THE ART AND SCIENCE OF EVALUATION IN THE ARTS THERAPIES

ABOUT THE AUTHORS

Bernard Feder received his Ph.D. from New York University in 1957, and has taught education and psychology at NYU, Adelphi, Hofstra and the University of Sarasota. Among his books are *The Complete Guide to Taking Tests* (Prentice-Hall) and (with David Stutz, M.D.) *The Savvy Patient: How to be an Active Participant in Your Medical Care* (Consumer Reports Books). In addition, he has edited a medical directory and guide and written numerous articles for general and professional journals. In 1990, he was invited to present a history of the arts therapies at the National Library of Medicine of the NIH.

Elaine Feder is a retired dance/movement therapist. Her background includes study with such pioneers in modern dance as Martha Graham, Jose Limon and Doris Humphrey; wide experience in teaching, performing and choreographing; a degree in expressive arts therapies; and extensive graduate study in the psychotherapies and movement analysis. She is a member of the American Dance Therapy Association and the Laban Institute of Movement Studies. Her most recent teaching position, until 1995, was in the Department of Dance at the University of Georgia.

The Feders wrote *The Expressive Arts Therapies: Art, Music and Dance as Psychotherapy* (Prentice-Hall, 1981, 1984), widely used as an introductory text in the arts therapies, and have written jointly for *Psychology Today*, *Human Behavior*, *The New York Times*, *The Chicago Tribune*, and other general and professional publications.

THE ART AND SCIENCE OF EVALUATION IN THE ARTS THERAPIES

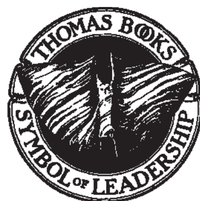
How Do You Know What's Working?

By

BERNARD FEDER

and

ELAINE FEDER



CHARLES C THOMAS • PUBLISHER, LTD.
Springfield • Illinois • U.S.A.

Published and Distributed Throughout the World by

CHARLES C THOMAS • PUBLISHER, LTD.
2600 South First Street
Springfield, Illinois 62794-9265

This book is protected by copyright. No part of it may be reproduced in any manner without written permission from the publisher.

© *1998 by* CHARLES C THOMAS • PUBLISHER, LTD.

ISBN 0-398-06873-9 (cloth)
ISBN 0-398-06874-7 (paper)

Library of Congress Catalog Card Number: 98-3547

With THOMAS BOOKS *careful attention is given to all details of manufacturing and design. It is the Publisher's desire to present books that are satisfactory as to their physical qualities and artistic possibilities and appropriate for their particular use.* THOMAS BOOKS *will be true to those laws of quality that assure a good name and good will.*

Printed in the United States of America
SM-R-3

Library of Congress Cataloging in Publication Data

Feder, Bernard.
The art and science evaluation in the arts therapies : how do you know what's working? / by Bernard Feder and Elaine Feder.
p. cm.
Includes bibliographical references and index.
ISBN 0-398-06873-9 (cloth). -- ISBN (invalid) 0-398-06874-7 (paper)
1. Art therapy. 2. Music therapy. 3. Dance therapy. 4. Mental illness--Treatment--Evaluation. 5. Psychodynamic psychotherapy. 6. Psychotherapist and patient. I. Feder, Elaine. II. Title.
RC489.A7F426 1998
616.89'165--dc21 98-3547
CIP

PREFACE

This book is designed as both an introductory text and a handbook in evaluation and assessment in the emerging arts therapies. We believe that it can be useful both for students in arts therapies training programs and for practitioners in the field who want guidelines for developing and implementing evaluation programs.

At present, there is paucity of materials that correlate the theory and practice of the arts therapies with generally accepted procedures in evaluation.

The first section of the work deals with those fundamentals and principles that apply to all evaluation, qualitative as well as quantitative.

This general treatment is followed by chapters that deal with specific approaches to evaluation: psychometric, clinical or intuitive, and behavioral. The implications for evaluation of the three major philosophical orientations (psychodynamic, existential, and behavioral) are explained.

The last section focuses on evaluation procedures in the individual nonverbal arts therapies. Existing procedures are reviewed and emerging trends are examined.

The basic theme of the book is the interrelationship between the creative and the scientific approaches to evaluation.

While the fundamentals and principles of evaluation are applicable to all the arts therapies, we have included in this last section only the three major nonverbal arts therapies: art, music, and dance/movement. Because psychodrama, drama therapy, and poetry therapy are essentially verbal, they tend to rely heavily on approaches that have been developed in psychotherapy or in verbal group therapies.

B. F.
E. F.

INTRODUCTION

There are two basic themes around which this book is organized. First, we believe that the argument over whether therapy is an art or a science is not only fruitless but counterproductive; it can only perpetuate divisions in a field in which both artistic creativity and scientific validation are necessary. Second, we contend that the therapeutic endeavor has little meaning if therapists cannot formulate defensible ways of ascertaining whether what they do makes a difference.

Let's begin with an overview of our first proposition.

"Art is I," wrote Claude Bernard (1813-1878); "science is we." In Isaac Newton's words, the scientist, seeking to uncover the laws that govern the operation of the universe, stands "on the shoulders of giants." Science is collaborative, incremental, and cumulative. Each scientist adds a bit of understanding to the body of knowledge, to enhance or to correct what had been discovered before. In this sense, even competitors are collaborators. What has been supplanted is either rendered obsolete or is incorporated into the revised perception of the reality of the world around us.

In contrast, the hallmark of art is independence and autonomy, the freedom to break from what was done before, and to create the new. Artists, of course, are not completely free agents; to some degree, they are in bondage to the technology of art, to the limitations of their materials, and to the demands of tradition. More easily than scientists, artists can break with tradition. However, unlike the scientist, the artist doesn't add to the body of knowledge so much as he or she transforms what has already been learned to create a unique statement. This statement does not necessarily detract from what came before. Michelangelo's masterpieces are not diminished by the works of van Gogh, or Monet, or Jackson Pollock.

While the cultures of art and science appear to be distinct, there is an interplay, and there are vast areas of overlap. Discoveries about the properties of clay or glass or marble, improvements in the quality of pigments or oils or tempera, the development of new materials for the manufacture of musical instruments, advances in the production of varnishes, all open up vistas for artists and provide them with the means by which they can conceive, and create, and execute their personal statements. Historian and former Librarian of Congress Daniel J. Boorstin (1994) writes of the symbiotic relationship between what he calls the culture of discovery and the culture of

creation during an age in which both flourished:

> Renaissance belief in the inspired unique creator elevated the painter, equipped with the newly discovered science of perspective, from craftsman to artist. . . . The technique that Giotto had applied by rule of thumb became a science in the hands of da Vinci or Durer. (pp. 24, 29.)

It would be a mistake to think of the artist only as the beneficiary of the fruits of scientific labor. During this age of exuberant discovery and creativity, we must remember, the quintessential Renaissance man was simultaneously discoverer and creator.

In da Vinci's notebooks we find questions and more questions, and we would be hard-pressed to know if these are the questions of an artist or a scientist. How does a bird fly? How does a man walk? How can the trajectory of a mortar shell be described? What does each of the ten ways he could draw a foot reveal about its structure and function? In these notebooks, we find a bewildering assortment of drawings: pumps, a self-locking worm gear, an air hose, a steam engine, a parachute, an airplane, a submarine, roller bearings, sprocket chains, a machine gun. Was this a man who used science to master the skills of the artist? Or a scientist who used art to probe the worlds of anatomy, and geology, and mechanics, and hydraulics?

While we cannot find many Leonardos, for whom creativity and discovery are indivisible, there is a constant interplay between the worlds of art and science. Just as the artist owes much to the discoveries of the scientist, there are significant bodies of scientific knowledge that have been induced by questions posed by artists. The field of "projective geometry," dealing with the images that figures create when they are viewed from different angles, was developed by mathematicians in the seventeenth century as a result of prompting by artists.

Modern psychotherapy owes much to both cultures. From art, it draws on the artistic creative impulse driven by intuitive insight, the ability to discern relationships, to develop the personal empathetic bond between therapist and patient that acts as catalyst in the interpretive and healing processes. From science, it derives the recognition that the creative proposition must conform with what has been discovered about the ways humans actually function, so that we can distinguish between a principled proposition and a whim.

The relationship between art and science in the modern practice of psychotherapy is a restless and disturbed one. With the increasing specialization of occupation, artist and scientist frequently speak in different tongues and have difficulty understanding each other. From what should be a harmonious chorus often comes a disturbing and dissonant cacophony.

