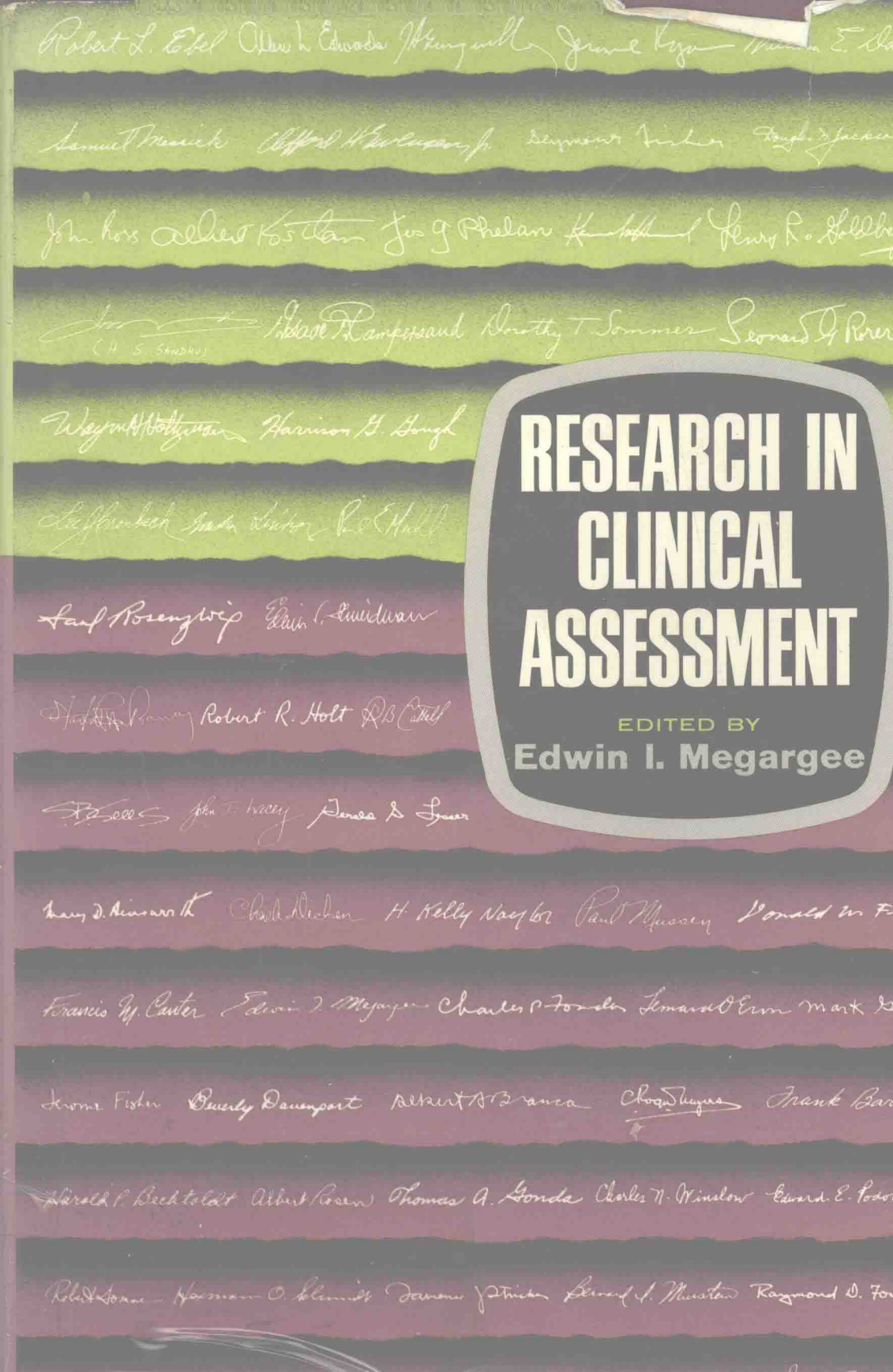
RESEARCH IN CLINICAL ASSESSMENT
EDITED BY
Edwin I. Megargee

RESEARCH IN CLINICAL ASSESSMENT

EDITED BY
Edwin I. Megargee
THE UNIVERSITY OF TEXAS

Harper & Row, Publishers
NEW YORK AND LONDON

Library of Congress Catalog Card Number: 66–17283

TO THE CLINICAL PSYCHOLOGIST—

AWARE OF HIS LIMITATIONS AND THOSE OF HIS INSTRUMENTS,

HE PERSEVERES IN HIS EFFORTS

TO DO THE BEST JOB OF ASSESSMENT

NOW POSSIBLE, WHILE ENGAGING

IN RESEARCH TO IMPROVE

ASSESSMENT TECHNIQUES FOR THE FUTURE.

