

CLINICAL JUDGMENT AND THE USE  
OF PSYCHOLOGICAL REPORTS

By

JAMES MARTIN SIWY

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL  
OF THE UNIVERSITY OF FLORIDA IN  
PARTIAL FULFILLMENT OF THE REQUIREMENTS  
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1984

## ACKNOWLEDGMENTS

This dissertation owes its existence in no small part to the encouragement and tutelage of Roger Blashfield, whose enthusiasm, suggestions and criticisms guided me through every step of the process. Roger's research acumen helped discern the essential from the nonessential; his patient nurturance sustained the belief that the project was not only worthwhile, but also possible.

In like manner, Lynn Robbins deserves great praise for her generous assistance with the data analysis. Her special effort in enabling a thorough, efficient analysis to be done was marked by personal encouragement and unselfish donation of her time and talent.

I would also like to extend my appreciation to fellow colleagues, June Sprock, Mark Gilbertson, Roberta Isleib, Rich Shapiro and Marna Cohen, for their contributions and helpful ideas in the development of this project. The thoughtful criticisms of Eileen Fennell, Mary McCaulley and Scott Miller assisted me in constructing a feasible research design. The wise advice

of Austin Creel was finally heeded and much appreciated for the completion of the manuscript.

Special recognition goes to Bonnie Jo DeCoursey and Larry Leshan for their creative portrayals of the interviewed patients. Others who provided skillful assistance were Susan Linn, Hank Rowland, and Dawn Lopresto. And, of course, the project was made possible by the effort of the clinicians who volunteered their time amidst busy schedules.

Finally, my deepest appreciation goes to my family for the motivation they provided: Janet, with her loving support and encouragement, and Elizabeth Anne, with her spirited insistence on entering the world just prior to the completion of this paper.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS . . . . .	ii
ABSTRACT . . . . .	vi
CHAPTER I. CLINICAL JUDGMENT. . . . .	1
Introduction. . . . .	1
The Significance of Judgment. . . . .	2
Definition of Judgment. . . . .	5
The Domain of Psychological	
Research on Judgment. . . . .	10
Results and Criticisms of	
Research on Judgment. . . . .	19
Research Suggestions. . . . .	32
CHAPTER II. CLINICAL PSYCHOLOGICAL TESTING . . . . .	36
Introduction. . . . .	36
Historical Background . . . . .	37
The Referral Process as an	
Interprofessional Relationship. . . . .	42
The Referral Process. . . . .	56
Empirical Research on the Use	
of Psychological Reports. . . . .	73
Applications in the Present	
Study . . . . .	90
CHAPTER III. METHOD . . . . .	94
Summary . . . . .	94
Hypotheses. . . . .	95
Subjects. . . . .	96
Materials . . . . .	97
Measures. . . . .	100
Procedures. . . . .	104

	Page
CHAPTER IV. RESULTS . . . . .	111
Subject Characteristics . . . . .	111
The Hypothetical Cases . . . . .	116
Hypothesis #1. . . . .	122
Certainty Measures . . . . .	136
Hypothesis #2. . . . .	140
Hypothesis #3. . . . .	144
Hypothesis #4. . . . .	147
CHAPTER V. DISCUSSION. . . . .	150
APPENDIX A. ADMITTING NOTES. . . . .	159
APPENDIX B. TAPED INTERVIEW TRANSCRIPTS. . . . .	163
APPENDIX C. NURSING REPORT SUMMARIES . . . . .	167
APPENDIX D. PSYCHOLOGICAL REPORTS. . . . .	169
APPENDIX E. BRIEF PSYCHIATRIC RATING SCALE . . . . .	179
APPENDIX F. SYMPTOM CHECKLIST. . . . .	183
APPENDIX G. DIAGNOSTIC IMPRESSION. . . . .	184
APPENDIX H. DIAGNOSTIC SELECTION . . . . .	185
APPENDIX I. ADDITIONAL INFORMATION . . . . .	189
APPENDIX J. RECALL NOTES . . . . .	191
APPENDIX K. FINAL QUESTIONS. . . . .	192
APPENDIX L. STATEMENT OF INFORMED CONSENT. . . . .	193
REFERENCES. . . . .	195
BIOGRAPHICAL SKETCH . . . . .	207

Abstract of Dissertation Presented to the Graduate School  
of the University of Florida in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy

CLINICAL JUDGMENT AND THE USE  
OF PSYCHOLOGICAL REPORTS

By

James Martin Siwy

August, 1984

Chairman: Roger K. Blashfield, Ph.D.

Major Department: Clinical Psychology

Psychological test reports have long been a common component of the information used to make psychiatric judgments. A review of the literature on human judgment and on the use of psychological reports revealed that reports had received little empirical study and that some of the common findings in judgment research had not been tested in the clinical area. Judgment research has emphasized persistent biases such as overconfidence and inability to use disconfirming information. Using hypothetical psychiatric cases in a naturalistic manner, this study tested how psychiatrists and clinical psychologists made use of psychological

reports that either confirmed or disconfirmed their previous judgments. In two sessions, clinicians were given increments of clinical information about two patients whose descriptions were designed to maximize diagnostic ambiguity. In the second session the psychological report for one patient confirmed the subject's first judgment about the presence of thought disorder; the report for the other patient disconfirmed that judgment. Contrary to the main hypothesis, the 32 subjects tended to follow the reports rather than to ignore disconfirming evidence. Clinicians judged in a rational, discriminating and conservative manner. Confidence did not increase over time as predicted. No significant differences in judgment were found according to profession or level of clinical experience. The results support a more rational model of human judgment and suggest that psychological reports can have a significant influence on diagnostic judgment.

CHAPTER I  
CLINICAL JUDGMENT

Introduction

The status of psychological testing within clinical psychology has generated controversy through much of the past thirty-five years. Questions of the appropriateness and usefulness of testing have often been linked to the issue of professional identity, particularly concerning the psychologist's role in the mental health field vis-a-vis other mental health professionals. In contrast to the apparent importance of psychological test reporting through the years, the literature has revealed very little substantial research on the use of psychological reports.

The activities that comprise a psychological assessment are complex. Numerous factors affect the referral, testing and consulting processes. These will be discussed in the latter part of the literature review. Of particular interest will be the use of test reports by the referring party or report "consumer." As the survey of the literature will reveal, little empirical work has been done on



how information obtained from a psychological consultation is integrated with other clinical information.

Consideration of the use of psychological test reports is informed by the more general studies of the use of information in person perception, judgment and decision making. These areas of research have become prominent in the burgeoning literature of cognitive psychology, which is the general domain of the present topic. The reading of a psychological report can be viewed as a psychological event with many variables that would potentially relate to the effects of the report. These include such processes as attention, comprehension, evaluation, information processing, memory and, finally, action. Another way of framing the topic is to say that the clinical use of test reports is a particular type of integration of information on person perception. Rather than extending the discussion to encompass the many topics of cognition, the following review will address the study of judgment as the most pertinent context in which to consider the proposed study. A discussion of definitions will help clarify and delimit the topic.

### The Significance of Judgment

Judgment has long held an esteemed position as one of the processes of the mind. Philosophers have, through

the centuries, commented upon it and described it, often as a simple, straightforward activity, untroubled by the complexities attributed to it by modern psychological research. Cassirer (1944), for example, in his discussion of the Stoic view of life, stated:

Judgment is the central power in man, the common source of truth and morality. For it is the only thing in which man entirely depends on himself; it is free, autonomous, self-sufficing. (p. 8)

Further citations would be unnecessary in establishing that the judgmental process has had a vital place in human affairs. Indeed, the use of judgment has been recognized through the ages as one of the distinctive demands of being human.

Clinical judgment, therefore, is not merely an academic expression of an opinion. It involves elements of commitment and responsibility. When clinicians make their judgments, they are in a sense establishing precedent as to what sorts of patients will be described in particular ways. To a certain extent these judgments may be evaluated by comparison with established criteria. However, the means to evaluate judgmental validity scientifically is limited. Thus, clinical judgment will be accepted partly, if not predominantly, according to the authority bestowed upon the clinician. The clinician must shoulder the commitment and responsibility for his judgment in much the same

manner as Polanyi (1962) has described judicial and research judgment:

The course of scientific discovery resembles the process of reaching a difficult judicial decision . . . . in both cases a passionate search for a solution that is regarded as potentially pre-existing, narrows down discretion to zero and issues at the same time in an innovation claiming universal acceptance. In both cases the original mind takes a decision on grounds which are insufficient to minds lacking similar powers of creative judgment. The active scientific investigator stakes bit by bit his whole professional life on a series of such decisions and this day-to-day gamble represents his most responsible activity. (pp. 309-310)

As will be discussed further, this personal element suggests that research on judgment should begin with examination of what clinicians are doing rather than leaping ahead to an assessment of how validly they are judging in comparison with a contrived standard. What is central to the judgmental process is the personal commitment of the judge:

Though every choice in a heuristic process is indeterminate in the sense of being an entirely personal judgment, in those who exercise judgment competently it is completely determined by their responsibility in respect to the situation confronting them. (Polanyi, 1962, p. 310)

The point here is not to debate whether such responsibility and authority should be granted to clinicians. In most clinical situations action does depend on such personal judgment. What is required is a deeper understanding of

what judgment is and how it in fact is used in regards to patients today.

### Definition of Judgment

Discussions of clinical judgment in the psychological literature seldom refer to definitions outside that literature. Before looking at how psychologists define judgment, it will be helpful to examine more general, formal definitions.

Dictionaries reveal that "judgment" has multiple meanings and usages. Most definitions are of a general nature, such as "that function of the mind whereby it arrives at a notion of anything; the critical faculty; discernment" (Oxford English Dictionary, 1971), or "the power or ability to decide on the basis of evidence" (Webster's Third International Dictionary, 1976).

Webster's divided the philosophical consideration of judgment into the Scholastic and Kantian approaches. The former concerns the realm of value: "the capacity to arrive at a decision about the value of things." The Kantian view is of two types. The first involves "the power of relating particulars to general terms or concepts." This can be "reflective judgment," which "proceeds from given particulars to the discovery of a general concept or universal principle under which the particulars may be

subsumed." Alternatively, the first type may be "determinative," which "proceeds from a general concept or universal principle and designates the particulars which are to be subsumed under the general." This first type would seem to describe inductive and deductive reasoning, respectively. The second Kantian definition in Webster's described judgment as "a capacity mediating between reason and the understanding; broadly the critical faculty" (emphasis in original).

A review of psychologists' definitions of judgment reveals that the Scholastic or value issue has tended to be avoided and that the Kantian definitions have been employed. Much attention has been given to the reasoning process, both reflective and determinative. In addition, there has been a recognition of judgment as something beyond reasoning.

Some psychologists who study judgment avoid explicit definitions. For instance, Rappoport and Summers (1973) expressed the presuppositions of a number of researchers who had attended a conference on judgment. They saw judgment as an aspect of thinking, based upon the goal of adaptation to uncertainty, that allows maintenance of organization and continuity in behavior while "going beyond perceptual and cognitive 'givens'" (p. 4). Judgment is used to mediate between a person's intentions and the uncertain environment; it is thus centered upon the

relationship of a cognitive system with an environmental system. Similarly, in lieu of a definition, Newell (1968) referred to judgment as "an umbrella term," comparable to such terms as "perception," "learning" and "cognition." These terms need not be well defined because they designate a general class of phenomena about which theories are developed. He preferred to classify judgment within "pretheoretical language" that eschews precise definitions. Instead, he offered a list of definitional strands that serve to circumscribe the topic.

Other psychologists have been more explicit in defining the term. For instance, Bieri, Atkins, Briar, Leaman, Miller and Tripoli (1966) saw judgment as the assignment of a stimulus to a response category. Johnson (1972) ascribed to a similar approach, with the emphasis that judgment functions to bring order to matters of uncertainty:

Judgment begins with unordered objects, events, or persons, assigns them to specified response categories so as to maximize the correspondence between the responses and the critical dimension of the stimulus objects, and thus ends with a more orderly situation. (p. 340)

An essential element in most definitions of judgment has distinguished the process from mere application of rules of reasoning. Newell (1968) pointed out that

Judgment fills the gaps in rational calculation. If the calculation could do it all, then no judgment is required. (p. 4)

He added that judgment has "a flavor of immediacy, and . . . nonrationality" (p. 4). Responses to complex situations cannot be formally programmed:

The upshot is that whatever can be analyzed, becomes part of the formal apparatus, and the remainder is left to be subjective judgment. Thus, although one may still assert that there is good judgment and bad--that a man can somehow arrive at his inputs to the system appropriately or not--judgment occurs in precisely those situations where the scientist is least able to attribute rationality to them. (Newell, 1968, p. 5)

Perhaps this aspect of judgment was best captured by Hammond, Stewart, Brehmer and Steinmann (1975), who described human judgment as "a cognitive activity of last resort" (p. 272).

Different views have existed as to whether judgment is to be distinguished from decision making. Slovic and Lichtenstein (1973) observed such a distinction to be "a tenuous one" (p. 16). Bieri et al. (1966) tended to agree with this view.

Others have preferred a separation of the topics (Einhorn, Kleinmuntz & Kleinmuntz, 1979; Hogarth, 1980, 1981). For example, Hogarth, who used "choice" in place of decision, drew the following relationship between judgment and choice:

Judgment can be thought of as providing a temporal background of mental activity that is punctuated by particular choices. This is exemplified in the target model, where aiming is used as the analogy for judgment, and the actual shooting is analogous to choice. This does not, of course, deny that one can shoot without aiming, just as one can ignore judgment in choice. (1981, p. 201)

In a more recent paper, Einhorn and Hogarth (1983) distinguished "diagnostic inference" from predictive clinical judgment. The former was defined as the inferences of the causal process that could have produced an observed set of "outcomes/results/symptoms" (p. 1). The authors were apparently dealing with the question of etiology as one aspect of diagnosis. Although not explicitly stated by these authors, it would seem that diagnostic inference is one part of clinical judgment to be considered in relation to the essential predictive aspect.

From the above discussion, several points apply to the consideration of clinical judgment in the proposed study. First, judgment is perhaps the most crucial cognitive activity of the professional. Judgment involves reasoning and assumption of responsibility that cannot rely on calculations alone. Second, because judgment is so closely related to decision, it should not be relegated to the refined category of cognition, but must be studied in the context of action. Finally, although questions of value have not generally been the focus in most research on



judgment, including the present one, the influence of a judge's valuing or devaluing has to be incorporated as an implicit background for judgments and choices. The effect of evaluation will be more evident in the consideration of judgmental biases. Although it may at times be helpful to distinguish evaluation from prediction, the person as judge never ceases to be an evaluative person. This point may seem obvious; it is included here as a matter of emphasis to contrast the objectivist bias that has dominated twentieth century psychology (Polanyi, 1962).

### The Domain of Psychological Research on Judgment

#### Introduction

In recent years the study of judgment has been a popular area for psychological research. As is true for any relatively new topic, the study of judgment has been approached from many perspectives. Among the reviews of this literature, there does not appear to be any consensus on how the topic may be divided. Newell (1968) gave a reminder of the most basic division between psychophysical judgment and the more "general" study of judgment. Different authors have employed various outlines in the discussion of the latter type, which is the subject of this proposal.

### A "Systems" Approach

One method of review has been the consideration of the judgmental task according to the number of known "systems" under observation. Using this approach, Rappoport and Summers (1973) classified human judgment according to a "general systems theory hierarchy" (p. 5). This taxonomy includes four classes: the single-system case, two-system case, three-system case and n-system case.

The single-system case consists of judgment made under uncertain conditions which lack adequate feedback. The authors cited both foreign policy decisions and psychodiagnosis as examples. The latter was described:

Given ambiguous symptomatic information, the diagnostician must categorize patients and recommend treatment. He may then never see the outcome or he may receive outcome information that can be attributed to many factors other than those leading to the diagnosis. (Rappoport & Summers, 1973, p. 5)

The authors provided support for the focus of research on the clinician's thinking rather than on characteristics of patients:

The single-system case must therefore be understood as largely focussed on the cognitive system of the judge, because the environmental or task system remains obscure. (p. 5)

Hammond et al. (1975) noted that the single-system case has been the one most studied in the research literature.

In the double-system case, the task outcomes, and thus the task structure, are known. One person forms judgments about one task system. The task system may be a person. Accuracy of judgment and learning to improve judgment have been topics frequently studied in this case. Hammond et al. noted that research in this format has shown the usefulness of "cognitive feedback" in contrast to outcome feedback in improving judgmental performance. When subjects were given information on their own judgmental system or the task system, their accuracy in judgment improved faster than if they were given only outcome feedback on their performance. This is an example of the use of "cognitive aids."

The triple-system case involves two judges and a task. In social judgment theory, the focus has been on cognitive differences between judges rather than the differential gain as the source of conflict (Hammond et al., 1975).

The final case is that of the n-system, in which more than three persons are studied and the task system may or may not be known. This case has examined the characteristics of different factions and their policies. Thus, it has encompassed the cognitive bases of conflict within a group.

### A Historical Perspective

A useful historical perspective was offered by Bieri et al. (1966), who examined the study of judgment as applied to clinical inference. The authors listed three phases of this research. The first was introspective reporting exemplified by Freud and, to use a more recent example, Erikson's (1964) account. The second phase, labelled the "reliability-validity phase," refers to the period of intensive study of psychological tests that took place during the decade following World War II. The preoccupation of clinical psychological research at that time was the reliability and validity of clinical inferences based upon projective methods of assessment. The third phase was the period of debate about clinical versus statistical prediction, led by Meehl's (1954) book. Bieri et al. (1966) concluded with what they called the "contemporary" phase of development of theoretical models of the judgment process. This phase can be divided into three domains of psychological research.

The first domain has been that of cognitive social psychology, which has studied clinical judgment in the context of social perception. Implicit personality theory has been a particular topic within this domain that is applicable to judgment theory (Schneider, 1973; Wegner & Vallacher, 1977).

The second contemporary domain has consisted of the historical influences of the psychophysical approach as applied to social judgment. This domain has not been easily separated from the others because of the enduring influence of psychophysical conceptual models. For instance, Bieri et al. (1966) drew heavily upon psychophysical constructs in conceptualizing processes of social judgment. They referred to a cue as an "input" which has multidimensionality, or "differentiation," that can be calibrated along intervals within a dimension. This last measurement is "articulation" (p. 16). More recently, Einhorn and Hogarth (1983) used a psychological model in their discussion of diagnostic inference. They began with the perceptual concept of figure/ground, which includes the importance of an "assumed causal background" to salient information. The concept of "net strength" of a pattern of cues is invoked. This refers to the relative amount of evidence for a particular diagnosis, analogous to the detection of a signal among competing signals. As in perception, diagnosis is a "constructive" process dependent in part on variables related to the diagnostician, such as expectations. Finally, like perception, diagnosis often is a rapid judgment that is made without awareness of underlying processes.

The third and final domain suggested by Bieri et al. (1966) has been the application of mathematical models to

judgmental processes. The literature within this domain has been so extensive that some reviewers see this approach as equivalent to the study of judgment. Most of the following research models, although not precisely delineated by Bieri, et al., have constituted this third domain.

### Research Models

Slovic and Lichtenstein (1973) provided one of the most comprehensive reviews of this area. Although focusing on the contrast and comparison of the regression and Bayesian approaches, the authors also delineated several other models of research. These models include process tracing (Kleinmuntz, 1968); multidimensional scaling (Wiggins, 1973); application of information theory (Bieri et al., 1966); "probability learning" and signal detection theory (Green & Swets, 1966). Slovic and Lichtenstein's discussion revealed the overlapping of different models. Their focus on the regression and Bayesian models accurately represented the main direction of the literature.

They divided the regression or correlational approach into two streams. The first has focused on the judge and his or her ways of combining information into a judgment. Hoffman (1960) was an early exemplar of this stream. Slovic and Lichtenstein saw Anderson's (1965, 1974) extensive application of integration theory as a continuing part of this tradition.