Science seeks underlying principles and the natural order of things. The

scientist wants to find the common elements that make humans human, that provide the grand structure of human nature. The individual who deviates from this order is literally the victim of a "disorder." The scientist wants to identify the nature of the disorder so that he or she can bring the victim back to normality–that is, conformity with the statistical norm, the natural order of things. The scientifically minded psychologist asks: What can we learn about depression or psychoses from studying the myriad of people who suffer from these disorders? Are we doomed to see each problem as floating in a vast void, unrelated to similar problems?

Art seeks the unique, the individual, the things that set humans apart. Why and how, the artist asks, is *this* human different from all other humans? How has this individual created his or her personal reality and structured his or her own world? The key to understanding the individual is to peer into that private world, to find the expression of his or her individuality. This ability to find the unique core of the individual constitutes the art of the therapist.

This thread–the uneasy relationship between the therapist as artist and the therapist as scientist–runs through this book. In our discussion of clinical judgment, we pull it to the surface so that it becomes the very fabric of Chapter 9. In this chapter, we explore the difficulties that clinician and researcher often have in communicating on the subject of evaluation.

The divergences may never be resolved, but they should be understood, because the elements of both art and science are essential to a meaningful practice of therapy. Without science, therapy can degenerate to the practice of superstitious ritual, in which each practitioner owes allegiance only to his or her personal myth of existence. Without art, it can lose the very humanity it seeks to examine.

This brings us to the second issue at hand: How can individual arts therapists ascertain the appropriate treatment for their patients or clients, and how can they know whether what they do works?

When we began this book, we lived in the small university city of Athens, Georgia. As we drove from Interstate 85, we would pass a large billboard that proclaimed: PRAY. IT WORKS.

It was difficult for us to pass this sign without comment. Occasionally, we would refer to the experience of Hans J. Eysenck, a psychologist at the University of London's Institute of Psychiatry, who had raised questions about whether psychotherapy "works." Almost a half-century ago, Eysenck published a number of articles in which he questioned the efficacy of psychotherapy, concluding that no method worked better than any other, and that no form of therapy improved on the recovery rate obtained through ordinary life experiences and non-specific treatment.

Eysenck's conclusions were the subject of intense debate among both

clinicians and researchers. We have no wish here to become embroiled in the substance of his studies, which were badly flawed in a number of respects. What was most interesting about the whole affair was the furor his work created at the time in the psychotherapeutic community. The very act of testing the effectiveness of psychotherapy, he reported, aroused emotional responses that he compared with those of a true believer against a blasphemer who had attempted a statistical test of the efficacy of prayer.

Since Freud's day, debate has raged over the effectiveness of psychotherapy. The debate has often been tumultuous and, at times, acrimonious. At one end of the debate stand clinicians who are impressed with improvements they see (or claim to see) in their patients, and are understandably eager to attribute such change to their efforts. At the other end stand the researchers who demand objective evidence that real change has actually taken place and that any such change is the result of the therapeutic intervention.

This book is designed for the individual arts therapist, for whom the issue is not whether there is a change in his or her patients. Change will occur whether a patient is in therapy or not. The central issues are to recognize and identify the nature of the change, and to know with some assurance the degree to which such change is the result of the therapy, and not coincidental with it.

Much has changed in the decades since Eysenck figuratively nailed his theses to the doors of the psychotherapeutic institution. Increasing numbers of both verbal and nonverbal therapists have come to accept the need for more than faith, zeal and uncorroborated anedoctal reports of cures in considering the effectiveness of their work.

This book explores a variety of approaches, both theoretical and methodological. Our purpose is not to provide formulas, which can be found in any basic textbook on psychological testing, or recipes, which abound in professional journals. It is to help therapists to relate their evaluation program to their goals, to identify what they are interested in evaluating and to design the kind of evaluation program that can do what the therapist wants it to do.

In the actual development of this book, Bernard was the designated writer. He was assigned the task of putting into words the ideas on which we had agreed during extended discussions. After each draft, we argued. Elaine, the intuitive enthusiast, and Bernard, the analytic skeptic, would spend hours debating points of contention until we arrived at a consensus. The one position on which we agreed from the beginning was that the arts therapies cannot legitimately lay claim to being professions until arts therapists can establish a credible method for evaluating (literally, ascertaining the value of) their services, and until they can develop ways of knowing that

what they do makes a difference to the troubled individuals with whom they work.

We believe that arts therapists are painfully aware of this problem. In large part, the problem has been brought to their attention through the demands of outsiders, such as insurance companies. In part, it is the result of the maturation process in a field undergoing an awkward adolescence. In recent years, virtually every professional conference includes panels and seminars on assessment, evaluation and research in the therapies. Yet, it is sobering to recognize how few arts therapies programs offer instruction either in research or in evaluation. The major problem now is not the resistance to assessment that Eysenck encountered in the 1960s, but the uncritical zeal with which many practitioners have come to embrace methods and instruments that offer the illusion of certainty, and often without any real understanding of their functions and limitations.

In this connection, it may be instructive to read the words of Oscar Buros, more than a generation ago. In the introduction to *Tests in Print* (1961), he wrote:

> At present, no matter how poor a test may be, if it is nicely packaged and if it promises to do all sorts of things which no test can do, the test will find many gullible buyers.
>
> . . . [Test users] seem to have an unshakable will to believe the exaggerated claims of test authors and publishers. If these test users were better informed regarding the merits and limitations of their testing instruments, they would probably be less happy . . . in their work. The test user who has faith–however unjustified–can speak with confidence in interpreting test results and in making recommendations. The well-informed test user cannot do this; he knows that the best of our tests are still highly fallible instruments which are extremely difficult to interpret with accuracy in individual cases. Consequently, he must interpret test results cautiously and with so many reservations that others wonder whether he really knows what he is talking about. (Buros, 1961, p. xxix.)

A decade later, Buros apparently found that little had changed since his earlier comments, and he wrote in apparent exasperation that "at least half of the tests currently on the market should never have been published. Exaggerated, false, or unsubstantiated claims are the rule rather than the exception" (Buros, 1972, p. xxvii).

We believe that assessment procedures will improve only if the creators and users of these procedures become more knowledgeable about evaluation and assessment than are most therapists today. It is our hope that this book will make some contribution in this regard.

For a number of reasons, this book is not a comprehensive primer on evaluation or a survey of assessment in the arts therapies.

First of all, practical considerations made it infeasible to try to develop a complete guide to evaluation. Such a book would have been prohibitively long and intimidatingly expensive.

In addition, the writing of such a book would have involved us in in-depth research in areas in which we were not comfortable, mainly because we were not familiar with their practical application. At the invitation of several faculty members of the University of Georgia, we considered applying for a grant to involve doctoral students in various areas in which we ourselves were deficient, but we decided that such an endeavor would have been too time-consuming and would have added only marginally to the book.

As a result, we chose to delimit the work in two major areas.

The first decision was to deal only with the assessment of individual clients and patients, and to refer fleetingly to the vast areas of couple, group, and family therapy. While there are some tangential points between the assessment of individuals and the assessment of families and groups, there are compelling reasons to view these areas as distinct categories in the field of the therapies.

The second decision, after a good deal of painful consideration, was to abandon the work we had already begun in examining such areas as psychodrama, drama therapy, and poetry therapy. Because these therapies are fundamentally verbal, evaluation procedures tend to rely heavily on approaches that have been developed either in individual psychotherapy or in couple, family, or group therapies.

We are obligated to those arts therapists who shared with us the evaluation procedures on which they had worked or were working. Many offered comments on their experiences, their philosophies, and their frustrations.

We owe a particular debt of gratitude to those who agreed to review and comment on the chapters in which they had particular expertise and interest. These include my friend and former colleague, Dr. John W. French, who had coordinated College Board research at the Educational Testing Service; Dr. Richard Graham, Director of the School of Music at the University of Georgia and former editor of the *Journal of Music Therapy*; Dr. Jerry Gale of the University of Georgia, whose area of interest is qualitative evaluation; Dr. Charles R. Martin of the Center for Applications of Psychological Types; and the numerous arts therapists, psychiatrists, psychologists, psychometricians and scholars in a variety of fields who offered criticisms and suggestions.