PREFACE

CLINICAL ASSESSMENT is a complex, frustrating, and fascinating activity. From small samples of behavior the clinician attempts to come to some understanding of each unique personality with which he deals, and, on the basis of this understanding, attempts to predict how this person is going to behave in the future. The intellectual pleasure this activity affords is tempered with the knowledge that this is not a game and that the decisions made can be vitally important. It is rather melodramatic to talk of "life or death" issues, yet the decision as to whether or not a test protocol indicates a brain tumor, whether a patient is suicidal, or whether a jealous husband is potentially homicidal are in sober fact nothing less than life or death issues. The fact that ultimate responsibility for a final decision may often rest in other hands does not absolve the psychologist from supplying whoever makes this decision with the most accurate data possible.

So it remains incumbent on us to be acutely aware of the strengths and limitations of our assessment techniques and to strive continually to test and improve them. Psychologists have responded to this challenge and have been producing an ever-increasing flood of reports on the validity of clinical assessment.

There is, however, a major problem in the communication of these results. To keep abreast of the current output of studies on the Rorschach it is necessary to read no less than three articles each week. Some seventy books on assessment are published each year. The practicing clinician, beset by service demands and hampered by inadequate library facilities, can quickly lose sight of new developments and issues in the field. The problem should be somewhat less acute for individuals in academic settings, but, despite ample library facilities, the research-oriented clinician all too often becomes an expert on one rather narrow topic and remains relatively ignorant of developments in other areas.

This book is designed to improve communication between researchers and practitioners and between those who specialize in one area and those who focus on another. The book, like Gaul, is divided into three parts. The first contains articles dealing with general issues basic to all research on clinical assessment, such as the criterion problem, the usefulness of the concept of construct validity, the influence of base rates on prediction, and whether or not assessment techniques should be validated in the first place.

The second part focuses on the validity of specific assessment techniques. The first section deals with structured tests and the second with unstructured or projective tests. Three chapters are included on special problems peculiar to both types of instruments, such as response sets, multiple scoring systems, and the like.

Part III deals with the integration of clinical data in assessment. The first chapter presents studies aimed at determining how well clinicians integrate data, the second a debate on whether the clinician or the actuary is the best man for the job.

Throughout the book there is an emphasis on diversity, not only in the techniques and approaches covered, but also in the research designs employed. Yet there is an underlying unity as clinicians of vastly different backgrounds and persuasions grapple with similar problems. The died-in-the-wool Rorschacher who

has never extended diplomatic recognition to a paper and pencil test may find with surprise that one of those "dust bowl empiricists" who favor the MMPI has come up with a strategy he is able to apply in his study of projective tests, and the MMPI fancier in turn might be stimulated by work on the TAT.

Not only is there diversity—there is also controversy. The editor feels that there is much to be learned from the head-on clash of differing viewpoints, so whenever possible articles have been selected to present several sides of each issue in an effort to stimulate the reader to think through the problem for himself or to read or experiment further to resolve the matter. Suggestions for further reading are made in the introductions to several chapters.

In selecting the personality tests to be covered, the first criterion was wide use in clinical practice. The MMPI, the Draw-A-Person, the TAT, and the Rorschach were chosen on this basis. Then other less widely used tests were added to provide coverage of important variations in the approach to assessment. Among the other tests covered in some detail are the California Psychological Inventory, the Sixteen Personality Factor Questionnaire, the Myers-Briggs Type Indicator, and the Holtzman Inkblot Technique.

For each test covered, an effort was made to include one study in which the test was used globally and another in which some individual scale or score was validated. Also, an attempt was made to include at least one investigation of the instrument's validity in a clinical setting and another study testing some aspect of the instrument in a laboratory setting.