The other stream has followed Brunswik's (1952, 1956) philosophy of "probabilistic functionalism," which is centered upon the adaptive interrelationship between the organism and the environment. Slovic and Lichtenstein noted that Brunswik was "the foremost advocate" of "'representative design'" (1973, p. 20), which calls for the study of an organism in realistic settings so that experiments may represent the "usual ecology" of the organism.

They described Bayes' Theorem as a normative model that prescribes optimal judgment based on specification of "certain internally consistent relationships among probabilistic opinions" (p. 31). The Bayesian approach produces a "distribution of probabilities over a set of hypothesized states of the world" (p. 30). It is aimed at enabling maximization of utility in decision making. Slovic and Lichtenstein noted that Brunswik could have chosen to use conditional probabilities (Bayesian approach), rather than correlations, for the assessment of relationships in a probabilistic environment. The implementation of either approach has involved mathematical modeling, which has been the dominant preoccupation in judgment research.

As in any topic of psychological study, some researchers have focussed on individual differences. Wiggins (1973) presented a "general individual difference model" that uses

multivariate statistics and multidimensional scaling of stimuli. The model attempts to derive ideal "types" of judges that reflect judgment by individuals more informatively than a group mean and more parsimoniously than purely idiographic analyses. Kaplan (1975) also looked at individual differences, using Anderson's (1974) basic model of information integration. In this model, each judge assigns a "scale value" to information along a particular judgmental dimension. This scale value can be described as the "meaning" given to the information in terms of the dimension of judgment. Information may have different meanings according to different evaluative dimensions; thus, stimuli are considered to be multidimensional. The importance for judgment is to what extent the information lies upon the same response dimension as does the judgment.

#### The Contextual Paradigm

A summary of the various models of judgment research indicates a general consensus on the use of a contextual approach (Jenkins, 1974). The diversity of models may be more apparent than substantial. Johnson (1972) observed that many models cannot be compared because of their specificity: "Most models are applicable only to those situations in which they were developed" (p. 350).

Most authors have paid homage to Brunswik's idea of representative design, if not the Lens Model itself. At



times new vocabularies have been applied. For instance, in their presentation of social judgment theory, Hammond et al. (1975) restated Tolman's idea of the "cognitive map" and Brunswik's concept of the "texture" of the organism and the environment. Simply put, relationships are the "fundamental units of cognition" (p. 272). What Brunswik referred to as "texture," these authors denoted as "causal ambiguity." This concept is in reference to the relationships among variables or cues. Social judgment theory distinguishes between "surface," which consists of given cues and "depth," which refers to the inferred conditions. An example would be a person's pattern of speech as the surface condition and thought disorder as the depth condition. The region between surface and depth is referred to as the "zone of ambiguity." This zone encompasses the often complex entanglements of interrelationships involving surface and depth. The authors noted that a single effect may have multiple causes and that one cause may have multiple effects. The causal ambiguity is due to three inherent conditions: (1) there is an imperfect relation between surface data and depth variables; (2) the relation between these two sets of variables may assume different forms, e.g. linear or curvilinear; (3) this relation may be organized according to a variety of principles, e.g. additivity or pattern. The environment is dominated by "causal ambiguity." People

apply their cognitive processes to reduce this ambiguity as much as possible. However, this active processing, which includes such processes as perception, thinking and learning, at times reaches its limit without resolution of the ambiguity. What remains is a "passive" means to arrive at a conclusion. This is the essential situation in which judgment occurs. Hammond et al. (1975) noted that social judgment theorists have made the reduction of causal ambiguity their primary goal. The method for doing this has been the externalization of the properties of the zone of ambiguity in each system, organismic and environmental.

This research paradigm has probably been the most widely followed one today. The most extensively published group that has continued in this tradition is located at the Center for Decision Research at the University of Chicago (Einhorn & Hogarth, 1978; Hogarth, 1980; Einhorn & Hogarth, 1982, 1983).

## Results and Criticisms of Research on Judgment

### Introduction

A brief review of the research will focus upon judgmental biases. First, some general conclusions about the status of knowledge on judgment will be quoted.

### General Conclusions

Hammond et al. (1975) concluded that the following "empirical regularities" exist both inside and outside the laboratory:

(1) People do not describe accurately and completely their judgmental policies, (2) people are often inconsistent in applying their judgmental policies, (3) only a small number of cues are used, (4) it is difficult to learn another person's policy simply by observing his judgments or by listening to his explanations of them, (5) . . . cognitive aids . . . can reduce conflict and increase learning, and (6) linear, additive organizational principles are often adequate to describe judgment processes. (p. 305)

Similarly, Slovic and Lichtenstein (1973) offered four generalizations about the state of knowledge of human judgment. First, responses of judges are highly predictable and can be represented in a quantifiable manner, often more accurately than the judge can describe his own processes. Second, because weighing and combining information are very difficult, judges use simplified decision strategies to "reduce cognitive strain" (p. 89). Third, the structure of the situation being judged has a major effect on judgment. This includes the sequences of information, manner of display and the nature of required response. Fourth:

Despite the great deal of research already completed, it is obvious that we know very little about many aspects of information use in judgment. (p. 89)

The authors added that "an enormous task" remains in integrating research and theory on judgment into "the mainstream of cognitive psychology," including such topics as attention, concept formation, learning, memory and problem solving (p. 89).

In more concise terms, Hogarth and Makridakis (1981) noted that cognitive psychological studies of human judgment have resulted in two major conclusions: "(1) ability to process information is limited; and (2) people are adaptive" (p. 116). Understanding of the context is a paramount factor. The authors emphasized that people are highly motivated to understand and control their environments. However, much error in human decision making can be attributed to superficial information acquisition and processing biases.

#### Common Shortcomings and Biases

Most reviews of judgment research have emphasized the failings of human judges. Einhorn and Hogarth (1978) cited several literature reviews and summarized the empirical documentation of the limited capacity in human judgment:

Although the study and cataloguing of judgmental fallibility have had a long history in psychology . . . an accumulating body of recent research on clinical judgment, decision making, and probability estimation has documented a substantial lack of ability across both individuals and situations. (p. 395)

They noted the low predictive ability shown in clinical settings, adding

It is apparent that simple statistical models for combining information consistently provide more accurate predictions than the judgments of clinicians. (p. 395)

Much of the research literature has been dominated by the documentation of common biases in human judgment. Hogarth (1980) presented an extensive list of such biases, e.g. primacy/recency effects (Luchins, 1957, 1958; Leach, 1974) and the ignoring of base rate information (Kahneman & Tversky, 1973). In lieu of reiterating the rest of the list, the following comment on clinical judgment can serve as a forthright reminder of the judgmental distortions which clinicians are commonly prone to commit:

There is always enough evidence in a rich source of data to nurture all but the most outlandish diagnosis. (Arkes, 1981, p. 326)

In contrast, Christensen-Szalanski and Beach (1984) presented evidence that a citation bias has existed in the literature on human decision making. Studies which have demonstrated the rationality of subjects' judgments have been undercited as compared to studies which have emphasized subjects' inadequacy in judgment. These authors argued that a balanced view of the research demonstrated mixed results in human judgment reliability, rather than only the more publicized biases.

In the present study, the bias of interest is the inability to use disconfirming evidence. Rappoport and Summers (1973) stated that problems in cue congruency, or disconfirming information, represented a "critical issue" in the study of judgment. They referred to such problems as "common." How a judge responds to cue inconsistency is not only theoretically important, but also of considerable practical significance" (p. 143).

Hogarth (1980) noted that the disregard for inconsistent information is a strategy born of the search for consistency. He cited the phrase, "'thirst for confirming redundancy,'" used by Bruner, Goodnow and Austin (1956).

#### Confidence in Judgment: Measurement and Findings

The difficulty shown by judges in adequately using disconfirming information has been one factor related to the problem of inappropriate confidence in judgment. The definition of confidence used here is taken from Blaser (1978): "the individual's emotional satisfaction and certainty about a judgment at which he has arrived" (p. 1276).

The study of confidence in judgment has received recent empirical attention within psychology. A brief review of this research will include measurement issues, empirical findings and explanatory discussion.

Johnson (1972) has been one of the few authors who has devoted any discussion to how confidence in judgment can be measured. He surmised that confidence can be measured either

semantically, e.g. with such terms as "very confident" or "quite certain," or on a rating scale, which is the usual procedure. Johnson suggested that subjects would feel most comfortable with a rating scale extending from 0 to 100 because people are able to conceptualize confidence in terms of percentage of certainty. He noted that distributions of confidence ratings are often U-shaped or J-shaped because of the frequent usage of end points. His review stated that in studies of psychophysical judgments, confidence has been shown to increase in concordance with an increase in the difference between the variable stimulus and the standard.

Slovic (1966) measured confidence in judgment by having subjects assign a probability estimate of the accuracy of their ratings. This represented essentially the same mode and range of response as suggested by Johnson (1972).

Most studies involving confidence ratings have employed a Likert-type rating scale. For example, Cantor, Smith, de Sales and Mezzich (1980) used confidence ratings made by clinicians who rated how poorly or how well, on a scale from 1 to 7, written case descriptions of patients fit into the diagnostic category chosen by the subjects. Eker (1981) and Blashfield and Sprock (1983) also used 7-point rating scales in the measurement of clinicians' confidence in diagnosis of hypothetical patients. In more general, non-clinical studies of the construct, "decisiveness," Weissman

(1980) used a 10-point scale for "Confidence in Decision," one of six components of decisiveness. Other studies have employed the same essential self-report rating (Oskamp, 1965; Ryback, 1967).

A major finding of a number of studies has been the negative relationship between confidence and accuracy. This has been obtained in both non-clinical and clinical research.

In a study that debunked popular attitudes about "women's intuition," Valentine (1929) reported a tendency for a negative relation between confidence in judgment and accuracy:

A feeling of unusual confidence, then, gives no clue to the reliability of the judgment: rather it might seem to act as a warning that the judgment is particularly unreliable. (p. 230)

Ryback (1967) found that, in a study of judgment involving comparisons of geometric designs in the absence of systematic feedback, undergraduates' confidence increased with experience. Accuracy did not increase; there was no relationship between confidence and accuracy. Malpass and Devine (1981) found that confidence and accuracy in eye-witnesses' selections of vandals from a lineup were unrelated. Witnesses who made choices tended to indicate high confidence in their selections, whereas those who rejected the lineup had low scores in self-ratings of confidence in their decisions. The authors noted that this is consonant with



the findings of several other studies of witness confidence and accuracy.

Oskamp (1965) studied psychologists' confidence in their clinical judgment. He reported that confidence increased steadily and significantly with increasing information about a case, but that accuracy did not increase. Similarly, Schinka (1976) found no significant relationship between clinicians' confidence and their accuracy in judgments.

Blaser (1978) studied the influence of different variables on the level of confidence in person perception. He confirmed previous empirical findings that increased information was related to increased confidence in person perception. Also the type of personality characteristic being judged was related to level of confidence. In his study of ratings by ten experienced psychotherapists (five analysts and five non-analysts), the rated variables ranked from most to least confident were

intelligence, likeability, suffering, contact, exhibitionism, endurance, maturity, defensiveness, suggestibility, readiness to accept interpretations, introspection, reality conformity.  
(p. 1278)

Blaser observed that these were each complex variables not accessible to direct observation, thus not making it clear as to why they would differ in terms of judges' confidence in their ratability. He also found that confidence level

related to the personalities of the judges, that is; some judges tended to be more confident than others in their ratings generally. The author suggested that this factor could relate to previous findings of no relation between accuracy and confidence of clinical judgment.

Eker (1981) found that confidence in diagnosis for American and Turkish clinicians was above the midpoint on a 7-point scale even when cases had inadequate diagnostic information. The high confidence was based upon knowledge of present symptoms. Eker criticized the implied tendency to not feel a need to search for other information, such as historical data.

One consistent finding in research on clinical judgment has been the rapidity at which clinicians arrive at a diagnosis (Gauron & Dickinson, 1969; Sandifer, Hordun & Green, 1970; Kendell, 1973; Clavelle & Turner, 1980). In these studies, clinicians often produced diagnoses within a very few minutes of being presented information. Although these findings did not necessarily imply overconfidence, they did raise the hypothesis of hasty decision making.

Several authors have offered explanations of the overconfidence phenomenon, often in relation to the problem of inattention to inconsistent information. Kahneman and Tversky (1973) demonstrated that people made predictions by selecting outcomes that were most representative of inputs. They suggested that confidence in predictions varied directly

with the extent to which a selected outcome was more representative of the input as compared with other outcomes. The authors stated that consistent information would increase confidence, even though it had no relationship to predictive accuracy. They concluded:

The foregoing analysis shows that, for example, consistency and extremity, are often negatively correlated with predictive accuracy. Thus, people are prone to experience much confidence in highly fallible judgments, a phenomenon that may be termed the illusion of validity. Like other perceptual and judgmental errors, the illusion of validity often persists even when its illusory character is recognized. (p. 249; emphasis in original)

Einhorn and Hogarth (1978) posed the following question in their review:

How can the contradiction between the considerable evidence on the fallibility of human judgment be reconciled with the seemingly unshakable confidence people exhibit in their judgmental ability? In other words, why does the illusion of validity persist? . . . Why does experience not teach people to doubt their fallible judgment? (p. 396)

The authors answered their questions by focusing on the environmental context in which judgments take place, specifically on the relationship of judgmental tasks and their outcomes. They noted:

In real-world situations, judgments are made for the purpose of choosing between actions. This means that outcome information, which is available only after actions are taken, is frequently the only source of feedback with which to compare

judgments. Therefore, to understand how people learn about their judgmental ability, it is necessary to consider judgments, actions, and outcome feedback together. (pp. 396-397)

The authors presented a mathematical model for confidence in judgment, which they redefined as "the strength of the learned concept 'my judgment is accurate'" (p. 401). Using learning theory principles, they derived an equation representing confidence as a function of the total feedback effect, which in turn was a function of the number of decisions made, the probability of selecting a particular action, the positive hit rate and the reinforcing value of negative feedback (pp. 401-402). They suggested that confidence in judgment was acquired slowly, then would experience a relatively rapid rise and finally would level off with large amounts of experience. The model suggested that once confidence in judgment was achieved, it would be highly resistant to extinction.

In another article, Einhorn (1980) offered the concept, "outcome irrelevant learning structures (OILS)" (p. 4), as an explanation for the pervasiveness of overconfidence in judgment. A major source of these structures is the inability to make use of disconfirming information. To overcome this common bias, a judge had to be willing to try to falsify his hypotheses, even when he felt certain of their correctness. Einhorn noted that seeking disconfirmation is a relatively recent methodology in the history of science. In general,

Einhorn posited three main factors to explain overconfidence in judgment:

(1) lack of search for and use of disconfirming evidence, (2) lack of awareness of environmental effects on outcomes, and (3) the use of unaided memory for coding, storing and retrieving outcome information. (p. 13)

Einhorn (1980) offered a number of solutions to the overconfidence problem. First, formal training in experimental design, use of control groups, and awareness of base rates could help offset the tendency to discount disconfirming evidence. Adopting a comprehensive model such as the Brunswik Lens Model could be helpful in gaining awareness of environmental effects on outcomes. Finally, keeping a record of judgments and outcomes could aid problems due to memory deficiency.

In general, judgment appears to be a complex task involving distillation of large amounts of information into a structured scheme of reasoning to aid the act of decision. Given the stress of the judgmental task, the judge tends to seek out reinforcing information that would confirm his efforts and conclusions. Perhaps Hogarth (1980) best described the motive for overconfidence in his statement that people were motivated to avoid "psychological regret" for either taking or failing to take an action.

Einhorn (1980) tried to temper the problem with the suggestion:

Perhaps overconfidence in judgment is in some ways functional--lack of confidence in judgment might result in too much analysis and a crippling of the ability to make quick choices. (p. 14)

However, the thrust of his writing delineated the essential dilemma of fallible judgment:

We are in the common but unenviable position of having to continually judge our judgments. Furthermore, if judgmental biases are as frequent as much research suggests. . . . we must seriously entertain the hypothesis that a great deal of what we believe about our judgmental ability is in error. (pp. 1-2)

Given these findings, the continuing empirical study of judgment and judge's confidence would seem to be an important project for clinical research. Adams and Adams (1961) articulated a compelling rationale for making clinicians and other judges aware of their biases:

It may be that realism of confidence is in many situations a more important variable than level of performance itself. To take one example, it might be more important, in terms of his future work, social interactions, confidence in himself, etc., for a student to be able to discriminate realistically between what he knows and what he does not know than it would be for him to know considerably more than he does know without such discrimination. Realism of confidence seems obviously of importance, not only for scholars but in general for individuals making decisions having consequences of importance, either for themselves or for others. (pp. 36-37)

### Research Suggestions

From the authors reviewed, a number of recommendations for research may be distilled. The most repeated recommendation has been for representative research designs. This could involve a number of factors often overlooked in previous studies. For example, Kleinmuntz (1968) suggested that a simulation experiment could better approximate a real-life situation if "noise," in the form of irrelevant or insufficient information were included. This same idea was expressed by Hammond, et al. (1975) in terms of the intrinsic ambiguity of the environment:

The methodological corollary is that such ambiguity among relations must be represented in the judgment tasks used to study human judgment. (p. 273)

Einhorn, Kleinmuntz and Kleinmuntz (1979) observed that most research on judgment involved "deliberative" responses in which subjects were given enough time and were encouraged to "think through the problem" (p. 469). Also, the response was usually in the form of a quantitative rating, which contributed to the use of compensatory rules and "focuses attention on the trade-offs between cues in order to be quantitatively precise" (p. 469). One could infer that a less deliberative and quantitative situation would be more representative of common judgmental tasks. The goal for research would be to ascertain the nature of the judgmental process.

Anderson captured the essence of such process by noting that a judgment was dependent upon a "train of inferential thought" (p. 91). Valid theory and research on judgment was likely to be an elusive goal, as was suggested by his major point:

In a very real sense, therefore, people do not know their own minds. Instead, they are continuously making them up. Knowledge and belief are not static memories, but typically involve active, momentary cognitive processing. (p. 89)

Along these same lines, Hogarth (1981) wrote of the importance of feedback and redundancy in the continuous judgmental processes that characterize everyday life. He reminded the reader that the main purpose of judgment was to facilitate action. Discrete judgmental tasks studied in laboratories have often suffered from a lack of both redundancy and feedback. He summarized the "one result" that "dominates" descriptive research in decision making over the years:

Judgment and choice depend crucially upon the context in which they occur and the cognitive representation of that context. However, both environment and mind interact in continuous fashion. . . . Theories of judgment and choice that lack a continuous perspective exclude one of the most important determinants of the behavior they purport to explain. (p. 213)

The objectives of Hammand et al. for social judgment reflected one of the most emphatic programs of



representativeness. This pragmatic approach was meant to be (1) "life relevant," (2) descriptive, rather than aimed at establishing laws of behavior or judgment, and (3) designed to create "cognitive aids" to help people make judgments.