CONTENTS

Page

Preface v
Introduction vii

Chapter

1. PURPOSES OF EVALUATION 3
   Functions of Evaluation 5
   Diagnosis and Assessment 6
      Diagnosis as a Guide to Decision Making 7
      The Meaning of Diagnosis 8
      The Diagnostic Babel 10
      How Useful is Diagnosis? 11
      On Being Sane in Insane Places 12
      Psychiatric Labels and Arts Therapists 15
      Monitoring and Summative Evaluation 17
   Quantitative and Qualitative Evaluation 18
      Quantitative Evaluation 18
      Qualitative Evaluation 19
   The Difference Between Evaluation and Research 21
   Summary 23
2. PSYCHOLOGICAL WARFARE:
      PSYCHOTHERAPEUTIC BIAS IN EVALUATION 25
   Is Anyone Out There Normal? 28
   Alternative Views of Psychotherapy 30
      Psychodynamics 31
      Behaviorism 34
      Humanism and Existentialism 37
   Another Look at the Elephant 40
   Theory-Driven Approaches to Evaluation 43
      Psychoanalytic Approaches 44
      Behavior Modification 46

Humanistic Psychotherapy 47
Summary 49
3. WHAT TO LOOK FOR IN EVALUATION PROCEDURES 51
Some Background: What's Intelligence and Who's Got It? 52
Validity 55
Content Validity 56
Construct Validity 58
Criterion Validity 60
Test Validation 63
Questions to Ask About a Procedure's Validity 69
Reliability 70
Reliability of the Test 71
Inter-Observer Reliability 74
Questions to Ask About Reliability 76
Practical Considerations in Choosing a Procedure 76
Utility 76
Directness 77
Nonreactivity 77
Questions to Ask About Practical Considerations 78
Summary 78
4. TRAPS, QUIRKS, AND FALLACIES IN EVALUATION 80
Problems of Insufficient or Biased Data 81
Accuracy is Affected by the Base Rate 81
Conditional Probability 82
The Unreliability of Small Samples 85
The Two Faces of Accuracy: Sensitivity and Specificity 88
Problems in Interpreting Data 88
Information Overload 88
The Bias Towards Overinterpretation 90
Believing is Seeing 92
Causes and Effects 94
Regression to the Mean 95
Questions About Questions 97
Whose Reality is More Real?: The Fallacy of Reification 100
Summary 102
5. THE MAGIC OF NUMBERS:
BASIC STATISTICAL CONCEPTS 104

The Ordering of Numbers 106
Nominal Scales 106
Ordinal Scales 107
Interval Scales 107
Ratio Scales 107
Cautions in the Use of Scales 108
The Frequency Distribution 108
Central Tendency 109
Reification of the Central Tendency 110
Variability 111
The Normal Curve 113
Proportions Under the Normal Curve 113
Derived Scores 114
Standardized Tests 116
Predicting From Statistics 117
The Standard Task 118
The Normative or Standardization Sample 119
Selecting a Standardization Sample 120
Standard Error 123
Correlation 127
Correlations to Measure Reliability 127
Other Correlation Coefficients 129
Correlation of Scores From Two Tests 130
Levels of Confidence 130
Summary 131
6. OBJECTIVE PERSONALITY TESTING 133
The Search for Typical Behavior 133
Performance or Situation Testing 134
Self-Report Questionnaires 135
Limitations of Self-Report Questionnaires 135
Approaches to Objective Self-Report Construction 137
Logical or Intuitive Test Construction 137
Criterion-Keyed Test Construction 140
Theory-Based Test Construction 145
Program Tests and Alternative Forms 150
The Test Manual as a Source of Information 151
Use and Application 152

Administration and Scoring 152
Technical Data 153
Summary 153
7. PROJECTIVE APPROACHES IN ASSESSMENT 155
The Projective Hypothesis 156
Projective Techniques 156
The Rorschach 157
Projective Drawings 162
Historical Development 163
Drawing Tests of Cognitive Ability 163
Figure Drawings as Tests of Personality 165
Auditory and Music Projective Tests 175
Projective Techniques in Dance/Movement Therapy 177
Psychodrama and Drama Therapy 177
How Valid are Projective Techniques? 178
Hypotheses About Figure Drawings 179
Can Projective Techniques be Tests? 181
Summary 182
8. BEHAVIORAL ASSESSMENT 184
Identifying Target Behaviors 186
Direct Behavioral Assessment 187
Frequency Measures 187
Interval Recording 189
Duration of Response 190
Self-Monitoring 190
Estimating Reliabilities for Direct Observation 191
Indirect and Analogue Approaches 192
Guided Imagery 192
Role Playing 192
Standardized Behavior Rating Scales 193
The Social Skills Rating System 194
Child Behavior Checklist 194
Cognitive Behavioral Assessment 195
Clinical Techniques for Recording Cognitions 196
Cognitive Behavioral Self-Reports and Rating Scales 197
Measuring Beliefs and Expectations 201
Psychophysiological Measurement 201

Problems of Reliability and Validity in Behavioral Assessment 204
Establishing Goals and Objectives 205
    Distinguishing Between Goals and Objectives 206
Summary 211
9. CLINICAL ASSESSMENT 213
The Meaning of Clinical Assessment 214
Clinical Judgment 215
    Alternative Definitions of Intuition 215
    Subjectivity and Bias in Clinical Judgment 217
Clinical and Statistical Approaches to Assessment 220
    The Search for Meaning 221
    Clinical Versus Statistical Accuracy 223
The Uneasy Relationship Between Clinicians and Researchers 228
    The Enthusiastic Clinician and the Skeptical Researcher 228
    Can the Gap be Bridged? 231
The Assessment Interview 233
    Purposes of the Assessment Interview 233
    Structure of the Interview 234
Summary 236
10. ISSUES IN ARTS THERAPIES EVALUATION 238
Why Arts-Based Assessments? 238
Art *In* Therapy or Art *As* Therapy? 239
Professional Identity and Levels of Assessment 244
    Diversionary Programs 244
    Clinical Activity Programs 245
    Arts Psychotherapy 248
Categories of Assessment Procedures 250
    Client Population 250
    Area of Functioning 251
    Theory/Model 251
    Technique 252
Who Administers the Arts Therapy Evaluation? 252
    Assessment by Specialists or Generalists 253
Cautions in Relying on External Evaluations 254
Psychopharmacology and Arts Therapy Assessment 256
Summary 258
11. VISUAL ART THERAPY 260

Psychodiagnosis and Personality Assessment 261
The Search for Meaning 261
Signs and Signals as Symptoms 265
Projective Drawing Tests 265
Art Therapists and Assessment Procedures 266
Attempts to Standardize Art Therapy Diagnosis 269
The Formal Elements Art Therapy Rating Scale 269
The Diagnostic Drawing Series 273
Is Patients' Artwork Enough for Assessment? 276
Family and Couple Drawing Assessments 277
The Kwiatkowska System 277
Landgarten's Family Art Psychotherapy Assessment 278
Mutual Drawings by Couples 279
Cognitive and Developmental Evaluation 280
The Goodenough-Harris Drawing Test 280
The Silver Drawing Test of Cognitive Skills and Adjustment 280
The Levick Emotional and Cognitive Art Therapy Assessment 281
Developmental Art Therapy Evaluation 282
Summary 283
12. EVALUATION IN DANCE/MOVEMENT THERAPY 286
Describing and Recording Movement 287
Effort/Shape Movement Analysis 288
Other Coding Systems 291
Personality Assessment 292
Marion North's Personality Assessment Through Movement 292
The Kestenberg Movement Profile 294
Espenak's Movement Diagnosis Tests 297
Movement and Jungian Personality Types 298
Psychiatric Diagnosis 299
The Movement Psychodiagnostic Inventory 299
Body Movement Scale for Autistic and Other Atypical Children 304
The Search for a Simplified Diagnostic System 308
Problems in Amassing a Data Base for Diagnosis 309

Functional Movement Assessment 309
Assessment for Movement Restructuring 309
Assessment for Rehabilitation 310
Groups and Activities Programs 311
Activity Groups 311
The Tameside and Glossop Screening Procedure 313
Family Systems Assessments 314
Dulicai's Nonverbal Family Assessment 314
Summary 317
13. EVALUATION IN MUSIC THERAPY 319
The Behavioral Bias in Music Therapy 320
Music-Based Assessments of General Status and Functioning 321
The Search for Music-Based Global Assessments 322
Music-Based Projective Tests 322
Music-Based Tests of Function and Ability 324
Assessment for Treatment Planning 327
Needs Assessments 327
Goals and Objectives 330
Summary 332
Bibliography 335
Index 369

THE ART AND SCIENCE OF EVALUATION IN THE ARTS THERAPIES

Chapter 1

PURPOSES OF EVALUATION

Especially in the arts therapies, many practitioners are often uncomfortable with the notion of evaluation. Some are unsure of the processes involved. Some assume that evaluation must involve the kind of counting and measuring that many of them find distasteful, mysterious, frightening, or even abhorrent. Others may believe that the programs in which they toil have such intrinsic and self-evident value that formal evaluation is not necessary. Some may be reluctant to put their work on the line by subjecting the results to scrutiny. Others may bristle at the notion of evaluation which they perceive to be a questioning of the value of their work. Many may be so involved in the day-to-day work of therapy that they lose sight of the reason for their labors.