In choosing among several well-designed studies, the editor would often select the one employing the most novel strategy or the most ingenious criterion measure if other factors were equal. The purpose was to maximize the number of designs employed throughout the book, not only in order to demonstrate the ingenuity of some investigators and the vigor of some instruments, but also, hopefully, to stimulate the use of new approaches by other investigators.

A bias in favor of good writing also guided article selection. All too often psychologists write as if obscurity and dullness were pathognomic of erudition. Whenever possible, papers were chosen which demonstrate that the reader does not have to be anesthetized in order to be informed.

E. I. M.

ACKNOWLEDGMENTS

EDITING a book is similar to managing a baseball team. It is the players who make the team a success, but the manager can make even the best team a failure. So, too, with a book of readings. Whatever merit this book has is a result of the many hours of thought and work put into each individual article by its authors. Without their generous permission and that of the publishers, these articles could not have been reprinted and this book would not have been possible. As a token of gratitude, part of the royalties have been allotted to the authors; as further evidence of their generosity, they have in turn stipulated that these funds should be turned over to the American Psychological Foundation.

While the responsibility for article selection is the editor's, his judgment was improved by the advice of several people. Chief among these was Wayne Holtzman, who was extremely generous with his time and suggestions. The criticism of Harrison Gough was also most helpful. Raymond Cattell and Allen Edwards also took the time to make suggestions which improved the coverage of their work on factor analytic test construction and social desirability, respectively.

When the chapter introductions were being prepared, Joseph Thorpe was kind enough to read them and make suggestions as to how they might be revised.

A major debt of gratitude is owed to the editor's wife, who typed and proofread the entire manuscript while caring for two small children and one medium-sized husband. In justice, she should be listed as co-editor; however, it is cruel enough to ask the reader to pronounce "Megargee" once without making him attempt it twice in succession.

CONTENTS

PART I

Problems in Validating Clinical Methods

Chapter 1

Problems of Research in Clinical Assessment

■ This chapter is designed to introduce some of the general problems of research in clinical assessment which will receive more detailed examination in the subsequent chapters. A secondary purpose is to provide the reader with some yardsticks by which he might judge the adequacy and relevance of the specific validation studies to be found in Parts II and III.

Many of the problems in validational research stem from difficulties in specifying the independent and dependent variables. In the typical laboratory experiment the situation is systematically simplified so that stimulus and response are well defined. This approach can also be applied to assessment, but the more the situation is simplified, the less relevance it bears to the clinical situation. At best this approach can only specify that under certain artificial conditions a lawful relationship exists between stimulus X and behavior Y. However, we have no way of knowing whether such a relationship will also be found when this stimulus is part of a complex interpersonal situation and the dependent variable is influenced by myriad other stimuli.

This can be demonstrated by a consideration of what takes place when a client comes to a psychologist's office. The attitude he brings with him will vary greatly depending on whether he is seeking help for himself or for someone else, whether he is attempting to escape anxiety or a jail sentence, whether he is middle class or lower class, and whether he is under tension or sedation. His initial reactions to the psychologist will be influenced not only by these factors, but also by his past learnings derived from contacts with physicians, social workers, school principals, and the like, as well as his perception of the psychologist as young or old, male or female, Jewish or Gentile, warm or cold, accepting or rejecting. When the psychologist opens the interview, asking, "What brings you here?" it is obvious that the client's response to this initial "stimulus" is going to be determined by a host of variables, and each succeeding response as question follows question and test stimulus follows test stimulus is going to be the product not only of the apparent stimulus, such as Card V on the Rorschach, but also of all that has gone before.

When the client leaves, the psychologist is left with his notes, his tests, his books, his biases, his training, his ringing telephone, and his worries about

the house payments.[1] His task is to isolate the relevant variables and come to some understanding of the client's personality. Then he typically must communicate this perception to another person in the form of a written report containing descriptive statements and, hopefully, behavioral predictions. His task is then finished; for the researcher who wishes to study clinical assessment the task has just begun.