Another recommendation for research has been the use of multiple methods. Slovic and Lichtenstein (1973) criticized the myopic view of some researchers who have been unaware of other methods of investigation. In the same spirit, Einhorn et al. (1979) compared linear regression and process tracing models. They concluded:

Although we believe that both approaches treat the underlying process at different levels of detail, some process models may not be seeing the forest for the trees, whereas some statistical modelers may not see any trees in the forest. (p. 483)

More specifically relating to clinical judgment, Eker (1981) recommended further research on "the amount and type of information" that clinicians view as satisfactory for diagnosis. Bieri et al. (1966) also discussed the crucial aspect of structuring response alternatives, the range of which is an important parameter in judgment:

The clinician who in an interview poses for himself the problem of determining what the client is like has a much broader array of response alternatives potentially open to him than does the clinician who must decide if the client is schizophrenic. . . . In the research situation, in order to achieve comparability of conditions

across judges, it is often necessary to confine all judges to the same set of response alternatives even though any given judge may feel that this is an irrelevant domain or not the most important one for this particular client. (pp. 10-11)

There is a lack of other specific recommendations to cite because of the limited literature on clinical judgment, which, in turn, has traditionally had little effect on clinical practice (Hammond, Hursch & Todd, 1964). A reason for this may be, as Slovic and Lichtenstein (1973) suggested, that most people regard judgment as an intuitive act that cannot be didactically taught. However, this attitude is counter to the whole enterprise that endeavors to create "expert" opinion. As Bieri et al. (1966) succinctly described:

Professional education in clinical psychology, social work, and psychiatry is based upon the assumption that it is possible to develop within the clinician capacities to render judgments which are both accurate and therapeutic. (p. 11)

In conclusion, Hogarth (1980) made three observations about decision making that can be borrowed as a good summary for the justification of research on judgment:

(1) people are generally unaware of how they make decisions and often why they prefer one alternative to others; (2) they show little concern for the quality of their own decision making processes (although the failures of others are often indicated with haste); and (3) the scientific study of decision making has not in my view attracted the attention it merits. (p. ix; emphasis in original)

## CHAPTER II CLINICAL PSYCHOLOGICAL TESTING

### Introduction

The processes of psychological assessment and consultation offer a practical arena in which to examine clinical judgment. As in most applied research, consideration of potential influencing factors reveals a host of variables that may be implicated in the relation being studied, e.g. in this case, between a psychological report and particular judgments or choices. A historical perspective is necessary for an appreciation of the meaning of the referral situation and the use of reports. This review will look at the emergence of testing in the history of clinical psychology. Included will be an examination of the interprofessional relationship between psychiatry and clinical psychology and the controversy that has surrounded clinical testing in the past three decades. A detailed description of the referral process will include discussion of the reasons for referral and potential factors influencing the use of assessment. Finally, empirical research will be

reviewed, including both surveys and the few studies that have endeavored to study the use of test reports directly. The need for more research will be emphasized, with general suggestions followed by an outline of the present study.

## Historical Background

### Introduction

The function and significance of psychological testing in clinical psychology can be understood only in its historical context. Changes in the use of tests parallel developments in the growth of the profession. A key factor in these developments has been psychology's relationship to psychiatry. Although other professionals, such as educators and nonpsychiatric physicians, have had an impact on testing, the relationship of interest here is with psychiatry, which has probably had the greatest influence on applied clinical testing.

### The Emergence of Testing in Clinical Psychology

Tallent (1965) provided an excellent discussion of the historical roots of clinical testing. He cited three sources: academic psychology, clinical psychiatry and clinical psychology. He compared the two differing approaches: (1) the academic-psychometric, based upon

operationalism and physicalism and (2) the clinical, based upon an idiographic and more impressionistic tradition. Similarly, Watson (1953) traced the development of clinical psychology to two general roots: the psychometric test tradition and the influence of dynamic psychology.

In the early part of the century, intelligence testing was the primary occupation of those psychologists who could be called "clinical." Watson gave an even more delimiting picture: "In fact, for years, the major task of the clinical psychologist was to administer the Stanford-Binet" (1953, p. 323). Until the 1940's, this role as "mental tester" remained the center of the psychologist's functional identity vis-a-vis psychiatry.

An expanded role for psychologists was signified by Brown and Rapaport's (1941) article in the Bulletin of the Menninger Clinic. The authors provided an overview of personality testing and explained the use of a psychological test battery in a psychiatric setting. Although the article focused on the various tests and their uses, its title, "The Role of the Psychologist in the Psychiatric Clinic," demonstrated the identification of clinical psychologists with the testing function. The authors emphasized that this role, which would now seem limiting to most psychologists, was a broadening of professional opportunity:

Recent advances in experimental psychology and the psychology of personality establish the psychologist in a new role in clinical psychiatry. Instead of limiting his activities to intelligence testing, the psychologist functions as an investigator of the personality by means of projective personality tests, concept formation tests, Lewinian techniques, and intelligence tests. (pp. 82-83)

The subsequent fifteen years brought about the complete acceptance of, if not infatuation with, personality assessment by clinical psychology. Watson (1953) cited the consensually agreed upon role of "diagnostic appraisal as a task of the clinical psychologist" (p. 341) as one of the stabilizing factors in the emergence of clinical psychology as a profession following World War II. Part of the initial enthusiasm for psychodiagnostic assessment emphasized the advantages of the "objectivity" of testing. For example, Brown and Rapaport (1941) declared:

To the clinical methods of the psychiatrist and the psychoanalyst [the psychologist] adds tools having a considerable degree of objectivity in investigating all the aspects of the total personality. (p. 83)

Psychologists were soon regarded by many as wielders of powerful projective techniques, further enhanced by the scientific responsibility of psychological research. Kubis' summary of clinical practice and research was a poetic testimony to the optimism of that era:

The projective tools offer the clinician an opportunity for the free exercise of his intuitions in a complicated life-space where the interactions of all forces cannot be controlled by exacting experiment. Though more restrictive, standardized psychological tests present the functioning individual side by side with a norm, thus enabling the psychologist or psychiatrist to evaluate this specific behavior in a more objective light. There is finally, the experimental procedure, uncompromisingly rigid and assuming the double role of alert sentry and discerning skeptic with respect to all claims demanding entrance into the citadel of science. It is clearly apparent that these diverse approaches to scientific truth are closely interrelated and often merge into a productive pattern well illustrative of the scientific method. Promising clinical insights are temporarily molded into a working tool, such as a projective test or questionnaire. These are, sooner or later, thrown into the experimental furnace and with the impurities dissolved, the remainder is often molded into a standardized form, a psychological test. Note here the torturous paths science compels clinical intuitions to take so that (ultimately) they may emerge in a purified form, independent of their creator and his special abilities. (1951, pp. 105-106)

Other authors were more skeptical. For example, Diethelm and Knehr (1951) emphasized that tests were aids to diagnosis and not substitutes for a clinician's judgment:

The well-trained clinician who takes sufficient time to study the patient will have little need for projective tests to establish the diagnosis. (p. 78)

These authors also relegated the use of psychological tests to diagnosis, "with only incidental usefulness

in therapy" (p. 73). They reasoned that, because most test procedures were cross-sectional in nature, they would be of limited use to dynamic psychiatry, which emphasized longitudinal aspects of etiological factors.

The subsequent disenchantment with testing will be detailed further. However, the point here is that psychologists' campaign for the acceptance of personality assessment was too much of a success. Rosenwald (1963) described the situation, albeit in somewhat extreme terms:

Psychodiagnosis became the signal function of clinical psychologists, even though some were active in other ways as well, and even though they were not the only ones to avail themselves of these tools. Before the mental health team became the practical reality which it now is in many hospitals and agencies, the boundaries of the psychologist's activity were tightly and impermeably drawn around testing. (p. 222)

Ivnik (1977) also stated that reviewers of the disputed value of psychological testing have agreed that "historically psychological tests have been intertwined with the clinical psychologist's professional identity" (p. 206). He added:

Tests provided an assessment procedure that differentiated the psychologist from other mental health professionals. Testing also defined a unique domain of professional functioning wherein the psychologist was the



unquestioned expert. Unfortunately for some, the testing boon soon became the clinical psychologist's bane. (p. 206)

Perhaps the most incisive remark has been that of Kahn, a psychiatrist, who lamented, "To see a whole generation of psychologists becoming testing technicians is one of our painful experiences" (1965, p. 1025).

In the developments since that time, psychologists have expanded their roles in many ways, particularly in the realm of psychotherapy. In the area of diagnosis, they have emphatically endeavored to establish themselves as "consultants" on an equal footing with referring professionals (e.g. Crary & Steger, 1972; Tallent, 1983). This concern for status has most often centered upon psychologists' relationships with psychiatrists (Shakow, 1949; Ausubel, 1956; Brody, 1956, 1959; Sheckman & Harty, 1982).

#### The Referral Process as an Interpersonal Relationship

As stated previously, the use of psychological tests has often been a focal point of defining inter-professional relationships. Bellak (1959) described psychological testing as "the main area of contact between psychiatrist and clinical psychologist" (p. 76). Brody (1959), who was writing as a psychiatrist, redefined the question of referral for psychological testing as a question of referral for the consultation of a psychologist:

The issue, then, is not why a patient should or should not receive psychological tests. It is why the psychiatrist does or does not refer his patient to a psychologist. (p. 88)

Brody underscored what may be called the social psychology of the psychological referral:

The psychiatrist's request for psychological testing is not simply a dispassionate, objective effort to obtain scientific evidence which will be useful in his attempts to understand and treat his patient. It is a social process involving his attitude toward psychologists, toward psychological testing, his identification with his own profession, and his insecurities about his own status. (pp. 88-89)

Brody asked, "What are some of the hidden functions of psychological test referral for the psychiatrist?" (pp. 89-90) He suggested that sometimes psychiatrists wanted to impress their patients with an approach that appeared to be scientific. Also, he noted that

Psychological testing provides the clinician with a rapid feedback without risking some loss of status in his patients' eyes by asking for consultation from another physician. (p. 90)

Another hidden function of the test referral, suggested by Brody, was the use of the consultation for the psychiatrist's defensive needs, such that the psychologist could become the target of the patient's aggression or libidinal desire.

Brody also speculated about hidden functions of the psychiatrist's decision not to refer a patient for testing. The delay in decision making due to testing may have detracted from the psychiatrist's sense of prestige as a professional who acted immediately. Also, a psychiatrist maybe would feel unconscious guilt in "trapping" a patient to reveal information about himself in a way through which he was unaware, i.e. by testing. Another reason for not referring for testing was that a psychiatrist perhaps would see the nonobjective nature of testing as meaning that he was merely requesting the clinical judgment of a psychologist; some psychiatrists would not want to take such a judgment above their own. Brody suggested a final reason for non-referral: to reduce the overload of information processing for the psychiatrist, who already would have plenty of data to integrate.

Brody summed up his discussion:

In summary, we are dealing here not with the established value of psychological tests as behavior-sampling devices. We are dealing with the generally unidentified factors which may result in the psychiatrist's depriving his patient of their benefit, or in his developing an unwarranted dependence on them. Some of these factors are economic. There are also circumstances in which avoidance of psychological testing or unusual dependence upon tests is justified and rational. Many of the factors, however, are involved in the psychiatrist's perception of the person

of the psychologist and of the nature of his tools. These influence the social process which is referral. (1959, p. 91)

Schafer (1954) also highlighted salient "nonscientific" factors in the testing referral. He discussed some common problems in referrals from psychiatric residents:

To some extent, and not always subtly, the residents ambivalently transfer to the tester major or full responsibility for clarification of diagnostic, dynamic, prognostic, therapeutic or dispositional problems in the case. It becomes the tester's "job" to settle the problematical issues. This difficulty becomes especially acute when cases are not routinely tested but are referred for testing only where major confusions or uncertainties exist in the psychiatrist's mind. (p. 10)

In general, he noted that the psychologist may be overvalued or undervalued in a given clinical setting. Both situations would affect his attitude towards his task and, ultimately, the responses of the patient.

Such comments advocate an increased awareness of the subtleties involved in a referral. In conjunction with this, psychologists have also urged each other to improve their written response to these various demands. Sargent (1951) presented different styles of test report writing and concluded with several questions, including

What is the purpose of a psychological test report? Do we who write them share the purposes of those who read them? If not, who is to educate whom to their fullest usefulness? (p. 186)

She answered her own questions:

The writer's position regarding all of these questions is that few can be answered outside the context in which a test is requested, given, and reported. (p. 186)

Sargent advocated improved communication, as did Klopfer (1959): "In a real sense the sole and only legitimate purpose of the psychological report is that of communication" (p. 88).

Thus, taken in a historical context of interprofessional conflict and problems in communication, the psychological testing referral can be seen as a complex social psychological event. A review of the vicissitudes of testing, from critics and proponents, will be followed by a summary of the empirical study of the prevalence of assessment. This in turn will be followed by a description of the processes and factors that may influence the referral situation. Finally, this second half of the literature review will conclude with an examination of the empirical research on the use of test reports.

#### The Reported Decline of Testing

The decline of psychodiagnostic testing as a task performed by psychologists has been much discussed, although not very well documented empirically, in the literature. Evidence for psychologists' devaluation of testing has been generally indirect, but nonetheless accepted by many authors. One of the first comments

about this trend was made by Carson (1958), in his letter to the editor of the American Psychologist. He voiced his concern that psychologists were devaluing their roles as diagnosticians by frequently relegating the assessment task to trainees.

A review of published research reveals a decline in articles related to diagnostic testing. Crenshaw, Bohn, Hoffman, Matheus and Offenbach (1968) comprehensively surveyed ten journals over an eighteen year period and concluded that the use of projective techniques in research peaked in 1955, dropped sharply in 1956 and 1957, and then stabilized through 1965. Tolor (1973) documented the declining interest in psychodiagnosis by reviewing five major clinical journals between 1951 and 1970. He found consistent significant decreases in the number of diagnostically related articles. These decreases were especially striking when compared with the general trends of scientific literature. Most scientific topics have exhibited an exponential growth in related publications (Price, 1963; MacRae, 1969; Griffith, Small, Stonehill & Dye, 1974; Small & Griffith, 1974).

What is most salient is not the decreased amount of testing activity that may or may not be inferred from the studies of Crenshaw, et al. (1968) and Tolor (1973). The implied devaluation of testing from psychologists' points of view is what is of interest here. The present section

of this review is concerned with highlighting the rhetoric of the assessment controversy; empirical studies will be examined further on.

As an example, the following comment by Blatt (1975) suggested that such a devaluation had become commonplace:

Testing remains, by and large, a second-class clinical activity rather than a highly valued way of studying cognition, perception, affect and the representation of interpersonal relationships that can contribute to both the clinical process and the systematic investigation of clinical phenomena. (Blatt, 1975, p. 328)

The reasons for this devaluation have been offered by a number of authors (Mosak & Gushurst, 1972; Tolor, 1973; Cleveland, 1976; Lewandowski & Saccuzzo, 1976; Ivnik, 1977; Korchin & Schuldberg, 1981). In the most comprehensive review, Korchin and Schuldberg (1981) listed seven "themes" in the decline of interest in diagnostic testing: (1) new alternative roles have emerged for psychologists; (2) some theoretical orientations, such as the behavioral and humanistic-existential approaches, have opposed testing on principle; (3) psycho-diagnosis was initially overrated; (4) the teaching of assessment skills to students declined, due to inexperienced or disinterested faculties; (5) assessment is "difficult work that can strain the clinician's identity" (p. 1149);

(6) society has challenged the use of tests; (7) testing is expensive or not sufficiently profitable. To this list could be added factors mentioned by the other reviewers: negative research findings; overemphasis on pathology in reports while downplaying personality strengths; beliefs that reports are ignored; the low quality of psychodiagnosticians. Of these themes, the main emphasis in the literature has been the critique of testing's lack of utility compared to its expense.

#### In Defense of Assessment

The mounting criticism and apparent decline in testing have prompted a continuing defense of psychodiagnostics by psychologists (Zubin, 1951; Klopfer, 1962, 1964; Rosenwald, 1963; Towbin, 1964; Tallent, 1965, 1983; Holt, 1967; Appelbaum, 1970, 1976, Gough, 1971; Craddick, 1972, Freedheim, 1972; Weiner, 1972; Blatt, 1975; Levy & Fox, 1975; Korchin & Schuldberg, 1981). The content of these rebuttals will not be detailed here. A sample of comments will serve to represent their positions.

Some writers have attacked the critics' arguments directly. Rosenwald (1963) attributed much of the criticism of the psychodiagnostic role of psychologists to the critics' discomfort with the "nonphysical, artistic and intrusive processes" (p. 237) involved in assessment. Ivnik (1977) noted that most criticisms of testing have referred to the



misuse of tests rather than to the utility of the tests themselves. He asserted that opinions by other professionals as to the value of psychological reports were not adequate criteria for assessing the value of tests themselves.

Other writers have reiterated the raison d'etre for testing. In the current Comprehensive Textbook of Psychiatry, Carr (1980) noted that psychological tests had two common features that made them advantageous over an interview: (1) they provided

a fairly objective means for comparing a relatively controlled sample of the patient's behavior with available normative data representative of a larger reference group (p. 940)

and (2) they could elicit responses by a patient to "a broad range of stimuli on the continuum of structure-ambiguity" (p. 940). In contrast to these behaviorally phrased arguments, most published proponents of testing have spoken from a psychodynamic point of view, e.g. Allison, Blatt and Zimet (1968).

Regardless of theoretical orientation, supporters of assessment have often deflected criticisms of the tests themselves with the common argument that assessment was a refined skill that leaves itself open to criticism when the results of mediocre practitioners were examined. As Klopfer (1962) emphatically stated:

There is no aspect of the job of the clinical psychologist which is more complex, more difficult, and more challenging. (p. 298)

Thus, from such a perspective, the object of criticism should be the incompetent assessing psychologist rather than assessment per se. This raises the empirical question of how many competent assessors are in practice, assuming such competency could be validly assessed. If such an evaluation could be made and there turned out to be very few psychologists truly competent to make the best use of assessment, then the value of assessment would still be questionable. Assessment would then properly belong only in the hands of an elite few, unless resources could be mobilized to transfer such skill on a more widespread basis. Theoretically, criteria for graduate degrees and professional licensure exist to insure such competency in assessment. However, the possibility remains that the use of these criteria has not succeeded in producing the caliber of psychodiagnosticians desired by the profession. Whether or not this hypothetical situation accurately describes the current state of affairs in clinical psychology is a question beyond the scope of this review.

Returning to the literature that has defended the use of testing, the reader will discover that, in general, supporters of testing have not denied the problems raised by critics. The most common review articles on testing

have acknowledged the "decline" or "uncertain status" of testing and have offered new approaches to improve the use of tests (Holt, 1967; Neuringer, 1967; Appelbaum, 1970; Hertz, 1970; Gough, 1971; Beutler, 1973; Lewandowski & Saccuzzo, 1976; Ivnik, 1977; Petzelt & Craddick, 1978; Sloves, Docherty & Schneider, 1979; Korchin & Schuldberg, 1981; Sterling, 1982). The many ideas of these authors will not be detailed here. What is significant, however, is that their discussions have demonstrated a continued desire by those who support assessment to address the limitations of assessment and to adapt to new clinical problems.

#### Empirical Study of the Prevalence of Testing

In response to the widespread notion that the use of testing has declined in clinical psychology, a number of surveys have been taken in the past decade. Prior to then, most surveys dealt with comparative frequency of use of particular tests, e.g. Sundberg (1961); Lubin, Wallis & Paine (1971). Sundberg (1961) did extrapolate from his results to estimate that, annually, nearly 7,000 clinical and counseling psychologists were seeing 1,300,000 persons for various services, including the administration of psychological tests to 700,000 clients. These figures did not include industrial, military, educational and private clinical settings; thus, they

represented a fraction of total services. The more recent surveys have endeavored to look at the role of assessment in psychologists' daily work.

Garfield and Kurtz (1974) reported, in a survey of one-third of the members in the Division of Clinical Psychology of the American Psychological Association, that an average of 10% of the professional clinical psychologist's time was devoted to diagnostic assessment. This was compared to time devoted to three other activities: individual psychotherapy (25%); teaching (14%) and administration (13%). Ivnik (1977) observed that this survey indicated that 24% of psychologists' direct clinical services was devoted to assessment. He argued that, although assessment may have been taking a smaller proportion of clinicians' time than previously, it still occupied a significant portion of their professional work.

Levy and Fox (1975) surveyed clinical employers who had advertised job openings during the 1971-72 year. They found contrary to some discussions of the decline of diagnostic work, that testing was "alive and well." In fact, 90.5% of their respondents indicated that they expected job applicants to have skills in testing.

Wade and Baker (1977) surveyed five hundred clinical psychologists about their attitudes toward and use of psychological tests. They reported that "the great

majority of responding clinicians devoted substantial time to psychological testing" (p. 879). This was true for clinicians of every theoretical orientation. The most endorsed reason for testing and for recommending that students learn testing was information that testing could provide on personality structure. Clinicians reported that their personal experience with tests was the main reason contributing to their decisions to use tests for a given client. This experience apparently was more influential than the negative research findings reported in the literature. Clinicians often reported that they used their own personalized assessment methods and that testing was an insightful process rather than an objective skill. Another reported factor in the use of tests was the lack of an appropriate alternative for assessment. Clinicians also justified their continued use of tests with methodological criticisms of assessment research.