As a result, until the last few decades, arts therapists have been slow in developing evaluation methods and procedures. However, several developments have spurred arts therapists to generate plans for assessing the needs of their patients and clients and for evaluating the results of their efforts.

One development is the increasing demand by third-party payers that claims for services specify the diagnosis of the patient or client. Arts therapists in private practice who submit claims for treatment to insurance companies, Medicare, and other third-party payers are usually required to use the diagnostic labels developed by the American Psychiatric Association or to ride the coattails of physicians who may prescribe arts therapies for physical or rehabilitative services.

A second development, affecting principally those arts therapists who work in hospitals, was the publication by the Joint Commission on Accreditation of Hospitals, in 1981, of the *Consolidated Standards Manual*, with new requirements for psychiatric, alcoholism, and drug abuse facilities. In addition to the medical or psychiatric diagnosis to be provided by the institution itself, "activity services" in such facilities were obliged to assess "the patient's needs, interests, life experiences, capacities, and deficiencies" (1981, p. 126). For the most part, arts therapists in such institutions are usually free to develop any "activity" assessment that satisfies the hospital administration. Such assessment procedures run the gamut from diagnosis to the identifica-

tion of learning preferences, personal abilities, social skills, and social needs.

A third is the result of the expansion of the arts therapies beyond their original base in psychiatric hospital wards into schools, community programs, and "wellness centers." Increasing numbers of arts therapists have been drawn into school systems as a result of the enactment of Public Law 94-142 in 1975. This law, Individuals with Disabilities Education Act (IDEA), mandated the establishment of programs to serve the needs of children with physical, development, or emotional problems. Schools, therefore, were required to develop a host of services addressed to the problems of exceptional children. In such settings, arts therapists may be involved in special education programs as teachers, therapists, or consultants. In these programs, educators are required by law to ascertain the developmental, physical, neurological or emotional problems of exceptional children, to identify their deficiencies and handicaps, and to develop individual educational plans (IEPs) designed to remediate or ameliorate these deficiencies.

There have been major problems in developing such assessment programs. Frequently, there may be no generally accepted criteria for ascertaining success or failure of a program of therapy; indeed, some therapists insist that success and failure of a therapeutic program or course of treatment are impossible to ascertain. Moreover, there is a good deal of uncertainty among arts therapists over the very nature of the emerging professions. Are these new endeavors *expressive* therapies, taking their cues from the older practices of psychotherapy? Or are they *creative* therapies that need to develop independent goals and criteria? Are their constituents patients–sick people who need medical attention–or are they clients, the myriad of those who need to be educated in coping with the overwhelming problems of daily living? Or are they essentially normal human beings who need guidance in sorting out their own identities, those who are often included in the category of "the worried well"? (In this book, we shall use the terms *patient* and *client* more or less interchangeably, although they represent different approaches to the people with whom arts therapists work.)

The approaches to evaluation are as varied as the philosophical approaches to the arts therapies. But, regardless of approach, the problem is the same: without some meaningful criteria for evaluation, we have no way of knowing whether a patient or client is receiving treatment (or training) that is appropriate for his or her problem; whether the treatment is helping, or has helped, the client; whether a therapist should augment, abandon, or change a method or an approach; whether a program is doing what it was set up to do; and whether it should be maintained or modified or abandoned.

There is no single best way to evaluate. Evaluation may be formal or informal, based on statistics or on intuition. Information may be gathered through the use of tests and measurement scales, through observation of

patient/client behavior or by asking patients about their thoughts and activities, through a qualitative assessment of a patient's drawing or movements or music-making, or through a convergence of impressionistic data. But, in terms of the definition of evaluation on which this book rests, they have a common denominator. Evaluation, for our purposes, is a method for collecting information on which to base decisions. And for some situations and for some purposes, some forms of evaluation are far more appropriate than others. Much of the skill of the evaluator rests on knowing the difference.

FUNCTIONS OF EVALUATION

There are five basic functions of evaluation:

(1) to ascertain the problems and needs of a person (a patient/client or a staff member), a program, or an institution;
(2) to predict future behavior;
(3) to monitor change;
(4) to know when to stop; and
(5) to learn how to improve treatment methods or techniques.

These functions are not independent and mutually exclusive. For example, without a baseline to establish the patient's status and need, monitoring is useless, since there is no way of knowing what change has taken place. Unless a therapist can predict a patient's behavior with some accuracy, there is no way to monitor change in any meaningful way, or to know when to terminate treatment, since any change that is noted might have been the result of the treatment, or it might simply have been a reflection of the natural course of the disorder.

All of these functions serve one fundamental purpose: to guide and direct treatment. Any evaluation program that does not contribute to planning or improving treatment, any program whose purpose is merely to label patients or to pigeonhole them, serves no useful purpose and may actually do considerable harm.

In addition, an assessment or evaluation is likely to be of little value unless the evaluator has a clear idea of why he or she is undertaking it. Not too many years ago, it was common for a psychiatrist to ask a staff psychologist to "do a psychological workup on Mr. Smith," leaving the psychologist to guess the hidden agenda. Was the purpose to validate a diagnosis? To choose an appropriate course of treatment? To identify likely problems, cognitive styles, capacity for insight, affective level, or potential resistances, in order to decide on an assignment to an activity? To identify problems that were likely to arise in the course of treatment? To find how much change

might have taken place during treatment? To probe Mr. Smith's suicide risk, or level of danger to others? To understand why a specific treatment is not working? Or to decide if a patient is ready for discharge from a hospital? Each purpose suggests a particular direction for the assessment procedure or perhaps even a different procedure. While such shotgun assessments are less common than they once were, they are still the rule of thumb in too many institutions. In addition, in many clinics or hospitals, it may be standard procedure to administer batteries of tests to every patient without a clear idea of the information that may be needed. The result may be that a good deal of information may be amassed that is irrelevant for the question or decision at hand, and much needed information may not be available to answer the question under consideration or to guide the decision.

The terms *evaluation* and *assessment* are often used interchangeably, because both refer to value judgments that are used as a basis for decision making. However, different groups, agencies, and institutions distinguish between the terms in a variety of ways. Some use the term assessment to describe the entire process of identifying a patient's or client's problems, and for monitoring his or her progress, while evaluation examines the efficacy of the treatment program. In most clinics and hospitals, however, assessment refers only to the initial determination of the patient's problems or needs, while evaluation describes the dual processes of monitoring patient progress and making judgments about the course of treatment.

DIAGNOSIS AND ASSESSMENT

In a medical setting, the initial assessment of a patient's problem is usually referred to as *diagnosis*, derived from the Greek words meaning "to separate" and "to know." It is the act of recognizing a disease by distinguishing it from others; in modern medicine, this process is sometimes referred to as *differential diagnosis.* The diagnostic function in psychotherapy rests squarely on the medical, or illness, model of mental distress and is often referred to as *psychodiagnosis.*

Until recent years, while psychiatrists and psychotherapists spoke of psychodiagnosis, they did not practice the differential diagnosis that is traditional in biological medicine: distinguishing between diseases that might have the same or similar symptoms. Psychiatric diagnosis was based almost exclusively on symptomology, while psychological testing was focused largely on ability or personality assessment, finding how close someone was to the norm, or average, rather than on diagnosis. As we shall see, this situation has changed dramatically since the 1950s.

While the term diagnosis is frequently used in the arts therapies, few arts

therapists are actually involved in ascertaining the psychiatric label to be applied to a patient. Over the years, however, there have always been those who have pursued the development of arts-based diagnostic procedures.

Diagnosis as a Guide to Decisionmaking

Years ago, the diagnosis of patients was not as critical as it is today in identifying mental disturbances, mainly because there were few treatment options. During the Middle Ages and the Renaissance, mental illness was attributed to possession by demons (Zilboorg, 1941) or was perceived as retribution for sin. The standard treatment consisted of exorcism or prayer. Court records dating as far back as the thirteenth century indicate that judges used mental status examinations to distinguish between the mentally retarded or "natural fools" and the mentally ill, or "lunatics" (Neugebauer, 1979, p. 481). The reason for the distinction, apparently, was to identify lunatics who might pose a danger to the community, and who were locked out of sight, with no real attempt at therapy (from the Greek words for "to nurse" or "to cure").