The researcher must attempt to isolate and define independent and dependent variables from this mass of complex behavior if he is to understand the situation or determine the validity of the instruments used. Typically he will focus on a single set of behavior which is repeated over a number of cases such as a psychological test. Given a psychological test protocol as the independent variable, and temporarily ignoring the fact that many variables other than the client's personality helped produce it, the researcher's next task is to find a criterion or some other independent measure of the same trait against which he can validate his test score. In some cases this is relatively easy. If a test is designed to predict flying success, for instance, a good criterion might be the number of aircraft destroyed by each candidate by the end of the training program. On the other hand, if the test is designed to assess strength of pre-Oedipal fixation the task of criterion selection is going to be more difficult.

Another difficulty is the fact that different cues can be used to derive similar predictions while similar cues, in different contexts, can lead to quite different predictions. The fact that the tests and predictions made are less than perfectly reliable makes the process even more confusing.

In this chapter Edwin Shneidman, Paul Meehl, and Kenneth Hammond discuss various specific problems resulting from the general nature of clinical assessment. Shneidman and Meehl deal with the criterion problem while Hammond focuses primarily on research strategies for coping with the complexity of the assessment process. In addition, Shneidman and Meehl propose standards by which we might judge the results of a study of assessment. Shneidman points out that asking if a test is valid is an overly simple approach. Instead, we should determine under what conditions and for what problems it is more or less valid. Meehl argues that the *minimum* requirement for a test is that it yield clear, accurate statements about a person. He indicates additional criteria which should be fulfilled if a test is to be judged useful as well as valid.

REFERENCE

APPELBAUM, S. A., & SIEGAL, R. S. Half-hidden influences on psychological testing and practice. *Journal of Projective Techniques and Personality Assessment*, 1965, **29**, 128–134.

[1] For an account of how extraneous factors may affect an assessment, see Applebaum and Siegal (1965).

Suggestions for the Delineation of Validational Studies

Edwin S. Shneidman

This symposium on "Current Aspects of the Problem of Validity" is sponsored both by the Division of Clinical Psychology and the Society for Projective Techniques. Our topic relates to one of the most timely issues that might have been selected by either group. It is one that is important to our practice and fundamental to our science. As chairman, my own contribution to this symposium is one that might be called "A Fugue and Variations on a Four-Part Invention by Cron-bach," and is essentially a suggested modification of the types of validity formulated by Cronbach's APA Committee on Test Standards in 1952.

One hardly needs to reiterate that the APA Committee on Test Standards proposed four types of validity, as follows:

1. Predictive validity, which denotes correlation between test measures and *subsequent* criterion measures.

2. "Status validity, which denotes correlation between the test and *concurrent* external criteria.

3. "Content validity, which refers to the case in which the specific type of behavior called for by the test is the goal of teaching or some similar activity, as in an academic achievement test.

4. "Congruent validity, which is established when the investigator demonstrates what psychological attribute a test measures by showing correspondence between scores on a test and other indications of the state or attribute. This type of validity is used for tests intended to measure a construct arising from some theory; the validation consists of evidence that the scores vary (from person to person or occasion to occasion) as the theory would imply. Essentially, in congruent validity the meaning assigned to test scores is substantiated by demonstrating that scores are consistent with deductions from the theory from which the meaning derived. This validation process is much the same as that involved in evaluating a theory itself."

My own suggested modification is essentially a spelling out of the several dimensions involved in any validational study. I think that I might go so far as to propose that each of these be indicated at the beginnings of any validational study. The headings I suggest are as follows:

1. Validity for *whom,* in which (a) the subjects, (b) the examiners, and (c) the judges would be described. It makes a lot of difference, for example, whether the 700 Rorschachs being reported in a normative matching study were given to Columbia sophomores or Colombian peasants; whether the examiner was an ABEPP grandfather or his comely young granddaughter, and whether the judges were tyros or Tyroleans. This notion is, of course, a sort of tautology in that it begs the question of what are the relevant dimensions in terms of which the subjects, examiners, and judges need to be described.

2. Validity for *when,* in which one (or more) of the three logical temporal

Paper presented at a symposiom on "Current Aspects of the Problem of Validity" at the American Psychological Association Convention, Washington, D.C., August 30, 1958.

Reprinted by permission of the publisher and the author from the *Journal of Projective Techniques,* 1959, **23,** 259–262.