Reynolds (1979) studied the relationship between frequency of test usage and "overall quality of psychometric refinement" (p. 326), as rated by 31 academic psychologists. These quality ratings were applied to the ten most frequently used tests. Although a significant rank-order correlation was found for data based on a 1974 survey, only one-third of the variance in test usage was accounted for by test quality. This supported other findings, e.g. Wade and Baker (1977), as to the importance of nonscientific reasons for the use of tests.

In another study, Piotrowski and Keller (1978) surveyed 93 clinics and mental health centers in the Southeastern United States. Responses from 61 centers indicated that testing was an important function in these outpatient facilities. They contrasted this fact with the reported decline in emphasis on assessment training in doctoral programs (Shemberg & Keeley, 1970; Thelen & Ewing, 1970; Thelen, Varble & Johnson, 1968).

The above review suggests that a number of psychologists vigorously support assessment, which continues to be a significant activity for many in the profession. Forces in opposition to testing also continue to press their case. Thus, no dogmatic conclusions about the outcome of this controversy would be appropriate. Howes' (1981) conclusion about the use of the Rorschach can be applied to testing generally:

It would be quite apparent that the use of the Rorschach is dying hard, if it is dying at all, although there are undoubtedly many who are eager to attend the funeral. (p. 347)

Korchin and Schulberg (1981) gave a somewhat mixed message in their review. They concluded that assessment will continue to play a vital role in clinical activity, but, in answer to the question as to whether clinical assessment was alive and well, they responded, "Alive, without doubt, though only in moderate, but we believe

improving, health" (p. 1156). Ivnik's conclusion was probably the most accurate:

Few definite conclusions can be drawn regarding the status of psychodiagnostic testing in the field of clinical psychology today. Many divergent opinions exist, none of which has received unequivocal research support. It would appear that despite its detractors, psychological testing continues to be a frequent activity for the clinical psychologist. (1977, p. 212)

Having established that psychological assessment is a controversial yet important clinical activity, this review will now examine the referral process more closely. The following section will suggest the many reasons for referral and factors that potentially influence the referral process. Included will be the author's perceptions of a general dichotomy in approach to assessment and then an emphasis on confirmation in the service of increasing the referrer's confidence in judgment as a primary usage of psychological reports.

### The Referral Process

#### General Description

The process of psychological consultation has received some attention in the research literature, e.g. Rhodes (1974); Swenson (1974). However, detailed analyses of the step-by-step procedure of what is involved in a

referral for consultation have seldom been discussed. The referral may be seen in terms of a series of steps involving interaction between the referrer, consultant and client. The author offers a general conception of these steps in Figure 1. It should be noted that this outline omits potential active involvement of the person being assessed, e.g. in the sharing of test results. This omission is for purposes of economy in the current discussion. The author otherwise endorses such active involvement and open communication with the client, as has been discussed by several psychologists (Richman, 1967; Fischer, 1970, 1972; Brodsky, 1972; Mosak & Gushurst, 1972; Craddick, 1975; Tallent, 1983).

### Reasons for Referral

An Overview. The third step in Figure 1, concerning the referral question, has received much attention in the literature. Table 1 is presented as a list representative of possible reasons for referral. The list is probably not exhaustive, but should encompass most referrals, based upon the author's experience and survey of the literature. Listed reasons are not meant to be mutually exclusive, as will become clear in the following discussion.

The first type of referral, the seeking of new information about the client, is the most common and plainly stated reason in virtually all referrals. A second reason,



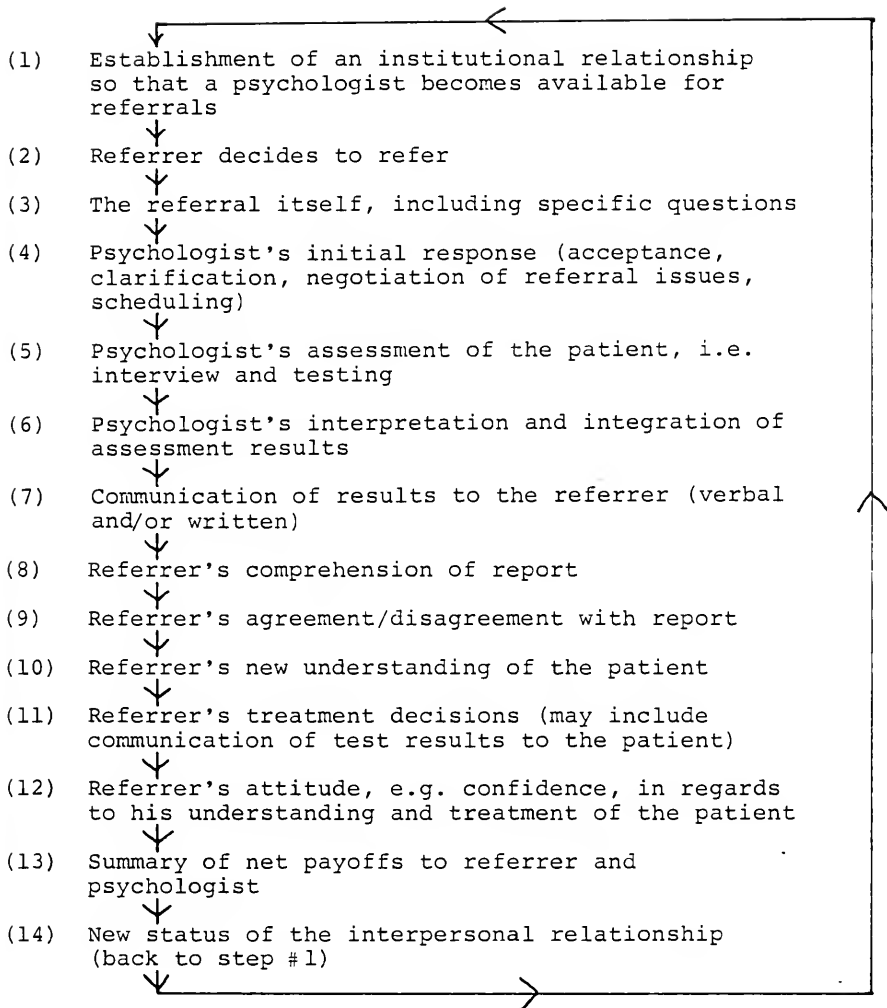


Figure 1. The Referral Process for Psychological Consultation.

seldom officially admitted, is that of bureaucratic requirement (Section "II" of Table 1). Psychological testing is done merely because of routine and not because the referrer or consultant believes it to be of intrinsic value for the referrer or the patient. An example of this might be a situation in which a test report is required as part of an application for admission to a facility, even though the test report itself contains no new material on the client. Testing for purposes of staff training would also fit into this category. Another reason for testing would be as a therapeutic intervention itself (Section "III"). This function has been gaining notice in recent discussions (Richman, 1967; Mosak & Gushurst, 1972; Tallent, 1983). A fourth reason for referral is for confirmation of current clinical judgment. The referrer seeks consensual validation for what he already believes. He is less interested in new information and more interested in redundant feedback, although he or she may not be aware of this. Confirmation would serve to simplify the clinician's judgment and probably would be reinforcing or anxiety reducing for him or her. This reason, which will be elucidated further below, would not be stated explicitly in a referral. It could refer to any of the aspects of clinical judgment mentioned in the first section of Table 1.

Table 1.  
Potential Uses of a Psychological Consultation

- I. Acquisition of new information
  - A. Descriptive diagnostic issues
    1. Differential diagnosis
      - a. DSM-III categories
      - b. Other dimensions, e.g. psychoanalytic
    2. Intellectual evaluation
      - a. IQ
      - b. Assessment of particular cognitive abilities
    3. Detection of organicity (neuropsychological assessment)
    4. Detection of thought disorder
    5. Assessment of psychodynamics (structure, defenses, conflicts)
    6. Assessment of level of depression
    7. Assessment of etiology
    8. Assessment of psychosexual level of development
    9. Assessment of impulsivity
    10. Assessment of sexual orientation
    11. Assessment of possible malingering
    12. Assessment of assets, strengths
    13. General description of patient
  - B. Predictive diagnostic issues related to treatment
    1. Prognosis
    2. Choice of treatment
    3. Progress in current treatment
    4. Assessment of suicide potential
    5. Assessment of potential transference reactions or defenses likely to surface in treatment
    6. Vocational/educational recommendations
- II. Required bureaucratic procedure
- III. Therapeutic procedure
- IV. Confirmation of the referrer's judgment

Utility versus Understanding. The preponderance of literature on assessment has been devoted to the issues listed under the first section, acquisition of new information. Although not precisely reflected in the divisions of Table 1's outline, a dichotomy of attitudes towards the purpose of assessment can be gleaned from an examination of the debate about the value of testing. As will be demonstrated, this dichotomy, which may be grossly contrasted as "utility" versus "understanding," is a subtle one that has not been previously clarified in the literature. The distinction of those who see assessment for utilitarian purposes from those who advocate the broader goal of "understanding" is a matter of degree. Both groups emphasize the new information that assessment brings. Those who fit in the "understanding" category would say that what they seek in testing is not only that which will be immediately useful, but also information that may not reveal its usefulness until a later time during the course of therapy. The differences between the two groups lie in what is considered useful, which is often a reflection of theoretical orientation. The meaning of the two approaches will be explained through description and example.

A number of advocates of revision in testing have fit clearly in the utilitarian group (Fulkerson, 1965; Cole & Magnussen, 1966; Breger, 1968; Arthur, 1969;

Lanyon, 1972; Howe, 1981) Fulkerson (1965) stated, "If tests are not being used to make decisions about patients, they should not be given" (p. 193). He recommended that tests should be developed by individual clinicians to correspond to criteria related to the specific decisions that need to be made. Cole and Magnussen (1966) proposed replacing traditional assessment with a model devoted to disposition and specific treatment for the patient. Thus, disposition and action would be the ultimate goal of the assessment, rather than diagnosis. The authors pointed out that such an orientation would serve to break down the artificial barrier between diagnosis and treatment. Such a "heuristic" approach, based on decision making theory, would take the context of the assessment into account and would involve the limited number of dispositions that are available in real situations. Arthur (1969) and Howe (1981) endorsed this approach, as did Lanyon (1972), who outlined an assessment scheme designed for efficiency.

Some observers who have not necessarily advocated this approach have acknowledged its impact. Klopfer (1968) observed that the typological interests of Rorschach and Jung were no longer as salient to practitioners as were concerns for predictive efficiency:

Thus the matter of predictive efficiency is more highly valued today as a criterion of validity

than is the concurrence of one test interpretation with another. (p. 404)

Tallent (1983) similarly described a recent trend in psychological reports:

Many reports now are sharply focused and highly prescriptive. What is the problem, and what can we do about it? In this tradition the behavior-oriented report is strictly a no-nonsense, business-like approach. Psychopathologically oriented reports increasingly reflect the fact that certain classes of medication target certain behavior symptoms. (p. 11)

One type of utilitarian approach has been the accommodation of testing to psychiatric classification. As psychiatry establishes a diagnostic category, psychological tests are soon applied in an attempt to develop test criteria for identification of patients who fit the new category. For instance, with the establishment of criteria for the borderline personality disorder in DSM-III, Snyder, Pitts, Goodpaster, Sajadi and Gustin (1982) developed an MMPI profile to discriminate a group of borderline patients from a group of dysthymic patients. Millon's (1982) inventory, applicable to Axis II of DSM-III, may also be seen as a direct response to a new nosology.

As contrasted to the emphasis on utility, some authors have stressed the need for a more general understanding of the person as the purpose of psychological

testing (Allison et al., 1968; Hertz, 1970; Holt, 1970; Weiner, 1972; Schwartz & Lazar, 1979).

Tallent (1965) noted a switch in emphasis from the goal of prediction to that of understanding: "The role of prediction is downgraded from its historical position of scientific royalty" (pp. 429-430). As noted above, however, his most recent review suggested that this trend was only temporary (Tallent, 1983). His earlier summary expressed the concept of interest here. At that time he cited the rise of construct validity as a more meaningful standard for clinical practice than predictive validity. Also, he quoted Cantril, Ames, Hastorf, and Ittelson (1949):

Emphasis on prediction alone can easily obscure the more fundamental aim of science covered by the word understanding. (Tallent, 1965, p. 430; emphasis in original)

Weiner (1972) referred to understanding as knowing "what people are like and how they got to be that way" (p. 536). He advocated the incorporation of personality constructs as intervening variables that generated explanations for behavior which may or may not be predictable. He stated that the purpose of psychodiagnostic assessment was the appraisal of personality processes rather than the prediction of behavior.

As a part of the continuing debate over the value of clinical versus statistical prediction, Holt (1970)

voiced a similar opinion. He observed that, in his view of science, prediction was not an end in itself, but a means to a larger end, which was "explanation through understanding" (p. 346). He acknowledged that other authors viewed prediction and control as the fundamental goals of science. This division of opinion, according to Holt, could be further linked to the basic dispute between phenomenology and positivism.

Schwartz and Lazar (1979) may have been speaking of the same dichotomy as they distinguished semantic interpretation from probabilistic interpretation. The former approach attempted to ascertain meaning, whereas the latter was interested in prediction. The authors asserted that projective assessment was based on semantic interpretation and should not be judged according to statistical, predictive criteria.

The aim of the clinician is to understand and describe, not to predict some future event. Prediction requires random sampling; description requires informal observation (p. 7). . . . [T]he clinician's tool is psycholinguistics, not statistical inference. When dealing with the responses of patients, we can rightfully speak of the art of clinical psycholinguistics. This approach applies to any clinical enterprise that attempts to understand another person's thought through his verbal productions. (Schwartz and Lazar, p. 11)

Similarly, Allison et al. (1968) described the purpose of an evaluation:



We believe that the major role of psychological testing should not be prediction, but, instead, the development of hypotheses which best describe the given individual and which constitute a "best fit" for the data and the observations which are available. The broader purpose of these hypotheses is to enrich and extend the understanding of the patient or client and to make available dimensions of the individual which may have been unavailable or not considered previously. (p. 2)

Such an emphasis on understanding over specific prediction should not be misconstrued as a rejection of utility altogether. Particularly in a long term intensive treatment situation, most statements made in a clinical test report can be construed as relating to useful clinical predictions. For example, Applebaum's (1977) description of the Menninger Foundation tradition of psychological testing embodied this approach. Following in the legacy of Rapaport and Schafer, Menninger psychologists have given essentially the same extensive battery of tests to each patient. This has included the Wechsler-Bellevue, Rorschach, TAT, Word Association, BRL Sorting and Story Recall (p. 24). A comprehensive report has been composed for each patient. The Menninger method would appear to be the prototypic descriptive evaluation that attempts to map the internal psychological functioning of an individual. Specific utility would seem to be a secondary consideration. Appelbaum, however, insisted that testing should be done with predictions for treatment as the foremost goal:

In our view, the true measure of "the validity of tests" is how well the psychologist who uses them is able to make useful clinical predictions. (p. 244)

What Appelbaum was referring to reflects the fact that a comprehensive, descriptive report will be utilitarian in such an intensive treatment context. In a different setting, such as the community clinic, the limited treatment options will not enable as much use to be made of such extensive interpretive descriptions.

The question of utility, then, is a matter of the nature of the treatment situation. Tallent (1965) cited Payne's (1958) ideas that pointed to a major refocusing of the validity question to that of utility:

Stated simply, the question is "What sort of clinical information is needed (The answer to this apparently simple question is not yet at hand and has hardly been investigated), and can the psychologist supply it?" (p. 430)

Tallent (1983) saw the ideal "case-focused" report as being able to answer Paul's (1967) specifications:

What treatment, by whom,, is most effective for this individual with that specific problem, and under which set of circumstances. (p. 111; emphasis in original)

Tallent added two additional requirements: "understanding of the person and the "why" of the report's recommendations (1983, p. 26; emphasis in original).

In setting up such an ideal, Tallent seemed to gloss over the distinction made here of the "utility" versus "understanding" groups. However, this synthesis would be appropriate if one recognizes that different referral situations require different levels of understanding and criteria for utility.

Communication centered upon the referral questions thus seems to be the essential problem of the assessment. Few authors in this area would dispute the following recommendation:

The usual referral for psychological testing should clearly specify exactly and what the referral source hopes to obtain from the consultation. The real value of the psychologist's report will lie in the emphasis which is placed on answering the specific referral questions. (Harlow & Salzman, 1958, p. 231)

This implies that the referrer would do well to be fully aware of his needs and that the assessing psychologist would also be aware of these. This follows the earlier discussion of the referral process as a social psychological event. As applied to the utility/understanding dichotomy, referrers will likely be placed into one group or another, reflecting the sort of questions they would like to have answered.

Confirmation and confidence in judgment. As previously stated, referral for testing has been influenced

by factors other than a rational, scientific search for new information. The referral is an interpersonal event multi-determined by the personalities involved, including especially the needs of the professional who requests the consultation. The main hypothesis of this study is that a primary factor in requesting a consultation is the desire by the referrer to have confirmed what he or she already believes, thereby increasing his/her confidence in his/her judgment. This corresponds to the fourth use of psychological reports listed in Table 1.

Very few writers have mentioned this use of an assessment. Allison et al. (1968) endorsed the worth of tests as confirmatory evidence because of the reassurance value for the staff to make decisions with more confidence. Rosenwald (1963) acknowledged the value of confirmation, but gave it secondary status:

A psychological report should partly confirm clinical findings already known, but in large part it should provide new information uniquely available through it at that time and not confirmable until perhaps much later. Finally, it may provide information which is perhaps not directly confirmable in any other way at any time. (p. 229)

Fulkerson (1965), who advocated an immediate decision making approach for assessment, did not see much value in assessment as confirmation. However, his analysis

suggested that much of assessment was done precisely for this purpose. He stated that the "mapping of an internal state of affairs" (p. 193) was, at best, only an intermediate goal. If this were the ultimate goal of testing, then a psychiatrist's judgment could serve as the validating criterion. Fulkerson concluded that the only value of testing in such a case would be to increase the psychiatrist's confidence in his judgment.

The point here, contrary to those who regard assessment from Fulkerson's perspective, is that this confirmatory function is valuable in itself. Appelbaum (1977) expressed this proposition in his discussion of assessment at the Menninger Foundation:

Tests have become a routine part of personality assessment at the Menninger Foundation as a result of many years of empirical observation of their utility. One of their uses is to corroborate other clinical judgments, an important though perhaps undramatic contribution. Clinical responsibility demands a high degree of confidence in the making of decisions that may profoundly affect other people's lives, and corroboration of inferences increases the confidence with which inferences and the recommendations based on them are held. (p. 2)

The proposed study will focus upon this "undramatic contribution." As discussed in the first part of this review, the seeking of confirmatory information is a common judgmental bias which has been found to correlate with an

increase in confidence in that judgment. Before proposing how this phenomenon will be tested, the present discussion will return to a summary of factors that influence the referral process and then a review of empirical research on the use of psychological reports.

#### Factors Influencing the Use of Assessment

A salient feature of the schema in Figure 1 is that it begins and ends with the status of the interpersonal relationship. This is a recognition of the point already emphasized, that is, the interpersonal aspects of the process.

Given that the referral process represents a complex interaction between individuals and, sometimes, institutions, then one can assume that numerous factors influence whether or not a referral is made for a psychological consultation. In view of a lack of any previous comprehensive literature discussion of factors that influence testing, the author offers a list of common influences (Table 2). Again, this list is not exhaustive. It is presented to further illustrate the complexity of the referral process. It should be noted that most of the options in Table 1 can be subsumed under part "C" of section "III" in Table 2.