With the advent of psychoanalysis, the "talking cure" constituted both diagnosis and treatment. Fundamentally, all patients were offered the same treatment, in the course of which their particular problems would emerge.

It has only been in the last few decades, with the remarkable growth in the use of psychiatric drugs, that the critical importance of diagnosis has become apparent. While such drugs as opium and morphine had been used for over a century in mental institutions, they were used almost exclusively to sedate patients, to keep them quiet, docile, and manageable, rather than to treat them (Brandt, 1975, p. 39). Even in the 1950s, when new drugs were found that could control psychotic symptoms, the need for accurate diagnosis was not recognized for some time. When tranquilizers were introduced, they were prescribed for a wide variety of mental disorders, on the ground that the common denominator of most disorders was anxiety (Pruyser & Menninger, 1976, p. 26). In time, it became increasingly apparent that many of the drugs that were being developed were not only useless but could be harmful if they were prescribed for the wrong disorder, or even for the right disorder, but during the wrong phase of that disorder. It was learned that different drugs acted quite differently on disorders whose symptoms seemed similar, like schizophrenia and some phases of manic depression, and that there were even metabolic differences between patients with symptom (or reactive) depression and those with illness (or clinical) depression that caused them to react very differently to drugs (Ayd, 1976, p. 146).

While the movement for greater precision in diagnosis was inspired by the advent of the psychiatric drugs, a good deal of information has accumu-

lated during the last few decades that has driven the endeavor to hone the diagnostic process. Researchers have found that patients who exhibit similar symptoms may be suffering from very different problems, requiring different forms of treatment. Moreover, even victims of the same general disorder may respond very differently to a particular treatment. As a result, differential diagnosis has become a basic principle in choosing the appropriate therapeutic methods and procedures.

The Meaning of Diagnosis

The subject of diagnosis has engendered a good deal of controversy. First of all, the term itself is ambiguous, and it has several meanings. Herbert Modlin (1976), a psychiatrist at the Menninger Foundation, pointed out that psychiatric diagnosis may have several meanings.

A Nosological or a Classification Label

The term *nosology* is often defined as a taxonomy, a system for classifying diseases as a basis for diagnosis. Some writers contend that, while nosology and taxonomy are related, they refer to different concepts. Paul Pruyser and Karl Menninger (1976) state that a nosology has to do with the way we conceptualize a disease–as a biochemical imbalance, for example, or a psychological response to stress. "Nosologists want to know what a disease is," they say (p. 13). Nosology involves creating theoretical constructs for disorders, erecting boundaries and parameters to distinguish these concepts from others. The bases for these conceptualizations range from the philosophical to the empirical. At one time, it was thought that what we now call *hebephrenia* (inappropriately "silly" behavior), *catatonia* (extreme changes in muscular tension), and *paranoia* (delusions centered on suspiciousness) were distinct disorders. Around the turn of the century, the German psychiatrist Emil Kraepelin recognized from his clinical observations that the three were more similar than they were different and that, in fact, victims could manifest symptoms of one or the other at different times. He conceived the nosological concept of a single disease that he called *dementia praecox*, meaning an early (precocious) form of dementia, or mental deterioration, to differentiate it from *senile dementia.* The Swiss psychiatrist Eugen Bleuler coined the name *schizophrenia* (splitting of mental functions) for the disorder (1911), from the Greek words for "division" and "mind", and he added a fourth subtype, *simple schizophrenia*, characterized by "negative" or "non-psychotic" symptoms (inattentiveness, flattening of affect, loss of appetite).

Therefore, nosology refers to the concept of the disease in terms of its attributes or characteristics: the way psychiatrists think of it. William

Greisinger, often linked with his maxim, "mental diseases are brain diseases" (Lunbeck, 1994, p. 118) conceptualized mental illness as a single psychosis with stages from melancholia to delusionary madness to dementia, and clinicians perceived each disorder as a distinct phase of the same organic problem (Pruyser & Menninger, 1976, p. 124).

A taxonomy or classification, in this view, refers to the ways psychiatrists organize the categories of the disorders and, usually, the diagnostic criteria or symptoms by which each may be recognized. *The Diagnostic and Statistical Manual* (DSM) of the American Psychiatric Association is a taxonomy, or classification system. The DSM provides the names (and numbers) of categories and subcategories by which psychiatrists and therapists identify mental disorders, as well as the diagnostic criteria for each.

The description of the symptoms on which the classification system rests often reveals the underlying nosological constructs. For example, the revisions of the DSM in the 1970s and 1980s defined schizophrenia largely in terms of "positive" symptoms, such as hallucinations and delusions; the 1994 revision reintroduced Bleuler's and Kraepelin's notion of "negative" or "nonpsychotic" symptoms, such as loss of drive, loss of ability to experience pleasure, and loss of emotional expression (Andreasen, 1994, pp. 345-346). While both the American DSM and the World Health Organization's *International Category of Diseases* (ICD) may use the same names for mental disorders, the lists of symptoms sometimes reveal differences in underlying concepts.

The classification label is simply a name that is used to describe a group of observable behaviors. Unfortunately, the labels don't help to predict the course of the disease. Some catatonics recover completely; others may be institutionalized for most of their lives. So additional categories have been established to differentiate between patients who suffer "reactive" forms of the problem (who are likely to recover) and those who suffer "process" or chronic forms of the disease (whose prospects are dismal).

The more recent DSMs do not suggest either the cause of the problem or the indicated treatment. Determination of the causes or etiology and decisions about indicated treatment are left to individual practitioners, for reasons that should become abundantly clear in the next few chapter.

The labels, too, aside from psychopharmacology, don't do much to offer guidance in the choice of treatment. What does a therapist who faces a patient diagnosed as manic depressive know about the patient's suicide risk, his or her cognitive style, or affective level, or capacity for insight, or preferred activity?

A Point of View

The diagnosis that suggests the etiology, or causes, of a problem, will inevitably reflect the assumptions that the diagnostician makes about the origin of disorders. "The obscurity of etiology in mental illness," says Modlin, "vivifies several unverified hypotheses, such as organic, neurochemical, psychodynamic, behavioral, interaction, and social explanations for our patient's deviations from theoretically constructed norms" (p. 153). Points of view may be influenced not only by theoretical allegiances but by the country or culture in which the psychiatrist practices. Some disorders that are diagnosed in the United Kingdom as neuroses may be diagnosed by American psychiatrists as psychoses (Modlin, p. 153).

The Diagnostic Babel

Compounding the ambiguities shrouding the nature of diagnosis are the confusions surrounding the language of psychiatric diagnosis. The classification systems and the names of the disorders themselves are riddled with inconsistencies. There is no single unified conceptual scheme for organizing or naming disorders. "Some disorders are grouped as mood disorders (depression, mania), others are called adjustment reactions; some are ideational aberrations (paranoia, obsessional neurosis), others are limited functional failures of memory (dissociative reaction); some derive from a historical and quite holistic view of personality or character (borderline condition, narcissistic personality) hardly conducive to precise symptom descriptions, others take their cue from a circumscribed habit (alcoholism, fetishism) or a 'special symptom,' as the DSM. . .calls it, such as enuresis or a speech defect" (Pruyser & Menninger, 1976, p. 16).

The constructs on which diagnostic labels are based may change with time, and constructs are formulated to fit theories. When the presence of fever was the distinguishing sign of a disorder, Hippocrates distinguished between *phrenitis* and *mania*. When sexual pathology was considered paramount as a cause of mental disorder, psychiatrists conceptualized such specific diseases as *hysteria*, *satyriasis*, *neurasthenia*, and *psychothenia* (Pruyser & Menninger, 1976, pp. 13-14). And when *ante-bellum* American physicians addressed the "maladies of the Negro race," they discovered the medical answers to recurrent problems among slaves in such diseases as *drapetomania* (from the Greek words for "runaway" and "madness"), and *dysaesthesia aethiopis*, an affliction characterized by disruptive behavior, usually accompanied by an "obviously pathological" change in the functioning of the nervous system that made the victim insensitive to pain when being punished (Chorover, 1973, p. 44).

Moreover, old diagnostic terms tend to persist long after the constructs on which they are based have been abandoned. The term *neurosis*, referring to a neurological disorder, is still in common use, although it was dropped from the DSM in the 1980s.

In addition, diagnosticians will often "describe" the disorder they have diagnosed in terms of their own theoretical affiliations, in effect projecting their biases. Psychoanalysts will use as descriptive terms words that are actually interpretations: transference, resistance, affect organization; biopsychiatrists will talk of bipolar depression and dyskinetic factors; those who subscribe to group therapies will describe splitting, loyalty and role; and gestalt therapists will talk about self, speaking out, and masking (Pruyser & Menninger, p. 19). Consider the following description by a psychoanalytically oriented movement therapist: "A couple entered my office. The husband smiled and sat down in a bulging forward, bipolar widened, dimensional side out, body attitude. . . . The wife sat in a vertically lengthened cross-armed and legged position. . . . She had projected her borderline enmeshed mother on him and was doing everything possible to encourage more differentiating oral- and anal-sadistic discharge in her spouse. . . . He, meanwhile, utilizing flexibility with an oral inner genital rhythm would try to 'understand,' for he had split off his aggression due to an early childhood trauma" (Lewis, 1990, p. 73).

Some authorities believe that diagnostic labels should do nothing more than describe the disorder (like phobias) and avoid attempts to explain causes or to suggest treatment, because the understanding of the problem may change and because of the large possibility of misdiagnosis. Asher (1972) offers illustrations of diagnoses which provided false information about etiology, or causes, and led to inappropriate treatment.

Matarazzo and Pankratz (1984) agree that many diagnoses contain premature or misleading information about causes. Most conditions, they contend, are more complex than was once imagined, and single model diagnoses may be misleading and are generally insufficient. Because of the complexity of disorders, there has been the rise of multidimensional approaches, like the "biopsychological" model and behavioral medicine (p. 372). In recognition of the many-sided nature of most mental problems, the more recent DSMs provide for multiaxial diagnoses that touch upon various manifestations of a problem.

How Useful Is Diagnosis?

The notion of diagnostic labels is not accepted universally. Many critics point out that psychiatric "symptoms" are actually interpretations based on the observations by the therapist and on reports from the patient. These

symptoms are subjective in nature; there are few, if any, truly objective tests to verify the existence of a disease or the degree of dysfunction, as there are in physical medicine. The reality of practice and the frequent difficulties in accurate diagnosis have raised doubts in the minds of many clinicians about the validity of the list of neatly bounded categories, each encompassing a distinct disease.

Many critics, including some who accept the illness model, believe that diagnosis is irrelevant, and they prefer to use the broader term *assessment*, by which they mean the process of determining an individual patient's needs. In behavioral theory, global diagnoses have no value, because to behaviorists the behavior itself is the problem.

Some humanists object to diagnostic categories on the ground that such categories strip the individual of his or her individuality. Abraham Maslow (1966), a founder of the humanistic movement, wrote: "I must approach a person as an individual unique and peculiar, the sole member of his class" (p. 10). The problem with ignoring the characteristics that the individual shares with others, respond the diagnosticians, is to recapitulate the errors of the past and to ignore the lessons of experience. "Would we not be totally ignorant of how to help each new patient," ask Shevrin and Schechtman (1973) "if all previous ones were also unique?" (p. 463).

In fact, contend Pruyser and Menninger (1976), "diagnosis is not only a necessary psychiatric activity, but by far the most important single *raison d'etre* for psychiatry as a profession." There are many who are eager to help the afflicted and who offer interventions, they say, "But unfortunately many interventions are made available and tried out without the least concern for their fitness to the condition to which they are being applied" (p. 25). The major function of diagnosis is to match a defined condition with an appropriate treatment. "Unless the diagnostic process facilitates treatment," says Modlin, "it is of little worth" (p. 157). As the menu of treatment choices has expanded in recent years, the numbers of categorical classifications has kept pace, and each category has been subdivided into increasingly narrow subcategories.

On Being Sane in Insane Places

The debate over diagnosis in psychiatry flared dramatically in the early 1970s with the publication of a report by D. L. Rosenhan (1973) in the journal *Science*, entitled, "On Being Sane in Insane Places". Rosenhan, a professor of psychology and law at Stanford University, sent eight pseudopatients to twelve psychiatric hospitals, where they gained admission on the basis of a complaint that they had heard voices for a period of three weeks.

All the applicants were admitted. At this point, they immediately ceased

simulating any symptoms of abnormality, and behaved as they normally did. When asked by attendants, they reported that they no longer experienced symptoms.

What Rosenhan found remarkable is that, while other patients often recognized the normality of the pseudopatients, the staff never did. Rosenhan wrote:

> Failure to detect sanity during the course of hospitalization may be due to the fact that physicians operate with a strong bias toward what statisticians call the type 2 error. This is to say that physicians are more inclined to call a healthy person sick (a false positive, type 2) than a sick person healthy (a false negative, type 1). The reasons for this are not hard to find: it is clearly more dangerous to misdiagnose illness than health. (P. 252.)

Beyond the tendency to call healthy people sick, wrote Rosenhan, "the data speak to the massive role of labeling" (p. 252). Labeling in psychiatry, he charged, carries a stigma that does not obtain in medicine; most medical illnesses are not pejorative. But psychiatric diagnoses carry personal, social and legal stigmas. "The tag colors others' perceptions of [the individual] and his behavior," Rosenhan wrote. "Once a person is designated abnormal, all of his other behaviors and characteristics are colored by that label" (p. 253).

Rosenhan was curious to see if misdiagnosis could occur the other way—that is, if the disordered would be diagnosed as normal. He informed the staff at a research and teaching hospital that at some time within the next three months, he would send one or more pseudopatients who would attempt to be admitted to the psychiatric hospital. Each staff member was asked to identify pseudopatients. The results: of 193 patients who were admitted for psychiatric treatment during the next three months, 41 were alleged, with high confidence, by at least one member of the staff, to be pseudopatients. Of these, 19 were suspected by at least one psychiatrist and one other staff member. Actually, not a single pseudopatient had been sent!

Rosenhan criticized the hospital personnel for not considering differential diagnoses, as would be done routinely in physical medicine, and he contended that the chances of misdiagnosis are great in psychiatry. Moreover, in reviewing the case summaries, he found that normal family histories were distorted and reinterpreted, probably unintentionally, to make them fit into a theoretical mold. In other words, the diagnostic labels not only became self-fulfilling prophecies of the way patients would behave but they colored the therapist's perception of what had happened in the past. Rosenhan concluded that such labels serve no useful purpose and do more harm than good. "We have known for a long time," he asserted, "that diagnoses are often not useful or reliable, but we have nevertheless continued to use them. We now know that we cannot distinguish sanity from insanity" (p. 257).

As was to be expected, Rosenhan's report was received with a combina-

tion of embarrassment and anger in the psychiatric community. However, it also stimulated a good deal of thoughtful self-examination among responsible psychiatrists. Robert L. Spitzer, later to head up the American Psychiatric Association's task force that revised the DSM, pointed out (1975) that admitting staff members are trained to assess symptoms, not to identify fraud. He identified a number of flaws in Rosenhan's study, and argued that "the clinical picture includes not only the symptom (auditory hallucinations) but also the desire to enter a psychiatric hospital, from which it is reasonable to conclude that the symptom is a source of significant distress" (p. 446). While he admitted that reliability in the diagnosis of mental disturbance ranged from barely satisfactory for alcoholism and organic brain syndrome to poor for most categories, he argued that psychiatry is not the only branch of medicine plagued by inaccurate diagnosis and poor interpretations of diagnostic data. However, he did concede the prevalence of a major diagnostic distortion that Rosenhan had noted: the revision of historical facts to achieve consistency with psychodynamic theories. He wrote in the *Journal of Abnormal Psychology* (1975):

> Here, for the first time, I believe Rosenhan has hit the mark. What he described happens all the time and often makes attendance at clinical case conferences extremely painful, especially for those with a logical mind and a research orientation. (P. 448.)

When Dr. Spitzer presided over the revision of the DSM, he labored to tighten the organization's definitions and symptomology, and he went to great pains to insist on the use of specific inclusion and exclusion criteria in the interpretation of symptoms. This move towards greater differential diagnostic precision has continued in the more recent revisions.

Fredrick Shechtman, Core Psychologist of Diagnostic Services at the Menninger Foundation, admitted to the validity of Rosenhan's complaint about the diagnostic procedure at the hospitals involved (1976), but he asked whether the pseudopatients were really diagnosed, or simply labeled. A psychiatric label should serve as a beginning, not an end point in diagnosis. Rosenhan was right in criticizing poor diagnostic practice, Shechtman said, but he was wrong in condemning the whole enterprise of diagnosis because some practitioners abused the process (pp. 43-44).