Table 2.  
Factors That Influence the Use of Psychological Assessment

- I. Factors related primarily to the referrer
  - A. Referrer's attitude toward psychological testing
    1. Referrer's attitude toward clinical psychology
    2. Referrer's knowledge of testing
  - B. Type of treatment performed by referrer
  - C. Referrer's theoretical orientation
  - D. Referrer's experience
  - E. Referrer's self-confidence
  - F. Appropriateness of referrer's referral questions
  - G. Quality of communication between referrer and psychologist
  - H. Type of clinical setting where referrer works
- II. Factors related primarily to the consultant (psychologist)
  - A. Availability for referrals
  - B. Consultant's attitude toward testing
  - C. Consultant's theoretical orientation
  - D. Consultant's expertise in assessment
  - E. Quality and utility of previous reports
    1. Clarity in report writing
    2. Extent to which referral questions were answered
    3. Time lag in answering previous referrals
  - F. Time constraints for assessment
  - G. More profitable or desirable alternative activities, e.g. psychotherapy
  - H. Quality of communication between referrer and consultant
- III. Other factors
  - A. Cost and affordability to the patient
  - B. Testability of the patient, e.g. willingness
  - C. Type of patient, i.e. specific problem; referral questions
  - D. Scientific validity of tests

## Empirical Research on the Use of Psychological Reports

### Introduction

The empirical research reviewed thus far has concerned mainly the percentage of psychologists' time devoted to assessment. Turning to the empirical study of psychological reports themselves, the limited literature may be divided into surveys and studies that directly used reports. Before examining these studies, an important observation must be made. Compared to the amount of published research on particular psychological tests such as the MMPI or Rorschach, the empirical study of psychological test reports is a sparse literature. In 1956, Tallent observed that much clinical time was devoted by psychologists to preparing reports. In contrast, he declared:

It is surprising then, that investigators concerned with the elicitation of psychological data number in the many thousands whereas only a handful of writing have dealt with the problem of how such data may be communicated so as to be maximally useful. (p. 103)

Almost three decades later he observed that there was still a "dearth" of research on the utility of psychological reports. He noted that the few available studies showed contradictory evidence that was open to multiple interpretation. This was based in part on the widely



varied research designs employed (1983, p. 19).

Bellak's statement that "amazingly little systematic attention has been paid to the writing of the psychological report" (1959, p. 76) remains applicable today.

### Surveys

Much of the reported research on the use of psychological reports has taken the form of surveys, usually conducted by mail (Tallent & Reiss, 1959a,b,c; Lacey & Ross, 1964; Hinkle, Nelson & Miller, 1968; Mintz, 1968; Moore, Boblitt & Wildman, 1968; Smyth & Reznikoff, 1971; Olive, 1972; Wiedemann & Mintz, 1974; Tallent, 1983). In most of these studies, professionals who use reports, such as psychiatrists, social workers and other psychologists, have been asked to rate the value of reports generally. Data have been gathered as to how frequently reports are used, what sorts of information reports provide and what shortcomings they typically have. Most of these studies have been essentially uncontrolled attitude surveys. Respondents have given their opinions about typical reports without having to consider any report in particular. An exception was Mintz' (1968) survey of twenty-five student therapists who rated the value of reports in current psychotherapy caseloads. These subjects presumably could think in terms of specific reports that they had read in

regards to a limited number of current patients. However, even this study was relying upon general opinions, influenced by biasing factors, such as inaccurate memory. There was no control or consideration of the reports themselves. The limitations of this survey method were seldom acknowledged by authors. Not only were the usual pitfalls of self-report research not mentioned, but also there was a noticeable omission of any discussion concerning potential demand characteristics of these surveys. Such discussion would seem to be appropriate because the surveys involved critical evaluation of an important activity of the surveyors' profession. It is reasonable to hypothesize different results if the subjects had believed that the studies were conducted by persons who were not psychologists.

Setting aside these methodological problems for the moment, one interesting conclusion emerged from two independent studies. Smyth and Reznikoff (1971) and Wiedemann and Mintz (1974) each concluded that reports functioned mainly as confirmation of diagnostic judgment. This finding supported the main hypothesis of the current proposal. However, given the methodological drawbacks already mentioned, one must be cautious in asserting that most or many reports are indeed used for this purpose.

### Experimental Studies

There are a few studies that have attempted to use psychological reports in an experimental manner. The author could locate only ten such references that had subjects directly use reports. Of these, eight attempted to measure what Meehl (1959/1977) has termed "incremental validity." This refers to the extent of new information that a reported procedure, in this case psychological assessment, provides. Meehl's discussion of incremental validity was a part of a larger exposition on the concurrent validity of psychological tests. He posited four levels of concurrent validity in the "phenotypic" characterization of a person. By "phenotypic," he was referring to the "descriptive or surface features of the patient's behavior, including his social impact" (p. 91). These levels varied in terms of the rigorousness in criteria for validity. The eight studies in question here generally aspired to measure the utility of test reports according to Meehl's second level of concurrent validity, which he described in the form of a question:

To what extent does the test enable us to make, reliably, accurate statements which we cannot concurrently and readily (that is, at low effort and cost) obtain from clinical personnel routinely observing the patient who will normally be doing so anyway (that is, whose observations and judgments we will not administratively

eliminate by the introduction of the test)?  
(Meehl, 1959/1977, p. 104; emphasis in original)

Thus, this level of validity refers to what unique information may be derived from a psychological report. Meehl's first level of concurrent validity, or "validation demand" (p. 106), referred simply to the accuracy of information, regardless of whether it added to the user's knowledge. The third level asked, in addition to the criteria and uniqueness of the first two levels, "how much earlier in time" (p. 106; emphasis in original) that the psychometric assessment enabled judgments to be made. The fourth and most practical, incorporated the criteria of the previous three with the additional standard of how much the procedure helped in the treatment of the person being assessed. Meehl observed that, to his knowledge, no validity studies of psychological tests had been carried out at the third and fourth levels; very few had been attempted at the second level.

One striking aspect about the eight studies to be discussed is their variety of methodology. No two are alike, either in overall design or in specific criteria used to assess the utility of reports. Also, with one or two exceptions, these studies have suffered from significant methodological flaws or have been published in such sketchy form that a complete critical analysis

is not possible. Keeping in mind their diversity and limitations, they can be examined as follows, grouped loosely into five different types.

Two studies can be grouped together because of their survey methodology. They are distinguished from the survey studies already reviewed in that they involved either the ratings of actual reports by clinicians who use reports or ratings of material derived from actual reports. This is in contrast to the other survey literature which assessed general attitudes without reference to particular report material.

Affleck and Strider (1971) sampled a total of 340 psychological reports on inpatient and outpatient services for both children and adults. Referring sources were surveyed upon receipt of the reports and then again two months later, when a more extensive interview was conducted. The surveys revealed that about two-thirds of the referral questions were answered by "new and significant information" (p. 177) or with confirmatory information that was previously suspected but not well established. Followup interviews revealed that 52% of the reports provided information that resulted in some modification of patient management or disposition. Also 18% of the reports confirmed current thinking about the patient. In addition, 4.5% were vaguely "helpful"; 2.3% had "minimal effect"; 18% had no effect;

1.9% were "erroneous, incomplete or detrimental"; 3.5% of the reports could not be completed because the patient went "AWOL" or quit (p. 178). Affleck and Strider cautioned that the usefulness of reports was evaluated by co-workers of the psychologists; as colleagues, they may have been biased in favor of the reports. The authors concluded that, considering an average of three questions for a typical psychological referral, there was a high probability of useful information derived from psychological reports.

Hartlage and Merck (1971) found that when psychologists were made aware of what sorts of statements were valuable to the users of reports, their future reports improved in utility for their users. The study involved psychological reports from a large Goodwill Industries rehabilitation facility. Their criterion for utility consisted of ratings of reports, on a five-point scale from "very poor" to "very good" (p. 460), by the supervisors for whom they were intended. The authors concluded that psychologists tended to "grind out reports with good theoretical consistency but little decisional value" (p. 460) until they became familiar with the actual uses of reports by those for whom they wrote.

A second type of methodology used for the assessment of incremental validity was an archival examination of

reports in the files of a particular agency. Two studies employed this method, albeit in quite different ways using different criteria.

Adams (1972) assessed the impact of psychological evaluations by comparing them with the initial and final psychiatric diagnosis in 137 psychiatric inpatients' records. His criterion for utility thus depended on his own judgment as to the similarity of psychological report conclusions and the different psychiatric diagnoses. He concluded that the evaluations had an impact on no more than 16% of the diagnoses. In 52% of the cases, the report was judged to have either confirmed the diagnosis or was consistent with it without definite statements of confirmation. In 30% of the cases, the report's disagreement with the initial diagnosis was not influential in changing the diagnosis. Adams concluded from these statistics that psychological evaluations did not have much influence on psychiatric diagnosis. He suggested that the data supported Gauron and Dickinson's (1966, 1969) findings that psychiatrists derived emotional comfort from confirmation of their diagnostic opinions by psychological reports. Adams admitted that psychological evaluations provided more information than diagnosis. . He speculated that psychiatrists found psychological evaluations useful for information not requested in the referral.

He also suggested that psychiatrists were not aware of the effort involved in a psychological evaluation. In conclusion, Adams recommended that referrals for psychological evaluation should be accepted only if there can be anticipated tangible results. He suggested two criteria for referral questions:

(a) there is more than one answer possible and that (b) different alternatives can be expected to lead to importantly different behaviors on the part of the referring person. (p. 566)

Howe (1981) applied Cole and Magnussen's (1966) dispositional perspective through a three-step "dispositional assessment evaluation scheme" (p. 112). In this model, a sequence of three questions were asked: (1) Was the referral meaningful? (2) Did the assessment provide an accurate understanding of the patient? (3) Were the resulting recommendations useful? If the referral question was not meaningful, which typically occurred if testing was ordered merely to rubberstamp what was already known, the author placed the referral in a "bureaucratic" category. In the study presented by Howe, 31.4% of the vocational referral cases fell into this category, causing the consultant to recommend that future referrals be examined for their appropriateness before creating the expense of unnecessary assessment.



If the second question, concerning accuracy of the assessment, was not answered positively, then the case would belong to the "nonpredictive" category, which corresponded to the traditional "miss" category for diagnostic assessment. If the assessment were accurate, however, the third question would be the essential one that determined the usefulness of the whole assessment effort. To decide whether an assessment were useful would require knowledge of the context of the assessment, the available choices for action and some means to assess whether an appropriate choice had been made as a result of the assessment findings. In his study of 51 vocational assessment cases, Howe determined that 62.9% of the "nonbureaucratic" cases were predictive and useful, 31.4% were predictive, i.e. accurate, but not useful, and 5.7% were nonpredictive, i.e. inaccurate. As in Adams (1972), the criterion for utility depended upon the author's judgment. In this study, Howe's judgment in terms of the answers to each of the three questions for each case was crucial to the assessment of utility.

The studies of Lakin and Lieberman (1965) and Hartlage, Freeman, Horine and Walton (1968) could be grouped together as laboratory studies that assessed the utility of reports in very indirect ways. This indirectness, coupled with basic inadequacies in the studies, made these approaches less attractive than others.

Lakin and Lieberman (1965) obtained Q-sort descriptions of a patient by 18 psychoanalysts who were given varying increments of information about the patient. A psychodiagnostic report Q-sort was one of the types of information presented, in addition to minimal identifying data, psychiatric intake report, social history and a therapy protocol Q-sort. The authors used a factor analytic method to ascertain different aspects of the analysts' conceptualizations of the patient. Unfortunately, their methodological description was unclear; three tables mentioned in the text were missing from the article. In their reported results, the authors stated that the information from the psychological report, as compared with information from the therapy protocol, tended to increase the homogeneity of the analysts' ratings of the patient. The test report had the effect of causing the raters to view the patient as more disturbed, with a less hopeful prognosis. The raters emphasized psychosexual issues in their conceptualizations of the patient, whereas the various pieces of information provided were primarily concerned with interpersonal issues, emotional status and self-concept. The authors concluded that changes in conceptualization appeared to depend mainly upon the analysts' theoretical orientation. Diagnostic information, although not directly related to these

psychosexual issues, likely served as a "releaser" for predisposed response sets. Each type of information, including the test report, mattered little compared to the influence of theoretical orientation. These findings suggested that if information in the psychological report had been conveyed in psychosexual terms concordant with analysts' theoretical orientation, perhaps the report would have accounted for more of the variance in ratings.

In the other laboratory study, Hartlage, Freeman, Horine and Walton (1968) extracted 4370 different content statements from 1000 psychological reports selected randomly from the files of a state hospital covering a ten-year period, 1957-67. The statements were condensed into 55 content statements that were rated independently by four psychiatrists in terms of their decisional utility for a patient's treatment plan. These ratings were found to be negatively correlated ( $r = -.50$ ,  $p < .001$ ) with the frequency of occurrence of the statements, using a Spearman correlation coefficient. The authors recommended that specific referral questions be asked in future assessments so that psychologists could supply more relevant information. They also recommended that psychologists learn what information is of high utility value to psychiatrists.

A fourth type of study, quite distinct from the others, was the assessment of the value of an extensive psychological assessment vis-a-vis long-term psychotherapy. As reported by Appelbaum (1977), a major finding in the 20-year psychotherapy research project of the Menninger Foundation was that psychological test reports were better predictors of patients' response to treatment and better assessors of global diagnostic understanding and treatment recommendations as compared to psychiatrists' judgments. This involved the study of 42 patients, mostly severely disturbed, in psychoanalytic or psychoanalytically oriented treatment. One "dismaying" aspect of this finding was that the psychiatrists' judgments were based in part on test reports, whose findings they must have disregarded or disagreed with when making their own judgments.

The fifth and final method of incremental validity assessment was represented by the oldest study of the group, reported by Dailey (1953), who endeavored to measure the utility of psychological reports in a VA psychiatric service. Utility was explicitly defined in terms of decision making usefulness.

Dailey divided his project into several steps, the first of which was a compilation of a list of 32 treatment decisions that frequently had to be made by the professional staff on the unit. This list was

obtained by simply asking the psychiatrists, psychologists and social workers to suggest items for such a list, which was then edited by the author. The second step involved the creation of an index of "stereotypy," which was to serve as a base rate control for decisions made about an "average" patient. This index was established by the ratings of ten psychologists familiar with the treatment unit. They indicated their answers ("yes," "no," or "don't know") to each of the 32 treatment decisions for a hypothetically average patient on the unit. The index thus was meant to reflect a priori judgments about a patient, in the absence of specific information on that patient.

Using two clinicians as judges, Dailey then calculated the clinical utility of psychological reports for nine patients. The judges indicated their 32 treatment decisions on the basis of reading the reports alone. Dailey then applied a weighted scoring system based upon the index of stereotypy to establish a measure of "new, useful decisions" (p. 294) made on the basis of psychological reports.

Dailey also assessed the clarity of reports by totaling the number of decisions on which the two judges agreed. He found that, on the average, the two judges could agree upon 53% of the 32 decisions after reading a report. Taking base rate information into account via the index of

stereotypy, 26% of the list of therapeutic decisions could be considered clear, new and useful decisions resulting from the use of psychological reports. Although Dailey acknowledged that there existed no criteria to judge whether such a proportion was satisfactory or not, the thrust of his concluding discussion was that this percentage indicated an insignificant effect for reports. He attributed the results to "the operation of clinical bias or rigidity" (p. 301). Dailey cautioned that his measures of utility and clarity needed improvement in reliability and that they may have been dependent upon the particular institution in question. Also, he noted that psychological reports had value not only in "gross decisions" (p. 302), such as those considered in his study, but also for "subtler problems" (p. 302) relating to psychotherapy. Another purpose of a report, "establishing in the therapist or decision maker a clear and memorable image of the patient" (p. 302), was not included in the study. Dailey also admitted that the study did not assess whether the decisions were valid or not. He concluded that the study's most important implication was that the usefulness of a report could be measured, although such methodology needed refinement.

The two other studies that directly employed reports but did not primarily assess incremental validity were

by Garfield, Heine and Leventhal (1954) and Cuadra and Albaugh (1956). Garfield et al. surveyed psychiatrists, psychologists and social workers in a VA setting. Each clinician provided letter grades (A, B, C, D and F) and spontaneous comments for two psychological reports randomly chosen from clinical records. One report was given to every clinician, each of whom also rated a second report given only to him or her. The author's conclusions about the utility of reports consisted mainly in summary comments of what aspects of reports should be improved, such as stylistic features and matters of content, e.g. the overuse of speculative inferences without supporting data.

Cuadra and Albaugh (1956) did not address the issue of utility, but instead examined the effectiveness of communication in reports. They compared report writers' intended meanings and emphasis with perceptions of clinical judges who read the reports. The authors found that the correspondence between the report writers' intentions and judges' interpretations was only 53%. The greatest source of communication breakdown was related to matters of emphasis and degree or amount of personality characteristics described.

Altogether, these ten studies represented a diversity of research methods attempting to answer diverse questions. Methodologically, the oldest study (Dailey, 1953) was

probably the most exemplary because of its control for base rates in incremental utility. For that reason it was described in more detail than the others. In general, the research findings are mixed. Some authors seemed more pessimistic in their conclusions than their data warranted. The studies may have been biased against the real value of reports. They tended to focus on the immediate measurable utility. The one study that examined utility from a long-term perspective (Appelbaum, 1977) placed reports in a very favorable light.

Tallent (1983) pointed out another important factor that these studies mostly ignored. He suggested that any evaluation of the utility of psychological reports should be done with a comparative reference to the utility of other diagnostic procedures. He noted that the utility of such procedures as the social history, mental status examination and nursing reports has not been empirically demonstrated. Borrowing Dailey's (1953) finding that psychological reports contributed to clinical decisions in 26% of referred cases, Tallent asserted:

But even if only 26 percent of psychological reports make a difference in patient management and disposition, we might regard the psychological report as a very effective instrument indeed. (1983, p. 21)



Another significant feature of the studies reviewed above was that the authors generally did not indicate an awareness of literature on judgment, decision making or the problems of combining information in clinical judgment. This may in large part be due to the dates of these publications; the research literature on judgment has been predominantly more recent than the literature on psychological reports. Also, although the use of reports for confirmation of judgment was a theme that was raised by these studies, none of them endeavored to examine the issue of confidence in judgment and the value of confirmation. Such omissions suggest opportunities for exploratory research, of which the present study is one example.

#### Applications in the Present Study

Although the breadth of the field of judgment research has been reviewed, only some of the conclusions can be applied to any one study. The present experiment obtains its impetus from various currents in the literature. Foremost is the vacuum of empirical research on the use of psychological reports. Second, the need to apply research on clinical judgment to representative situations is addressed. As such, this study belongs

more to the applied, as opposed to the basic, pole of the research spectrum. However, this need not detract from potential scientific value. As Hammond et al. (1964) noted:

Clinical psychology provides an excellent research site for the study of man as an inferring organism. (p. 438)

The study does not employ mathematical models in the manner of contemporary research on judgment (Slovic and Lichtenstein, 1973; Hogarth, 1980). This is due to the limited number of subjects compared to the multiplicity of response variables. Also, the ecological validity of psychological reports is unknown. Assessment of accuracy or optimality of judgments is not the concern here. This study is concerned with cue utilization, that is, how subjects will respond to a given set of cues. Ecological validity of the cues or overall functional validity of responses is not at issue because there are no established criteria with which such validities could be assessed.

Therefore, the present study can be classified as a "single-system" case (Rapoport & Summers, 1973; Hammond et al., 1975). It also is best described, "discovery," as opposed to a systematically controlled experiment. The general issues are how clinicians use a psychological

report and how different information influences clinical judgment and confidence in that judgment.

Questions of professional differences are of secondary interest, prompted by the inclusion of two professional groups among the subjects. The study does not look at institutional relationship factors in the use of psychological tests, i.e. it is not, for example, a study of psychiatrists' attitudes toward psychologists. However, the author does hypothesize that psychiatrists will tend to be more confident than psychologists because the latter group has had more of an emphasis in acquiring skeptical attitudes in clinical thinking. This is in line with Brody's observation:

The graduate student in psychology lives in an atmosphere of research and skepticism, in which theories are constructed and abandoned, in contrast to that of the medical student, intern, and psychiatric resident who learn to apply already established knowledge and are reluctant to attack established theories, especially those espoused by respected figures in the field. (1956, p. 107)

A multi-method approach, involving structured and unstructured response tasks, is followed. As is true for the multiple sources of information on patients, the variety in the response format is intended to reflect the variety common in real clinical situations. This includes both descriptive and predictive judgment. An

audiotape is employed to approximate an actual clinical interview, which is the central activity in the diagnostic judgment process.