Since the controversy over Rosenhan's study, the whole issue of diagnosis may come to represent a diminishing concern to therapists. It has been estimated that nearly half of those who come to therapists for treatment "do not meet the diagnostic criteria for any defined mental disorder and that a portion of them suffer only from 'problems in living'–the annoyances and anxieties thrown up by day-to-day life over which . . . psychiatrists had attempted to assert their disciplinary authority" (Lunbeck, p. 309). What this suggests is one of two possibilities. Either increasing numbers of therapists

will find the illness model inapplicable to their practice with essentially "normal" clients, or the psychiatric establishment will continue to expand its classification system to redefine such problems as manifestations of underlying disorders, thereby turning "problems-in-living" into diseases. There are signs that both trends are actually evolving simultaneously. The relative development of the two is likely to be heavily influenced by what third-party payers are willing to underwrite as medical expenses.

Psychiatric Labels and Arts Therapists

There is a wide range of opinion among arts therapists on the utility of the psychiatric label.

Arts therapists who have been trained in hospitals or who work in hospitals are affected by psychiatric labels in a number of ways.

First of all, much of the treatment in the hospital is based on the diagnostic label, which is usually selected from the DSM menu: patients are assigned to wards, to activities, to methods of treatment largely on the basis of the label assigned by staff members who have conducted the intake assessment in the spirit of the medical model: interpreting symptoms in terms of an underlying disease.

The diagnostic label will be the basis for determining which medications will be administered, and the arts therapist must be familiar with both the intended effects and the incidental or unintended side effects of the various medications on patients in planning therapeutic interventions.

Moreover, arts therapists themselves may well be involved in assigning the psychiatric label. Their own assessments of the behavior of patients sometimes contributes to the staff's decision. To this end, many movement therapists, art therapists, and music therapists interpret the ways patients behave in their studios in psychiatric terms; the body dysfunctions, the distortions in patient drawings, the arrythmicity that are observed by arts therapists are viewed as specific manifestations of psychiatrically classified disorders. Some research has been conducted in seeking correlations between the ways people move, or draw, or produce or respond to rhythm, and the various psychiatric diagnostic categories.

The application of psychiatric labels in arts therapies assessment has not met with universal approval. Many arts therapists contend that such labels are of limited utility in guiding the treatment of patients in their disciplines. Some resist what they consider the subordination of the arts therapies to the verbal therapies, making the arts therapies "handmaidens of psychiatry." To some arts therapists, the dysfunctions they observe, even if they parallel psychiatrically defined symptoms, are expressions of problems that can be identified and treated without knowing the psychiatric label.

Some object to the dichotomous nature of diagnostic labels; someone has a disorder or doesn't have it. Kanfer and Saslow (1969) suggest that the difference between "normal" and "abnormal" behavior is one of degree, not of kind. Someone who exhibits "compulsive-obsessive" behavior, for example, usually does what all of us do–but to excess. A painfully shy person has the fears and concerns that we all have–but exhibits insufficient assertiveness to overcome them. So Kanfer and Saslow's "functional" approach ignores the diagnostic label and identifies behavioral excesses and deficits. The treatment, then, is directed toward having the patient or client increase the behavior in which he or she is deficient, and decrease the behavior in which an excess is identified.

On a practical level, arts therapists working in hospitals must conduct their own assessment of a patient's needs in planning their own therapeutic interventions, for which the psychiatric labels may offer limited guidance at best. For example, a music therapist who faces a new patient diagnosed as schizophrenic must decide which activities or programs are most appropriate: instrumental group improvisation for reality orientation to address the patient's delusional thinking; assignment to an instrumental performance ensemble as a mnemonic device to deal with the patient's impaired memory; guided music listening to evoke feeling responses as a way of addressing the patient's flattened affect; or the use of music activities to provide themes for later verbal psychotherapy. In fact, in many institutions, the arts therapist may not even be informed of the psychiatric diagnosis but only of the "symptom," such as the patient's need for increased socialization or inability to communicate coherently.

Arts therapists in private practice or in non-medical settings constitute a far more diverse group. Some consider themselves "arts psychotherapists," who may use the art experiences mainly as sources for the assignment of psychiatric or personality labels. They may elicit images, feelings, or behavior in the artistic media as a basis for verbal therapy. Others may practice "creative," "catalytic," or "process" arts therapies in which the art, movement or music experience itself constitutes both the raw materials for assessment and the therapy. In the view of those who practice "creative" therapies rather than "expressive" therapies, the behavioral manifestation of a problem can be identified and treated directly.

Still another group of arts therapists remains disengaged in the controversy over diagnostic labels. Many of those who consider their programs "activity services" rather than arts psychotherapies subordinate or simply ignore the psychiatric diagnostic labels. The *Music/Activity Therapy Intake Assessment* developed at Loyola University, for example, concentrates on skills areas within the music therapy program and deals with such areas as activity preference, organizational involvement, attitude survey, and post-

interview observations (Braswell et al., 1983). Still others, disdaining "a formula approach" to therapy, seek an art-oriented "glimpse of the inner life of the patient-in-crisis" (Moon, 1992, p. 138), looking for clues in the patient's preference for and choice of media, procedures, tools, postures and verbalizations as guides in determining the patient's needs.

Moreover, considerable numbers of arts therapists deal with clients and patients whose problems have little to do with traditional "disease" categories of mental disorders as they are normally conceived. In recent decades, arts therapists have rediscovered some of the ancient uses of the medicinal muses and have applied them in new ways. While minstrels may no longer play the harp softly in the boudoirs of ladies suffering from melancholia, hospitals are piping music into delivery rooms, operating rooms and recovery rooms to ameliorate trauma and expedite healing. And while musicians may no longer play *tarantellas* so that afflicted persons can dance off the venom of the tarantula, music and dance therapists help clients control stress and anxiety. Increasing numbers of arts therapists work with medical patients recovering from injuries or disease, with those suffering from developmental disabilities, and with those who need assistance in working out their relationship problems.

Therapists who deal with mental retardation, with victims of perceptual problems, with the physically handicapped, or with behavioral problems, such as drug abuse, may see little value in psychiatric labels. For one thing, many of these therapists may view their work as applications of education or training rather than of psychotherapy. When such therapists use the term *diagnosis*, they tend to refer to the identification of their patients' needs, or the deficiencies for which they may develop training schedules or individual educational plans.

Monitoring and Summative Evaluation

Once the assessment is made and treatment begins, the process of evaluation has not ended. In 1967, Michael Scriven, an authority on curriculum evaluation, distinguished between the kind of ongoing evaluation that he called *formative* and the kind that is conducted at the end of a program, which he called *summative*.

Formative evaluation is a continuing process. It provides information on changes in a patient's condition, so that the clinician can make adjustments in the treatment program until the treatment is to be terminated.

It is not easy for a clinician to know when to stop testing or to stop treatment. The basic purpose of evaluation is to help us make decisions, and there are times to recognize when additional information will not affect treatment. By the same token, there are times to recognize that there is little or

nothing more to be gained by further treatment. The decision to terminate treatment may be one of the most difficult for a therapist to make; sometimes, it is hard to know when he or she has done as much as is reasonable or even possible.

When the clinician considers whether the patient is ready to terminate the treatment or to be discharged from an institution, evaluation questions will probably include: Has the patient improved about as much as he or she is likely to improve? Have the objectives of the treatment program been achieved? Does the patient consider that he or she is ready to leave therapy? Is it safe for the patient to be discharged, both for the patient and for the community?

In addition, a major function of summative evaluation is to help the clinician learn from the experience. The end of treatment should be an opportunity to review the experience and to ask appropriate questions: What methods and techniques were particularly useful for this patient or for this group of patients, and which seemed to be ineffective? What should be done differently the next time? How useful were the diagnostic or assessment procedures in shaping treatment?

Many of the evaluation methods and techniques discussed in this book can be used either for formative or summative evaluation. What distinguishes them is the purpose for which they are employed.

QUANTITATIVE AND QUALITATIVE EVALUATION

We must be careful to avoid viewing quantitative and qualitative evaluation as antagonistic. Writing in the *HIV/AIDS Newsletter* of the Center for Disease Control, Assistant Surgeon General Gary R. Noble (1991) wrote:

> The best way to plan for evaluation is to establish programs that incorporate specific, measurable goals. Quantitative evaluation can then tell us *what* effect we are having (how much, where, who, when); qualitative evaluation can tell us *why* the program is effective or why there is a problem. Both are equally important. (P. 2.)