In general, this two-part literature review has revealed the need to apply laboratory-based methods of studying judgment and decision making to the practical issue of how psychological reports are used. The present study necessarily looks at only a small part of the many issues touched upon here. The complexities of the topic have been delineated to justify the breadth of experimental measurement and to inform the interpretation of subsequent findings.

## CHAPTER III METHOD

### Summary

Two clinical situations common to psychiatric hospitals were simulated. Subjects were clinicians from the professions of psychiatry and clinical psychology who were exposed to different sets of information about two hypothetical patients. A psychological test report was included as part of the information. Using a pre- and post-test design assessing clinical judgment and confidence in that judgment, the effects of the reports were assessed.

Each subject was asked to assume the role of treating clinician for the hypothetical patients. Information about each patient was given in four parts, two for each experimental session. In the first session, the subject read a brief admitting note describing the first patient and circumstance of hospitalization. Then the subject listened to a brief audiotape of the patient during an interview portrayed by an actor and actress. Following this, dependent measures were administered to obtain the

subject's assessment of the patient in terms of symptomatology and diagnostic impression. Measurement of the subject's certainty in these judgments was also made. The subject then had an opportunity to write himself or herself a reminder for recalling the patient at the second session approximately two weeks later. The same procedure was repeated using the other patient. For the second session, the subject was given a nursing report summary of each patient's stay in the hospital during the previous two weeks when the therapist (subject) had been on vacation. Then psychological reports of the patients were given to the subject, followed by a recent administration of the dependent measures. The psychological report for one of the patients confirmed the subject's initial judgment of presence or absence of thought disorder, whereas the report for the other patient disconfirmed this judgment. Data were analyzed by comparing responses from the first session with those of the second and by analyzing scores according to subject variables such as profession and experience.

### Hypotheses

#### Main Hypotheses

1. Clinicians would adhere to their original judgments even when these judgments were contradicted by a psychological test report.

2. Clinician certainty in judgment would increase with additional information. This would be independent of the content of the judgment and would occur even when the information upon which the judgment was based was inconsistent.

### Secondary Hypotheses

3. Psychiatrists would be more certain more quickly than psychologists. This would be most evident in the assessment of diagnoses.

4. Clinician certainty in judgment would vary according to experience.

### Subjects

The sample included psychiatrists, psychologists and those in advanced training for those roles. All subjects were clinicians at either the J. Hillis Miller Health Center of the University of Florida or the Gainesville Veterans Administration Medical Center. An attempt was made to get a representative cross-section of clinical experience in the sample. A total of 32 subjects participated, with 16 each from psychiatry and clinical psychology. Psychiatry was represented by 8 members of the Department of Psychiatry faculty and 8 psychiatric residents in that department. Eight faculty or affiliate faculty members

of the Department of Clinical Psychology participated, along with 8 non-faculty affiliates of the same department. This last group consisted mainly of advanced graduate students who had completed their internships but had not finished their dissertations.

The author personally recruited each subject, who was asked to donate his or her time for a "study of clinical judgment." Written informed consent was given by each participant. Subjects were assured that their responses were confidential and that the study was not a test of their aptitude.

#### Materials

The design of this study centered upon the creation of the hypothetical psychiatric patients. The patients were meant to represent, in a plausible way, hospitalized young adults who presented diagnostic problems. The different sources of information that presented an account of personal history, history of present illness, mental status, current functioning and other data were intended to represent realistically those found in a contemporary clinical setting. The goal was to have cases that would reasonably, if not typically, prompt a referral for psychological evaluation. The hypothetical cases included one female (Case S) and one male (Case F).



The question of presence or absence of a thought disorder was the key content of interest. The information presented in the initial session was designed to be highly ambiguous as to whether each patient had a thought disorder. Ideally, the subjects would be evenly split in their judgments about thought disorder at the first session. For purposes of simplicity, a social work evaluation was omitted from the clinical information.

#### Admitting Note

Each patient was described in the standard format of a psychiatric admission note, written by another doctor who was on duty at the times of the patients' arrivals at the hospital. The notes were descriptive and did not include the usual diagnostic impression or treatment recommendations. They are reproduced in their entirety in Appendix A.

#### Interview

Subjects listened to brief audiotaped interviews of each patient, portrayed by an actor and actress, respectively. The author assumed the role of interviewer. The interviews were designed to raise questions of diagnosis rather than to answer them. The patients expressed some unusual beliefs, displayed some mild loosening of associations, but no clear formal thought disorder. As in the admitting

note, multiple diagnoses were suggested. The interviewer's participation was minimal, but realistic. Complete transcripts are provided in Appendix B. The lengths of the interviews were 4 min 56 s for Case S and 6 min 55 s for Case F.

#### Nursing Report Summary

Each summary (Appendix C) given at the beginning of the second session was designed to help the subjects further visualize the patients without clarifying diagnostic issues. They also served the narrative purpose of tying together the two sessions of the study in a manner that would be plausible for the portrayed situations.

#### Psychological Report

Each subject received two types of psychological reports. One report confirmed the initial judgments of the presence or absence of a thought disorder for one of the patients; the report for the other patient disconfirmed this judgment. Thus, the type of report included in the second session depended upon the subject's responses in the first session.

The reports, presented in Appendix D, were designed to be representative of acceptable reports found in clinical practice. The issue of thought disorder was embedded in the body and conclusion of each report. The intention was

to make the psychologist's judgment about thought disorder clear but not obviously distracting compared to the rest of the reports' contents. The author constructed these reports with the aid of previous reports taken from actual clinical cases.

### Measures

#### Brief Psychiatric Rating Scale (BPRS)

The BPRS (see Appendix E) is a widely used rating instrument for psychiatric symptoms. Developed by Overall and Gorham (1962) on the basis of previous cluster and factor analytic studies, it was designed to be:

a highly efficient, rapid evaluation procedure for use in assessing treatment change in psychiatric patients while at the same time yielding a rather comprehensive description of major symptom characteristics. (p. 799)

Hedlund and Vieweg (1980) concluded in their comprehensive review of 177 studies using the BPRS that the scale was "a relatively simple, easy-to-learn, comprehensive rating scale that is quite compatible with traditional clinical concepts" (p. 56). These authors also stated that the scale had been shown to be valid over a wider range of psychopathology than just the functionally psychotic patients for whom it was originally intended. Median interrater reliability coefficients,

using the Pearson Product Moment Correlation, ranged from .63 to .88 for each item and an estimated .85 for total pathology score. Item #4 ("Conceptual Disorganization"), which is of interest in this study as a rating of thought disorder, had a reliability range of Pearson  $r$  from .64 to .97, with a median reliability of .80 across five studies. Four higher-order factors have been derived from the scale. The most reliable factor, "Thinking Disturbance," also to be used in the present study, had a reported range of .91 to .96, with a median reliability of .94 for the same five studies. This factor includes items #4, #12, ("Hallucinatory Behavior") and #15 (Unusual Thought Content"). Hedlund and Vieweg reviewed a number of validity studies involving mainly concurrent and construct validity. Many of these involved the use of the BPRS to evaluate behavioral change following pharmacological treatment. The measure has also been used extensively in classification research as an aid to the development of diagnostic subtypes (Overall & Woodward, 1975).

#### Symptom Checklist

This is a 15-item scale, listing psychiatric symptoms in a presence/absence format, with a rating of certainty for each item (see Appendix F). The author constructed this list mainly as a format within which to ask subjects whether or not a thought disorder was present in each case.

Again, the thought disorder measure was placed unobtrusively within a broader range of symptoms. The certainty ratings served as an overall measure of subjects' confidence; they were used to test the hypothesis that confidence would increase with increased information. The term "certainty" has been used because "confidence" may invoke notions of self-esteem not implied by "certainty." Thus, subjects may have been more willing to express a lack of certainty rather than a lack of confidence, even though both words express a subjective probability estimate of a correct judgment.

#### Diagnostic Impression

This unstructured measure (Appendix G) was included as a means of capturing the subject's diagnostic thinking prior to the structured prompting of subsequent measures. It did not have a direct bearing on the main hypotheses. It may have also facilitated the subject's focus of attention onto diagnosis and away from the specific issue of thought disorder.

#### Diagnostic Selection: Axis I and Axis II

These two measures (Appendix H) provided the structure to assess whether the different treatment conditions made a difference in diagnostic judgment, which is the crucial judgment in clinical practice. In each case for both

Axis I and Axis II, the ten diagnoses listed were plausible possibilities, based upon the presented information and DSM-III criteria. The rank-order method was included to see how many and which diagnoses the subject was considering. As mentioned before, these measures were meant to highlight the formal diagnoses and serve to attract less attention to the thought disorder issue per se.

#### Additional Information

This measure (Appendix I) included a list of 22 pieces of information that subjects could request to help make their judgments "sufficiently certain." It was intended as an indirect measure of subject confidence in judgment. The author constructed the list, based partly on Gauron and Dickinson's (1966) study of psychiatric judgment and use of clinical information.

#### Recall Note

This note (Appendix J) was meant primarily to serve its stated function: as a memory aid to link the two sessions for the subjects. It was designed to help subjects remain cognizant of their earlier judgments, as they would have been in a normal clinical situation. It also provided a glimpse of what information was highlighted by subjects. This may be of only anecdotal interest for the present study, but could contribute to ideas for further research.

### Final Questions

This last measure was intended as a means of recording demographic information and attitudes about psychological reports. It served to divide subjects according to professional and experimental status and to compare how their attitudes toward psychological reports related to their use of information in the experimental reports. The measure was simple and brief (see Appendix K). It was given at the end of the study so that subjects would not be previously aware both of the dimensions by which they would be grouped and also that the use of psychological reports was the central concern of the study.

### Procedures

#### Recruitment

Subjects were recruited via phone calls by the author. The study was briefly described as exploratory research in clinical thinking. Subjects were asked to donate one hour of their time for each of the two sessions of the study. Because the subject population was one for whom time was at a premium, subjects were given materials to take home and complete at their convenience. The presentation of materials and instruments of measurement were designed to take no more than 40 or 50 minutes for each session.

Session #1

The test packet was arranged so that the subject would use the materials in the designated order. The first page contained the following introduction:

Welcome to the first of two sessions in this study of clinical thinking. The following materials are being presented as part of an exploratory research program. There are no established right or wrong answers to the questions that you will be asked. The purpose of this study is to discover how clinicians will respond to these questions.

The materials have been carefully arranged in a specific order. Please read and complete one page at a time without looking ahead. The instructions will indicate when you are to listen to the enclosed cassette tape. Estimated time for completion of this packet is 40 to 50 minutes. When you are finished, please place all materials in the envelope, seal it and return to Susan Linn (Dr. Blashfield's secretary), Box J-256, JHMHC (392-4196).

This was followed by the directions on the next page:

Please imagine that you work as a staff psychiatrist (psychologist) on an inpatient unit of a small private psychiatric hospital. The unit includes adolescent and adult patients representing a wide range of psychiatric disorders. Average length of stay is approximately three to four weeks for adults and two to three months for adolescents. You are the clinician who will have primary responsibility for treatment of the two patients whose cases will be described in the following materials. Description of the first patient will begin with the admitting note from the patient's current medical record.



The admitting note from the patient's chart, as described in the Materials section, was then presented. Another brief set of instructions introduced the use of the audiotape and dependent measures:

As the primary therapist, you would normally be interviewing this patient that morning. However, due to the limitations of the study, you will instead listen to a five-minute audiotape interview of the patient. Please play the enclosed cassette tape, beginning at Side 1.

When the interview has ended, stop the tape player and leave the tape at that position (the interview of the second patient will then be ready to begin when required later in this session). Please complete the following measures in the order in which they appear. You may refer back to the presented materials and you may also use any additional references, e.g. DSM-III. However, the study is designed so that you will have little need for any additional materials.

Upon finishing the first case, the subject next turned to the directions for the second case:

Description of the second patient will be given in a similar manner. An admitting note will be followed by a brief audiotape interview. The interview will begin on the tape approximately 30 seconds following the end of the interview of the previous patient. As before, please complete the subsequent measures in the order in which they appear. You may refer to any materials, e.g. DSM-III, that you wish.

Session #2

Approximately two weeks after the first test packet had been returned, each subject received a second packet, which began with the following instructions:

In this second and final session you will be given additional information about the two patients described in the previous session. Please remember that there are no established right or wrong answers to the questions that you will be asked. The purpose of this study is to discover how clinicians will respond to these questions.

As before, the materials have been arranged in a specific order. Please read and complete one page at a time without looking ahead. Each case begins with the reminder notes that you completed at the previous session. Estimated time for completion of this packet is 45 minutes. When you are finished, please place all materials in the envelope, seal it and return to Susan Linn (Dr. Blashfield's secretary), Box J-256, JHMHC (392-4196).

These were followed by the Recall Note, which had been completed by the subject for each patient during the first session. Then the following directions were given:

Since the day of the patients' admissions you have been away on a two-week vacation. It is the morning of your first day back at the hospital. You learn that both patients are still in the hospital and that you are their therapist. Prior to attending the morning staff meeting, you read the nursing report summary and psychological report in each of their charts. You begin with one of the two patients.

The Nursing Report Summary and Psychological Report were presented, respectively, followed by the instructions:

Please complete the following measures in the order in which they appear. As before, you may refer back to the presented materials and you may also use any additional materials, e.g. DSM-III. Please base your responses on all the information presented about the patient, both in this session and in the previous one.

At the end of the first case of the second session, the following directions introduced the second case:

Description of the other patient will be given in a similar manner. The nursing report summary will be followed by the psychological report. As before, please complete the measures in the order in which they appear and base your responses upon all the information presented in both sessions.

Then the Recall Note was followed by the Nursing Report Summary and Psychological Report. The dependent measures concluded with the Final Questions, as described in the Materials Section.

An outline of the procedure for both sessions is given in Table 3. Table 4 lists the experimental design, which balances the various combinations of presentation order for the two cases along with the assignment of confirming/disconfirming psychological reports to the cases.

Table 3  
Procedure Outline

I. Session #1

Case S

Admission Note

Taped Interview

-----

BPRS

Symptom Checklist

Diagnostic Impression

Diagnostic Selection

Additional Information

Recall Note

Case F

Admission Note

Taped Interview

-----

BPRS

Symptom Checklist

Diagnostic Impression

Diagnostic Selection

Additional Information

Recall Note

II. Session #2

Case S

Recall Note

Nursing Report

Psychological Report

-----

BPRS

Symptom Checklist

Diagnostic Impression

Diagnostic Selection

Additional Information

Final Questions

Case F

Recall Note

Nursing Report

Psychological Report

-----

BPRS

Symptom Checklist

Diagnostic Impression

Diagnostic Selection

Additional Information

Final Questions

Table 4  
Experimental Design

<u>Subject</u>	<u>Order #1</u>	<u>Order #2</u>	<u>Psychological Report</u>
S1	S-F	S-F	C-S, D-F
S2	F-S	F-S	D-S, C-F
S3	S-F	S-F	D-S, C-F
S4	F-S	F-S	C-S, D-F
S5	S-F	F-S	C-S, D-F
S6	F-S	S-F	D-S, C-F
S7	S-F	F-S	D-S, C-F
S8	F-S	S-F	C-S, D-F
S9	S-F	S-F	C-S, D-F
S10	F-S	F-S	D-S, C-F
S11	S-F	S-F	D-S, C-F
S12	F-S	F-S	C-S, D-F
S13	S-F	F-S	D-S, D-F
S14	F-S	S-F	D-S, C-F
S15	S-F	F-S	D-S, C-F
S16	F-S	S-F	C-S, D-F
S17	S-F	S-F	C-S, D-F
S18	F-S	F-S	D-S, C-F
S19	S-F	S-F	D-S, C-F
S20	F-S	F-S	C-S, D-F
S21	S-F	F-S	C-S, D-F
S22	F-S	S-F	D-S, C-F
S23	S-F	F-S	D-S, C-F
S24	F-S	S-F	C-S, D-F
S25	S-F	S-F	C-S, D-F
S26	F-S	F-S	D-S, C-F
S27	S-F	S-F	D-S, C-F
S28	F-S	F-S	C-S, D-F
S29	S-F	F-S	C-S, D-F
S30	F-S	S-F	D-S, C-F
S31	S-F	F-S	D-S, C-F
S32	F-S	S-F	C-S, D-F

Order #1 = Order of presentation of cases in Session #1

Order #2 = Order of presentation of cases in Session #2

C-S, D-F = Confirming report in Case S; Disconfirming in Case F

D-S, C-F = Disconfirming report in Case S; Confirming in Case F

## CHAPTER IV RESULTS

### Subject Characteristics

Subject selection included eight faculty members each from the Departments of Clinical Psychology and Psychiatry; eight psychiatric residents; and eight non-faculty affiliates of Clinical Psychology. This last group included post-doctoral fellows, but was comprised mainly of advanced graduate students who had completed all requirements for the Ph.D. except the dissertation. All subjects except one psychologist, one psychiatrist and one psychiatric resident completed all materials they had begun. These three subjects were each replaced by another clinician in their respective categories. No attempt was made to balance the sample by sex. The larger number of males (78.1%) was higher than the sex distribution of the two departments. Because obtaining free clinician time was difficult and because there was no reason to suspect that sex would be a factor in this study, no constraints of representativeness by sex were placed upon the sample.

Table 5 lists the subject groups by sex and mean years of experience since receiving either an M.D. or Ph.D. degree. Four of the non-faculty psychology subjects had not yet received the Ph.D.; they were listed as having zero years experience. As Table 5 demonstrates, the residents and non-faculty psychologists formed a more homogenous group in terms of experience compared to the two faculty groups, each of which had considerable variation in experience. Analyses of results relating to experience were thus done according to the two groups: faculty versus non-faculty/residents.

Table 5.  
Subject Groups by Sex and Experience

	<u>M</u>	<u>F</u>	<u>Total</u>	<u>Post-Degree Years Mean</u>	<u>SD</u>
<u>Clinical Psychology</u>					
Non-faculty	6	2	8	2.8	1.4
Faculty	6	2	8	11.8	8.4
<u>Psychiatry</u>					
Residents	6	2	8	2.8	1.4
Faculty	7	1	8	14.0	9.8
Total	25	7	32	7.5	8.4

Table 6 demonstrates the diversity of clinical orientation in the sample. Subjects were asked to select one of

eleven terms, including "other," that best described their clinical orientation. Nine of the eleven descriptors were selected. Only "Gestalt" and "Systems Theory" were omitted. Over one-third of the subjects endorsed "Eclectic," which also accounted for one-half of the lesser experienced groups.

A more striking display of diversity was found in the list of authors cited by subjects as "most representative" of their clinical orientation. Twenty-eight different authors were listed. Six subjects did not provide names; several others named more than one. Only five authors were named more than once: Beck, Ellis, Horney, Langs and Lewinsohn, each of which was mentioned by two subjects. The diversity of names reflected the diversity of reported clinical orientations.



Table 6.  
Subjects' Clinical Orientation

	<u>Clinical Psychology</u>		<u>Psychiatry</u>		<u>Total</u>
	<u>Non-Faculty</u>	<u>Faculty</u>	<u>Residents</u>	<u>Faculty</u>	
Behavioral/ Learning Theory	1	2			3
Cognitive	1	1	1		3
Eclectic	3	2	5	3	13
Gestalt					
Humanistic/ Existential		1			1
Interpersonal		1	1		2
Pharmacological/ Biological			1	1	2
Psychoanalytic	1			2	3
Psychodynamic	2	1		1	4
Systems Theory					
Other				1	1
Total	8	8	8	8	32

Subjects were less diverse in their attitudes towards psychological reports. They were asked to rate, on a scale of one to six, the usefulness of psychological reports generally and the reports used in this study in particular. Table 7 provides the mean ratings according to four different subject groups.