Quantitative Evaluation

When we speak of quantitative evaluation, we are really referring to two distinct processes: measurement and evaluation.

Measurement is a quantitative description of a behavior or thing. When we measure something, we compare it against a standard: an inch, a degree, a pound or a meter. Or we measure the frequency with which something

happens. The measurement, often gathered through tests and expressed in a number, is useful for describing something, but in itself, it tells us little about whether the thing being measured is big or small, cold or warm, heavy or light, healthy or unhealthy, or good or bad. Measurement is usually considered objective, in that the standards are fixed and there is little personal judgment involved in the process, although, of course, there is a good deal of personal judgment involved in deciding what to measure and how to measure it. The bias in measurement is far less obvious than it is in qualitative description.

Measurement involves a particular way of collecting data on which to base either research findings or value judgments. In itself, it is neither research nor evaluation.

Evaluation, and *assessment*, by definition, involve value judgments. Again, we usually need standards for comparison, but these are usually clearly subjective and frequently they are fuzzy or elastic. Both DSM-I and DSM-II specified the symptoms that indicated the presence of a psychiatric disorder, but left it to the judgment of the psychiatrist to decide how many symptoms had to be present for a diagnosis. The task force preparing the revised DSM-III tried to reduce the ambiguity by specifying an arbitrary number in the criteria for diagnosis.

While assessment in the therapies often involves testing (measurement), there are clear distinctions between "testing" and "assessment," not only in terms of definition, but in terms of the therapist-client relationship. In his 1990 presidential address to the American Psychological Association, Joseph D. Matarazzo (1990) said:

> . . . objective psychological *testing* and clinically sanctioned and licensed psychological *assessment* are vastly different, even though assessment usually involves testing. . . psychological *testing* [is] an activity that has little or no continuing relationship or legally defined responsibility between examinee and examiner. Psychological *assessment*, however, is engaged in by a clinician and a patient in a one-to-one relationship and has statutorily defined or implied professional responsibilities. (P. 1000.)

Qualitative Evaluation

Any evaluation that does not involve measurement could be called qualitative. Used in this sense, it would include intuitive, impressionistic, and clinical techniques. However, in education and in some of the therapies, the term is often used to refer to a specific group of approaches, sometimes called "descriptive," "naturalistic," or "goal-free."

In contrast with quantitative description, qualitative information gathering is highly personal, since the investigator is the primary instrument. The

evaluator needs a considerable tolerance for ambiguity, because there are often no set procedures or protocols, nor is there any single "correct" way to proceed, and each next step flows from the evaluator's perception of what is unfolding and what needs to be clarified.

Whereas those who quantify strive to be as objective as possible, the very concept of objectivity is suspect to many who use qualitative methods. To many neo-Freudians, and to even more phenomenologist therapists, each individual is unique; subjectivity is the key to understanding how a subject perceives reality.

The evaluator, therefore, must be sensitive to the clues that reveal a subject's views and perceptions, or the function that is served by group behavior, or to the reasons why a program may not be operating as the clinician or the administrator had hoped. The evaluator must be sensitive, also, to the questions that arise as the procedure develops: What does the information tell you? Where is it suggesting you go? How can it direct you to the next question or observation?

Guba and Lincoln (1981) wrote that qualitative evaluators do not measure. They "do what anthropologists, social scientists, connoisseurs, critics, oral historians, novelists, essayists, and poets throughout the years have done. They emphasize, describe, judge, compare, portray, evoke images. . ." (p. 149).

Obviously, given the subjective nature of this approach, bias is inevitable. But, contend the proponents of qualitative evaluation, bias exists in every form of evaluation. It is simply more obvious in qualitative than in quantitative approaches, because it cannot be as easily concealed.

The fundamental protection against bias is recognition. "The best cure for biases," wrote Guba and Lincoln, is to be aware of "how they slant and shape what we hear, how they interface with our reproduction of the speaker's reality, and how they transfigure truth into falsity" (p. 148).

As the theorists of this approach describe qualitative evaluation, a number of principles emerge.

1. Human behavior is always bound to the context in which it occurs–historical, social, genetic, or environmental. Any attempt to divorce behavior from context is "context-stripping" (Mischler, 1979).
2. The search for meaning is constructed by the subject or inheres in the situation. It is not imposed by the evaluator.
3. The purpose of the evaluation is to focus on what is actually happening, rather than to see if the intent of the educational or therapeutic program is being achieved. Purpose is not just in the what, but the why and how, so the evaluation is interpretive and explanatory.
4. The collection of data and their analysis occur simultaneously; meanings emerge and change as more data are gathered. Consequently,

the evaluator must rely on hunches and working hypotheses that change as more data are gathered.

THE DIFFERENCE BETWEEN EVALUATION AND RESEARCH

The word "evaluation" is ambiguous. When Raymond Corsini was compiling the various editions of his *Current Psychotherapies*, he asked the contributing authors to deal with a number of topics, including evaluation. Fewer than half of the contributors dealt with judgments about patients; many interpreted the topic as an invitation to extol the effectiveness of their particular forms of therapy, and cited research findings and outcome studies to bolster their claims.

The term *evaluation* is sometimes used to describe a process of finding answers to general questions: Which of two methods of treatment is most effective in dealing with alcoholics? Can we identify similar components in different forms of therapy? Which types of therapy lend themselves most effectively to group situations? Can comparable results be obtained through short-term therapy as are achieved in an extended course of treatment? To what degree does personal rapport between patient and therapist affect the outcome of treatment? Are there significant outcome differences between therapies that combine nonverbal and verbal approaches and those that rely mainly on the nonverbal aspects?

Because we shall be referring to both research and evaluation, we prefer to use the terms with more precision. The methods of investigation in research and in evaluation may overlap; in some cases, they may be identical. Both may use tests and measurements, interviews, self-reports, behavioral observations, or checklists. In fact, the process of standardizing a test *must* involve research. What distinguishes the two is purpose.

Research involves a systematic inquiry that is designed to broaden our understanding of the subject under study. The approach is often based on an inductive logic, working from the particular to the general. From an examination of specific cases, we can develop generalizations that can be applied to large numbers of situations (in quantitative approaches, this process is referred to as *inferential statistics*). The validity of these generalizations is often established by applying deductive predictions that are verified by testing them in specific cases and under a variety of circumstances, particularly in the case of the exceptions that "prove" the rule (from the original meaning of *proof* as a test or trial).

Research findings in themselves are of limited value. Simply knowing that eight out of ten persons who are tense feel better after they play a vigorous game of racquetball may be useful information. But, unless we plan on playing racquetball each time we go to the dentist or take a test or open

a tax notice, we want to know *why* playing racquetball helps relieve tension. Developing a broader understanding of the subject under study is useful beyond the specifics. Research findings, therefore, are almost always accompanied by conclusions or interpretations that attempt to explain or account for the findings. The explanation itself is subjective, and its accuracy is usually tested and validated through replication or additional research.

Perhaps the outstanding characteristic of research, as distinguished from evaluation, is its *nomothetic* quality or "generalizability": the search for general statements that describe relationships between classes of phenomena. At the highest levels of abstraction, such statements will hold up over time and space, and will accurately describe large numbers of specific situations or cases. Such generalizations are often referred to as principles or, if they have held up over long periods of time and differentiated applications, as laws.

Examples of generalizations that develop from research may include the following:

- Art productions of schizophrenic patients often use single images to describe elaborate and complex sequences of ideas.
- Listening to music produces changes in blood pressure, changes in posture, pulse rate and general activity in schizophrenics, and measurable mood change in all listeners.
- Schizophrenics and autistic children tend to exhibit little of the self-synchrony between their own body parts, or the interactional synchrony with others that characterizes communication between normal individuals.

While the results of research may be used in the development of policy decisions and program designs, the researcher's fundamental purpose is to advance knowledge and understanding in the field, and not to come up with practical applications in specific cases. He or she is not likely to address the question of whether a particular method or technique is likely to help a particular patient.

Evaluation and *assessment*, on the other hand, refer to the process of gathering information on the basis of which we make specific decisions about specific programs for a particular patient or an identifiable group of patients or clients in a specific setting. In contrast with research, evaluation deals with the *ideographic*, or the particular. Whereas research culminates in *conclusions*, evaluation is the basis for *decisions*.

Although both concern the evaluator, he or she is usually more interested in knowing *whether* a program or a method works than in learning *why* it works. The function of evaluation is to provide information on which to base decisions: What treatment is most likely to help a particular patient? Is a