Table 7.  
Ratings of Usefulness of Psychological Reports

	<u>Reports Generally</u>			<u>Reports Used in Study</u>	
	<u>N</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
<u>Clinical Psychology</u>					
Non-Faculty	8	5.12	0.99	3.88	1.13
Faculty	8	3.88	1.25	3.25	0.89
<u>Psychiatry</u>					
Residents	8	4.62	1.06	4.25	1.28
Faculty	8	4.62	0.74	3.50	1.20

Ratings of reports in general were favorable, especially by the non-faculty psychologists. The high rating given by the non-faculty psychologists may have represented the fact that members of this group were more likely to have been writing such reports as part of their regular duties. Because producing such reports is a difficult task, a high rating affirms the value of such effort and avoids the cognitive dissonance that a lower rating would represent. It is interesting to note that the most experienced clinical psychologists gave the lowest ratings for psychological reports.

Table 7 also indicates that the reports used in this study were viewed by most subjects as being less useful than reports generally. Whether these lower ratings related to the perceived quality of the reports or

to the confirming/disconfirming variability of the reports is uncertain.

### The Hypothetical Cases

The two cases had been created with the intent of making diagnosis difficult. In fact, information provided at the first session could rationally support the assignment of any of the diagnoses in the multiple choice Diagnostic Selection measures for both Axis I and II. The ideal situation would have been an equal distribution of thought disorder judgments for both cases. However, the results revealed that this attempt at diagnostic ambiguity succeeded for Case S, but was not successful for Case F. Detailed analyses of the two cases will be presented as the specific hypotheses are discussed; some general differences will be noted at the outset.

The best way to compare the cases is by the scores at the first session because at that point all subjects had received the same information. Attempts to compare final responses would be ambiguous because of the different combinations of initial judgments and experimental conditions.

In terms of the initial thought disorder judgments, Case S was almost equally divided (thought disorder: 15; no thought disorder: 17). Case F, on the other hand, had 27 endorsements of thought disorder versus only 5 who said no thought disorder.

Initial Diagnostic Impressions were varied, especially for Case S, which produced 15 different DSM-III diagnoses for Axis I and 5 different for Axis II. The most frequently cited categories were Cyclothymic Disorder, given by 7 subjects for Axis I, and Borderline Personality Disorder, cited by 6 subjects for Axis II. Some of the other categories were highly related, differing mainly in specific subcategories, such as Bipolar Disorder, Depressed versus Bipolar Disorder, Mixed. The most frequently listed diagnoses for Case S are listed in Table 8.

Table 8.  
Diagnostic Impression: 1st Session Case S

<u>DSM-III Diagnosis</u>	<u>DSM-III #</u> <sup>a</sup>	<u>Frequency</u> <sup>b</sup>
<u>Axis I</u>		
Cyclothymic Disorder	301.13	7
Bipolar Disorder, Manic	296.4	5
Major Depression, Single Episode	296.2	3
Bipolar Disorder, Depressed	296.5	3
Bipolar Disorder, Mixed	296.6	3
No Axis I disorder		3
<u>Axis II</u>		
No Axis II disorder		20
Borderline Personality Disorder	301.83	6
Histrionic Personality Disorder	301.5	3

<sup>a</sup>Similar Axis I diagnostic subcategories have been combined by the first decimal digit.

<sup>b</sup>Diagnoses are listed only if mentioned by at least three clinicians.

For Case F, ten different Axis I diagnoses were given at the first session; only two different Axis II categories were named (Table 9). Schizophreniform Disorder and Schizoid Personality Disorder were the most frequently given diagnoses for Axis I and Axis II, respectively.

Table 9.  
Diagnostic Impression: 1st Session Case F

<u>DSM-III Diagnosis</u>	<u>DSM-III #<sup>a</sup></u>	<u>Frequency<sup>b</sup></u>
<u>Axis I</u>		
Schizophreniform Disorder	295.4	12
Major Depression, Single Episode	296.2	10
Schizoaffective Disorder	295.7	4
Schizophrenia, Undifferentiated	295.9	3
<u>Axis II</u>		
No Axis II disorder		23
Schizoid Personality Disorder	301.20	8

<sup>a</sup>Similar Axis I diagnostic subcategories have been combined by the first decimal digit.

<sup>b</sup>Diagnoses are listed only if mentioned by at least three clinicians.

Initial diagnoses were also assessed in a more structured multiple choice format (Diagnostic Selection). For Case S, 9 of the 10 Axis I diagnoses were selected by at least 2 clinicians each; 7 of the 10 Axis II diagnoses were chosen by at least one clinician (Table 10).

Table 10.  
Diagnostic Selection: 1st Session Case S

<u>Diagnosis</u>	<u>Frequency</u> <sup>a</sup>
<u>Axis I</u>	
Bipolar Disorder, Depressed	9
Cyclothymic Disorder	7
Factitious Disorder with Psychological Symptoms	3
No Axis I disorder	3
<u>Axis II</u>	
Borderline Personality Disorder	12
Histrionic Personality Disorder	6
No Axis II disorder	5
Mixed Personality Disorder	3
Schizotypal Personality Disorder	3

<sup>a</sup>Diagnoses are listed if selected by at least three clinicians.

For Case F, 7 of the 10 Axis I diagnoses were chosen in the first session; 5 of the 10 Axis II diagnoses were also selected. Table 11 lists the most often cited categories for Case F.

Table 11.  
Diagnostic Selection: 1st Session Case F

<u>Diagnosis</u>	<u>Frequency</u> <sup>a</sup>
<u>Axis I</u>	
Schizophreniform Disorder	12
Major Depressive Episode with Psychotic Features	7
Schizoaffective Disorder	5
Schizophrenic Disorder, Undifferentiated	4
<u>Axis II</u>	
Schizoid Personality Disorder	21
Schizotypal Personality Disorder	6

<sup>a</sup>Diagnoses are listed if selected by at least three clinicians.

In general, at the first session, subjects predominantly judged Case F to be psychotic, whereas Case S was more variable, tending to be considered as having an affective disorder or disordered personality, but not as likely to be psychotic. The greater clarity provided by Case F was also expressed in higher certainty scores, especially in the first session (Table 12).



Table 12.

Mean Certainty Scores: Case S versus Case F

	<u>Session 1</u>				<u>Session 2</u>			
	<u>Case S</u>		<u>Case F</u>		<u>Case S</u>		<u>Case F</u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
TDDC <sup>a</sup>	4.02	1.41	4.50	1.24	3.88	1.54	3.81	1.42
SCTOT <sup>b</sup>	62.08	9.86	68.11	9.03	64.30	9.70	65.81	8.86
DIDC <sup>a</sup>	3.72	1.08	4.10	1.04	3.84	1.07	4.13	1.06
DSIDC <sup>a</sup>	3.41	1.19	4.06	0.98	3.88	1.01	4.19	1.06
DSIIDC <sup>a</sup>	3.56	1.24	3.97	1.28	3.59	1.10	3.94	1.31

TDDC: Degree of certainty for Thought Disorder item.  
 SCTOT: Total Sum of certainty ratings on Symptom Checklist.  
 DIDC: Diagnostic Impression degree of certainty.  
 DSIDC: Diagnostic Selection degree of certainty (Axis I).  
 DSIIDC: Diagnostic Selection degree of certainty (Axis II).

<sup>a</sup>Scores ranged from 1 (Uncertain) to 6 (Certain).

<sup>b</sup>Possible range of scores: 15 (Uncertain) to 90 (Certain).

Hypothesis #1

The first hypothesis stated that subjects would maintain their initial judgments about presence of thought disorder. This would occur even when given a disconfirming report.

Results did not support this hypothesis. Contingency tables were constructed, pairing judgments of thought disorder, as measured by the Symptom Checklist, for the two

sessions. For judgments made in the confirming condition, in which the psychological reports confirmed the initial judgment, results were generally as expected (Tables 13 and 14).

Table 13.

Thought Disorder Judgments: Confirming Condition Case S

		<u>Session 2</u>		
		<u>No TD</u>	<u>TD</u>	
Session 1	No TD	9	0	9
	TD	3	4	7
		12	4	16

$\chi^2 (1, n = 16) = 6.857, p < .05$ , Fisher's Exact Test.

Table 14.

Thought Disorder Judgments: Confirming Condition Case F

		<u>Session 2</u>		
		<u>No TD</u>	<u>TD</u>	
Session 1	No TD	2	0	2
	TD	0	14	14
		2	14	16

$\chi^2 (1, n = 16) = 16.000, p < .01$ , Fisher's Exact Test.

As Table 14 indicates, for Case F, all subjects maintained their judgments on thought disorder when the psychological report confirmed them. This pattern did not hold as strongly for Case S, in which three subjects changed their judgments from positive to negative for thought disorder.

It was unclear as to why such a change was made. Consequently, the original protocols of these three subjects were examined. For two of these three, the initial judgment had been thought disorder at a low degree of certainty (2 on the 6-point scale). These two subjects had also given initial Diagnostic Impressions and Axis I Selections of Bipolar Disorder and Cyclothymic Disorder, neither of which includes thought disorder in its formal criteria. Conceptual Disorganization ratings on the BPRS were also low (3 and 2, respectively). Thus these two subjects displayed uncertainty and inconsistencies in their initial judgments related to thought disorder. The third subject was relatively confident (rating = 4) of his thought disorder endorsement, which was reflected in his Diagnostic Impression and Selection, Brief Reactive Psychosis. However, his initial Conceptual Disorganization rating was also low (2), as was the three-item Thought Disturbance Syndrome Factor (6). He also requested a high number of items on Additional Information (12).

For the second session, the three subjects paradoxically changed their opinions about thought disorder, which contradicted their first judgments and the psychological report. The first two subjects increased their degree of certainty for this judgment (4 and 3). However, their ratings for Conceptual Disorganization also increased. One subject decided upon a nonpsychotic Axis I diagnosis (Dysthymic Disorder) while the other one indicated no Axis I disorder as his final judgment. The latter gave Borderline Personality Disorder as the primary diagnosis, which was ranked second for the former's Axis II diagnoses. The third subject indicated much uncertainty (rating = 1) for his final judgment of no thought disorder. This was paired with the doubling of the scores for Conceptual Disorganization and Thought Disturbance Syndrome Factor. His final Diagnostic Impression included Brief Reactive Psychosis, Mixed Personality Disorder and Adjustment Disorder with Mixed Disturbance of Emotions and Conduct.

In summary, these three anomalous cases included internal inconsistencies and uncertainties reflecting the diagnostic difficulty of the case. Although these subjects did not precisely follow the confirming psychological report, due to their omission of thought disorder as a present symptom at the second session, their other judgments suggest that they were close to making such

a judgment. Thus their judgments do not stand as glaring exceptions to the general trend of following the psychological report. In general, for all subjects, the judgments in the confirming condition made rational sense. They suggest that the subjects were attending to the information.

The results for thought disorder judgments in the disconfirming condition are less clear. Tables 15 and 16 indicate this by each case.

Table 15.

Thought Disorder Judgments: Disconfirming Condition Case S

		<u>Session 2</u>		
		<u>No TD</u>	<u>TD</u>	
Session 1	No TD	4	4	8
	TD	7	1	8
		11	5	16

$$\underline{X}^2 (1, n = 16) = 2.618, n.s.$$

Table 16.

Thought Disorder Judgments: Disconfirming Condition Case F

		<u>Session 2</u>		
		<u>No TD</u>	<u>TD</u>	
Session 1	No TD	1	2	3
	TD	6	7	13
		7	9	16

$$X^2 (1, n = 16) = 0.163, n.s.$$

The experimental hypothesis had predicted that the contingency pattern for the disconfirming condition would be similar to the results for the confirming condition. In Case S, almost the opposite result was obtained for those subjects who had initially indicated thought disorder. However, a more ambiguous result was found for subjects who had initially judged no thought disorder. Only half of these clinicians were willing to change their opinions and follow the psychological report. This might suggest that these subjects were more confident in their first judgments of no thought disorder than were the subjects who initially judged Case S to have a thought disorder.

Certainty levels were thus examined for thought disorder ratings of subjects who were in the disconfirming

condition. Table 17 presents these data; those who had initially judged no thought disorder for Case F are excluded because of the small number of observations.

Table 17.  
Mean Thought Disorder Certainty Ratings: Disconfirming Condition

<u>Case</u>	<u>1st Judgment</u>	<u>N</u>	<u>Session 1 Certainty</u>			<u>Session 2 Certainty</u>		
			<u>Mean</u>	<u>SD</u>	<u>Range</u>	<u>Mean</u>	<u>SD</u>	<u>Range</u>
S	No TD	8	4.38	1.19	3-6	4.00	1.31	2-6
S	TD	8	4.12	1.13	2-5	3.75	1.58	2-6
F	TD	13	4.77	0.93	3-6	3.31	1.32	1-5

The certainty levels for Case S were thus highly similar for the differing initial judgments. Thus, the subjects who were less willing to change their opinions did not express more confidence in their judgments.

An alternative explanation for the results of Table 15 could be that these subjects might have represented clinicians' preference for conservative judgments. For Case S, they were more willing to switch to a judgment of no thought disorder than to a positive judgment.

Case F was quite different in that subjects were not as willing to change their initial impressions that thought disorder was indeed present. This may have reflected the different characteristics of the two cases.

Whereas Case S was evenly divided at the first session on subjects' judgments of thought disorder (46.9% indicated yes), Case F was overwhelmingly endorsed (84.3%) as having a thought disorder. As Table 17 indicates, subjects who assigned a thought disorder to Case F in the first session did so with a high level of confidence.

A comparison was made between the six subjects who changed their judgments and the seven who did not after each group had received a report disconfirming their initial positive judgments of thought disorder (Table 16). There appeared to be little difference between these subjects according to profession. However, all three of the clinical psychology faculty in this group maintained their initial judgments. In contrast, three of the four non-faculty psychology subjects changed their judgments in response to the reports. This difference was consonant with the above reported differences between these two groups in their valuation of psychological reports (Table 7). The six subjects who changed their judgments are compared with the seven who did not in Table 18.



Table 18.  
Comparison of Subjects Who Assigned Thought Disorder  
to Case F in 1st Session: Disconfirming Condition

	<u>Thought Disorder Judgment: 2nd Session</u>			
	<u>No TD<sup>a</sup></u>		<u>TD<sup>b</sup></u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
TDDC Session 1	4.50	1.05	5.00	0.82
TDDC Session 2	2.50	1.38	4.00	0.82
Usefulness of Reports Generally	5.17	0.75	4.86	0.69
Usefulness of Reports This Study	4.33	0.82	3.57	0.98

TDDC: Degree of certainty for Thought Disorder item,

<sup>a</sup>n = 6.

<sup>b</sup>n = 7.

Table 18 consists of subjects who initially indicated thought disorder for Case F and then received a psychological report stating that no thought disorder was present. The subjects who did not follow the reports tended to have more certainty in their judgments and saw reports as less useful than did the subjects who followed the reports. Confidence decreased from Session 1 to Session 2, especially for those who followed the report. These data suggest that the disconfirming nature of the report had an impact on both groups. Confidence in the initial impression of Case F as a thought disordered patient was quite high, such that over half the subjects did not change their judgments. Those who did change did so with hesitation, as expressed by their drop in certainty.

Dependent measures of thought disorder derived from the BPRS were also used to test the first hypothesis. These measures included the rating of Conceptual Disorganization (DC) and the Thought Disturbance Syndrome Factor (TDSF). Table 19 summarizes these results according to type of psychological report received.

Table 19.  
Mean Ratings at Session 2: Confirm versus Disconfirm Condition

<u>1st Judgment:</u>	<u>CD</u>		<u>TDSF</u>	
	<u>No TD</u>	<u>TD</u>	<u>No TD</u>	<u>TD</u>
<u>Case S</u>				
Confirm	2.00	4.71 <sup>a</sup>	4.00	10.57 <sup>b</sup>
Disconfirm	3.75	1.75 <sup>a</sup>	7.29	4.38 <sup>b</sup>
<u>Case F</u>				
Confirm		4.86 <sup>c</sup>		12.43 <sup>d</sup>
Disconfirm		3.54 <sup>c</sup>		8.38 <sup>d</sup>

CD: Conceptual Disorganization item of BPRS.

TDSF: Thought Disturbance Syndrome Factor of BPRS.

TD: Thought Disorder item on Symptom Checklist.

Note: Data for Case F, No TD, was not included because of small n.

Note: CD has a range of 0 (Not Present) to 6 (Extremely Severe).

Note: TDSF has a range of 0 (Not Present) to 18 (Extremely Severe).

Note: Means with the same subscript differ significantly at  $p < .01$ , Mann-Whitney U Test, Confirm versus Disconfirm.

For both cases, CD and TDSF ratings differed significantly between subjects who received confirming and disconfirming reports only if the initial judgments were

positive for thought disorder. Under those conditions the ratings corresponded to the psychological reports. For the subjects who had initially indicated no thought disorder the differences at the second session between the two conditions were not statistically significant, although the trends were in the same direction. There were too few subjects to analyze the data of those who first judged no thought disorder for Case F.

To test whether certain aspects of the experimental design had extraneous effects on judgments about thought disorder, contingency tables were constructed. For the first session, Tables 20, 21, 22, and 23 demonstrate no relation between the order in which the cases were presented and the thought disorder judgments for the two sessions. Because no significant relationships were found using Chi Square and Fisher's Exact Test, the exact test statistics have been omitted from Tables 20 through 27.

Table 20.  
Order Effects on Thought Disorder Judgments: Session 1  
Case S

		<u>No TD</u>	<u>TD</u>	
Case Order	S - F	7	9	16
	F - S	10	6	16
		17	15	32

Table 21.  
Order Effects on Thought Disorder Judgments: Session 1  
Case F

		<u>No TD</u>	<u>TD</u>	
Case Order	S - F	2	14	16
	F - S	3	13	16
		5	27	32

Table 22.  
Order Effects on Thought Disorder Judgments: Session 2  
Case S

		<u>No TD</u>	<u>TD</u>	
Case Order	S - F	13	3	16
	F - S	10	6	16
		23	9	32

Table 23.  
Order Effects on Thought Disorder Judgments: Session 2  
Case F

		<u>No TD</u>	<u>TD</u>	
Case Order	S - F	3	13	16
	F - S	6	10	16
		9	23	32

This lack of relationship also held true when judgments were compared with different orders of presentation from Session 1 to Session 2 (Tables 24 and 25). Half the subjects received the materials in the same

order for each session, e.g. Case S followed by Case F. The other half were given the cases in the opposite order of what they had received in the first session.

Table 24.  
New or Same Order Effects on Thought Disorder Judgments:  
Session 2 Case S

		<u>No TD</u>	<u>TD</u>	
Order	New	12	4	16
	Same	11	5	16
		23	9	32

Table 25.  
New or Same Order Effects on Thought Disorder Judgments:  
Session 2 Case F

		<u>No TD</u>	<u>TD</u>	
Order	New	4	12	16
	Same	5	11	16
		9	23	32

Interaction effects were also tested to see whether judgment of thought disorder for one case related to the judgment for the other case. For Sessions 1 and 2, Tables 26 and 27 demonstrate no significant relationship.

Table 26.

Case Interaction Effects: Session 1

		<u>Case F</u>		
		<u>No TD</u>	<u>TD</u>	
Case S	No TD	3	14	17
	TD	2	13	15
		5	27	32

Table 27.

Case Interaction Effects: Session 2

		<u>Case F</u>		
		<u>No TD</u>	<u>TD</u>	
Case S	No TD	6	17	23
	TD	3	6	9
		9	23	32

Thus, clinicians did not show a tendency to be biased in the assignment of thought disorder. Their judgments discriminated according to the two different cases. The significance of this finding is that it serves as evidence against the notion that the clinicians would tend to make similar judgments for different individuals. Within the author's experience, some clinicians have developed reputations for favoring a particular diagnosis or for consistently interpreting behavior in a certain way. Such a trend did not occur for this sample concerning the judgment of thought disorder.

In summary, the first hypothesis received no support. The clinicians followed the statements of the psychological reports rather than adhering to their own first judgments. Their certainty and consistency in judging the presence of thought disorder increased in response to confirming information. Disconfirming information was not disregarded; its effects appeared to be reasonable. There was a tendency for those who judged no thought disorder to be less likely to change their judgments and follow the reports. Clinicians who indicated the presence of thought disorder tended more to follow the new information. Subjects' attitudes toward reports also related to how they responded to them. These findings do not appear to relate to any extraneous aspects of the experimental method. Subjects demonstrated an ability to discriminate between the hypothetical cases.

#### Certainty Measures

A number of measures of certainty were used to test the second, third, and fourth hypotheses. These measures included certainty for the Thought Disorder item on the Symptom Checklist (TDDC); total summed score of certainty ratings on the Symptom Checklist (SCTOT); Diagnostic

Impression degree of certainty (DIDC); Diagnostic Selection Axis I degree of certainty (DSIDC); Diagnostic Selection Axis II degree of certainty (DSIIDC); and total number of items checked on Additional Information (ADDTOT). These measures were correlated with each other for each case (Tables 28 and 29). As these tables show, Additional Information proved to be unrelated to the other measures. Thus, it was not used as a measure of certainty for any of the hypothesis testing.

For certainty measures in Case S (Table 28), more significant intercorrelations occurred in the first session than in the second. As expected, the strongest relationships were found between the Diagnostic Impression and Diagnostic Selection Axis I ratings for both cases and sessions. In Case S, the relation between the diagnostic and Symptom Checklist ratings that was present in the first session decreased significantly in the second session. The thought disorder item certainty rating (TDDC) maintained its correlation with the Symptom Checklist (SCTOT), of which it was a part, but otherwise did not bear any relation to the other certainty measures in Case S. In comparing the ratings across sessions, two of the measures (TDDC and DSIDC) had low correlations.

The intercorrelations for Case F (Table 29) showed a much greater consistency within the second session.



Table 28.  
Correlation of Certainty Measures: Both Conditions Case S (N = 32)

	<u>Session 1</u>						<u>Session 2</u>					
	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>	<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
1. TDDC	.49*	.08	.23	.23	-.19	.07	.15	.25	.01	.16	-.06	.05
2. SCTOT		.46*	.51*	.51*	.14	.01	.20	.34*	.41*	.18	.34	.06
3. D IDC			.71	.71	.62*	-.10	.19	.28	.55*	.10	.60*	.02
4. DS IDC				.32	.32	-.11	.21	.23	.50*	.27	.54*	-.00
5. DSI IDC						-.12	.09	.21	.34	.08	.62*	-.02
6. ADDTOT							.05	.06	-.03	-.08	-.03	.47*
<u>Session 2</u>												
1. TDDC								.48*	.06	.08	-.08	-.02
2. SCTOT									.10	.14	.05	-.04
3. D IDC										.47*	.63*	-.00
4. DS IDC											.18	.04
5. DSI IDC												.05
6. ADDTOT												

TDDC: Degree of certainty for Thought Disorder item on Symptom Checklist.  
SCTOT: Total sum of certainty ratings on Symptom Checklist.  
D IDC: Diagnostic Impression degree of certainty.  
DS IDC: Diagnostic Selection degree of certainty (Axis I).  
DSI IDC: Diagnostic Selection degree of certainty (Axis II).  
ADDTOT: Total number of Additional Information items checked.

\* $p < .01$ , Kendall's Tau B.

Table 29.  
Correlation of Certainty Measures: Both Conditions Case F (N = 32)

	<u>Session 1</u>						<u>Session 2</u>					
	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>		<u>1</u>	<u>2</u>	<u>3</u>	<u>4</u>	<u>5</u>	<u>6</u>
1. TDDC	.32	.32	.30	.34	-.02		.42*	.26	.49*	.49*	.37	-.01
2. SCTOT		.43*	.41*	.32	-.01		.12	.55*	.24	.18	.19	-.10
3. DIDC			.89*	.48*	.16		.26	.43*	.36	.33	.32	.22
4. DSIDC				.44*	.11		.24	.45*	.38	.38	.34	.16
5. DSIIDC					-.12		.08	.30	.44	.44*	.58*	-.04
6. ADĐTOT							-.13	-.00	-.03	-.15	-.12	.41*
	<u>Session 1</u>						<u>Session 2</u>					
1. TDDC								.24	.52*	.59*	.45*	-.21
2. SCTOT									.38*	.31	.35	-.01
3. DIDC										.90*	.73*	-.10
4. DSIDC											.67*	-.17
5. DSIIDC												-.09
6. ADĐTOT												

TDDC: Degree of certainty for Thought Disorder item on Symptom Checklist.  
 SCTOT: Total sum of certainty ratings on Symptom Checklist.  
 DIDC: Diagnostic Impression degree of certainty.  
 DSIDC: Diagnostic Selection degree of certainty (Axis I).  
 DSIIDC: Diagnostic Selection degree of certainty (Axis II).  
 ADĐTOT: Total number of Additional Information items checked.

In general, for both sessions, the certainty measures were more related with one another for Case F as compared to Case S. This included the relation of TDDC with the other ratings.

### Hypothesis #2

The second hypothesis, based upon the research literature, was that clinician confidence would increase from the first to the second session. This increase would be independent of whether subjects received confirming or disconfirming information.

To test this hypothesis, certainty scores for the two sessions were compared with one another. Table 30 lists the mean scores by session for both conditions combined. Tables 31 and 32 list the scores for the confirming and disconfirming conditions, respectively.

Most of the ratings did not change significantly between sessions. The only significant change was the decrease in certainty for the Thought Disorder item for Case F. This decrease primarily occurred in the disconfirming condition. Thus, the second hypothesis was not supported.

Table 30.  
Mean Certainty Scores: Both Conditions

<u>Measures</u>	<u>Session 1</u>		<u>Session 2</u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
<u>Case S (N = 32)</u>				
TDDC	4.02	1.41	3.88	1.54
SCTOT	62.08	9.86	64.30	9.70
DIDC	3.72	1.08	3.84	1.07
DSIDC	3.41	1.19	3.88	1.01
DSIIDC	3.56	1.24	3.59	1.10
<u>Case F (N = 32)</u>				
TDDC	4.50	1.24	3.81*	1.42
SCTOT	68.11	9.03	65.81	8.86
DIDC	4.10	1.04	4.13	1.06
DSIDC	4.06	0.98	4.19	1.06
DSIIDC	3.97	1.28	3.94	1.31

- TDDC: Degree of certainty for Thought Disorder.  
 SCTOT: Total sum of certainty ratings on Symptom Checklist.  
 DIDC: Diagnostic Impression degree of certainty.  
 DSIDC: Diagnostic Selection degree of certainty (Axis I).  
 DSIIDC: Diagnostic Selection degree of certainty (Axis II).

\* $p < .01$ , Wilcoxon Matched-Pairs Signed-Ranks Test,  
 Session 1 versus Session 2.

Table 31.

Mean Certainty Scores: Confirming Condition

<u>Measure</u>	<u>Session 1</u>		<u>Session 2</u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
<u>Case S (n = 16)</u>				
TDDC	3.78	1.64	3.88	1.71
SCTOT	60.22	11.30	63.62	10.31
DIDC	3.56	1.26	3.62	1.31
DSIDC	3.12	1.31	3.75	1.06
DSIIDC	3.31	1.49	3.31	1.30
<u>Case F (n = 16)</u>				
TDDC	4.69	1.14	4.56	0.96
SCTOT	66.88	9.84	67.06	8.70
DIDC	4.33	0.82	4.50	0.89
DSIDC	4.25	0.93	4.62	0.72
DSIIDC	4.00	1.20	4.25	1.06

- TDDC: Degree of certainty for Thought Disorder.  
SCTOT: Total sum of certainty ratings on Symptom Check-list.  
DIDC: Diagnostic Impression degree of certainty.  
DSIDC: Diagnostic Selection degree of certainty (Axis I).  
DSIIDC: Diagnostic Selection degree of certainty (Axis II).

Table 32.  
Mean Certainty Scores Disconfirming Condition

<u>Measure</u>	<u>Session 1</u>		<u>Session 2</u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
<u>Case S (n = 16)</u>				
TDDC	4.25	1.13	3.88	1.54
SCTOT	63.94	8.10	64.97	9.33
DIDC	3.88	0.89	4.07	0.70
DSIDC	3.69	1.01	4.00	0.97
DSIIDC	3.81	0.91	3.88	0.81
<u>Case F (n = 16)</u>				
TDDC	4.31	1.35	3.06*	1.44
SCTOT	69.34	8.26	64.56	9.13
DIDC	3.88	1.20	3.73	1.10
DSIDC	3.88	1.02	3.75	1.18
DSIIDC	3.94	1.39	3.60	1.50

TDDC: Degree of certainty for Thought Disorder.  
 SCTOT: Total sum of certainty ratings on Symptom Checklist.  
 DIDC: Diagnostic Impression degree of certainty.  
 DSIDC: Diagnostic Selection degree of certainty (Axis I).  
 DSIIDC: Diagnostic Selection degree of certainty (Axis II).

\* $p < .01$ , Wilcoxon Matched-Pairs Signed-Rank Test,  
 Session 1 versus Session 2.

Hypothesis #3

The third hypothesis predicted that psychiatrists would show more confidence than psychologists. As in the previous hypothesis, different certainty scores were used to access this prediction. Table 33 lists mean certainty scores for the two professional groups for each session, respectively. The non-faculty and faculty groups were combined for each profession to form each of the professional groups. To test the hypothesis, their scores were compared using the Mann-Whitney U Test.

The third hypothesis was supported for three of the measures of Case F. Diagnostic Impression (DIDC), Axis I Diagnostic Selection (DSIDC) and Axis II Diagnostic Selection (DSI IDC) for Session 1 each were rated with more certainty by the psychiatric group ( $p < .01$ ). None of the certainty ratings in Case S or in the second session for Case F showed significant differences by profession. Given that Case F was the more disturbed patient, the psychiatrists demonstrated more confidence in their diagnostic judgments of severe psychopathology.

Although Additional Information has been determined not to measure degree of certainty, it is interesting to note that psychiatrists consistently requested more

information than did psychologists. This was likely due to requests for medically related forms. Table 34 lists the mean number of items requested per case for each session by the two professional groups.



Table 33.  
Certainty Means by Profession

Case S	<u>Session 1</u>				<u>Session 2</u>			
	Clinical Psychology <sup>a</sup>		Psychiatry <sup>a</sup>		Clinical Psychology <sup>a</sup>		Psychiatry <sup>a</sup>	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
TDDC	4.00	1.25	4.06	1.61	4.00	1.32	3.75	1.77
SCTOT	61.59	10.43	62.56	9.58	65.25	10.68	63.34	8.85
DIDC	3.75	1.18	3.69	1.01	3.53	0.99	4.12	1.09
DSIDC	3.25	1.29	3.56	1.09	3.56	1.03	4.19	0.91
DSI IDC	3.62	1.36	3.50	1.15	3.44	1.21	3.75	1.00
 <u>Case F</u>								
TDDC	4.44	0.96	4.56	1.50	3.81	1.47	3.81	1.42
SCTOT	67.75	8.11	68.47	10.12	65.50	10.18	66.12	7.64
DIDC	3.47	1.06	4.69*	0.60	3.94	1.12	4.33	0.98
DSIDC	3.50	1.03	4.62*	0.50	4.00	1.03	4.38	1.09
DSI IDC	3.40	1.18	4.50*	1.15	3.80	1.32	4.06	1.34

- TDDC: Degree of certainty for Thought Disorder item on Symptom Checklist.  
 SCTOT: Total sum of certainty ratings on Symptom Checklist.  
 DIDC: Diagnostic Impression degree of certainty.  
 DSIDC: Diagnostic Selection degree of certainty (Axis I).  
 DSI IDC: Diagnostic Selection degree of certainty (Axis II).

<sup>a</sup>n = 16 for each group.

\*Psychiatrists' certainty means greater than psychologists' (p < .01, Mann-Whitney U Test).

Table 34.  
Number of Additional Information Items Requested by  
Profession

<u>Profession</u>	<u>Case S</u>				<u>Case F</u>			
	<u>Session 1</u>		<u>Session 2</u>		<u>Session 1</u>		<u>Session 2</u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
Clinical Psychology	8.38	2.92	4.74	1.44	8.06	3.09	4.94	1.69
Psychiatry	11.12	4.96	7.19	3.75	10.04	4.91	8.56	5.18

Note: Additional Information ranged from 0 to 22 items.

#### Hypothesis #4

The fourth hypothesis predicted that there would be a relationship between certainty in judgment and clinical experience. To test the hypothesis, the non-faculty clinical psychologists' scores were combined with the psychiatric residents' scores for comparison with the combined faculty ratings. Table 35 lists the mean scores for these groups for each session, respectively. Of the twenty comparisons (five certainty measures within two sessions and two cases), the only significant difference was for the degree of certainty for the Thought Disorder item (TDDC) in the first session of Case S. For this rating, the experienced clinicians expressed more certainty in their judgments. An interesting aspect of

Table 35.  
Certainty Means by Experience

	<u>Session 1</u>				<u>Session 2</u>			
	<u>Non-faculty<sup>a</sup></u>		<u>Faculty<sup>a</sup></u>		<u>Non-faculty<sup>a</sup></u>		<u>Faculty<sup>a</sup></u>	
	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>	<u>Mean</u>	<u>SD</u>
<u>Case S</u>								
TDDC	3.31	1.45	4.80*	0.94	3.44	1.21	4.31	1.74
SCTOT	60.22	8.46	63.94	11.05	61.84	6.28	66.75	11.92
DIDC	3.50	1.15	3.94	0.94	3.87	1.25	3.81	0.91
DSIDC	3.12	1.15	3.69	1.20	3.75	1.00	4.19	0.75
DSIIDC	3.50	1.26	3.62	1.26	3.69	1.01	4.13	1.25
<u>Case F</u>								
TDDC	4.38	1.20	4.62	1.31	3.88	1.54	3.75	1.34
SCTOT	67.38	10.22	68.84	7.92	66.59	10.00	65.03	7.82
DIDC	4.13	1.13	4.06	1.00	4.00	1.00	4.25	1.12
DSIDC	3.94	1.18	4.19	0.75	4.12	1.15	4.25	1.00
DSIIDC	3.81	1.33	4.13	1.25	3.73	1.28	4.12	1.36

TDDC: Degree of certainty for Thought Disorder item on Symptom Checklist.

SCTOT: Total sum of certainty ratings on Symptom Checklist.

DIDC: Diagnostic Impression degree of certainty.

DSIDC: Diagnostic Selection degree of certainty (Axis I).

DSIIDC: Diagnostic Selection degree of certainty (Axis II).

<sup>a</sup><sub>n</sub> = 16 for each group.

\*Faculty's certainty means greater than Non-faculty's (p<.01, Mann-Whitney U Test).

this difference is that this particular symptom decision was one of the most evenly divided choices in the study. It would make sense that certainty would

be low. Thus, in this instance, the more experienced clinicians appeared less reasonable in their confidence. Other than this one rating, there was no significant support to the prediction that degree of certainty would vary with clinical experience.

## CHAPTER V DISCUSSION

Readers of the human judgment literature have grown accustomed to revelations of judgmental fallibility. Recent reviews, e.g. Hogarth (1980), have been virtual catalogues of common errors found in laboratory studies. Each type of judgmental error has received a name so as to become an entity in itself and an object for citation and replication. Thus, ignorance of base rates (Kahneman & Tversky, 1980), overconfidence (Einhorn & Hogarth, 1978), primacy and recency effects (Luchins, 1957, 1958) and "thirst for confirming redundancy" (Hogarth, 1980) are examples of biases commonly found among subjects and attributed to the general population. At times, the research appears to have as its aim the documentation of every manner in which people fall short of a rational computer model of judgment and decision making. Although the stated purpose of researchers has been to improve human judgment and provide aids toward that end, the process taken has been highlighted by the discrediting of that judgment.

This state of affairs finds a parallel in the clinical versus statistical prediction controversy in clinical psychology (Meehl, 1954; Holt, 1970). Whereas actuarial prediction was meant as an aid to the clinician, the history of the controversy found the two methods pitted against one another. The actuarial side consistently "won" each comparison, according to the methods by which they were contested. What could have been recognized as ways to improve prediction soon became more known for the discrediting of clinical judgment. The clinician was then held up as the "loser" instead of the beneficiary of new technology.

In a similar sense, the human judge has been portrayed as a defective instrument marred by irrational biases. The literature responsible for this portrayal persuaded the author to advance the hypotheses of this study. The tendency to ignore disconfirming data appeared to be robust and likely to occur in clinical judgment. Thus, the first hypothesis was formulated, stating that clinicians would retain their initial judgments about the presence of thought disorder, even in the face of disconfirming evidence.

What occurred instead was quite striking. Clinicians responded in a rational, conservative manner, without exhibiting the predicted irrational bias to maintain their

opinions. Their judgments were discriminative of the information presented to them and were consistent with their reported opinions of psychological reports.

Moreover, in the instances when disconfirming information was not followed, there seemed to have been several rational factors involved, rather than biased adherence to the original judgment. First, in this ambiguous situation, the clinicians opted for the more conservative choice, i.e. no thought disorder. Second, the presence of thought disorder in Case F, based on the initial information, was so compelling that the disconfirming psychological report could not change the first impression. Third, some subjects did not value psychological reports highly enough to be persuaded to alter their judgments. This last reason applied to a subgroup of the faculty psychologists who were less impressed with the usefulness of the reports compared to other subject groups. The use of any of these three factors by the clinicians who did not follow the reports represented a rational basis for making a judgment. Thus, the irrational behavior of ignoring information, often cited in the literature, did not occur in this study.

The second hypothesis, that confidence would increase with additional information, received no support either. Even in the confirming condition there was no increase in certainty. In the disconfirming condition, one of the

ratings decreased from the first to the second session, which is a result that makes rational sense, but is contrary to the hypothesis. Oskamp's (1965) results of increasing certainty over time were thus not obtained here.

The third and fourth hypotheses, concerning certainty differences according to profession and clinical experience, respectively, also were not generally supported. Although the psychiatrists did show more confidence in their initial diagnosis of the more disturbed patient, there were no substantial differences between the two groups when all the possible comparisons in confidence were taken into consideration. Even fewer significant differences were found between levels of experience.

These last two hypotheses were not empirically based or strongly suggested by the literature. Rather, they were offered by the author as interesting areas to explore. The first two hypotheses, however, were drawn from the cited literature. That they were so soundly refuted comes as a surprise. The trained clinicians who participated in this study neither ignored information nor did they overstate their confidence, as has occurred in other studies (Oskamp, 1965; Schinka, 1976).

Instead, judgments tended to be discriminating and consistent. DSM-III categories were specifically selected



in a differential manner, often with the inclusion of the "fifth digit" for a precise diagnosis. The cases were logically designed for many possible diagnostic interpretations; the subjects indeed provided a diverse set of diagnostic judgments. Clinicians also named several different diagnoses that included thought disorder as a presenting feature, rather than equating the presence of thought disorder with schizophrenia.

In addition, clinicians clearly distinguished the two cases, which brought different patterns of responses. There was no tendency to make stereotypic judgments across cases. Thus, the judges' responses for one case did not affect their responses for the other case.

These results support the claim of Christensen-Szalanski and Beach (1984), who highlighted the rational behavior of human judges. These authors demonstrated that research which showed irrational biases of judges had received disproportionate citation in the literature. They advocated a more balanced emphasis in future publications. Perhaps a reason for this imbalance is that systematic irrational behavior is more intriguing than practical rational behavior. Just as a psychiatric patient's abnormal behavior attracts more attention than his or her normal responses, even though the latter often constitute a greater proportion of total observed behavior, so also do judges'

aberrations become more noticeable than their correct judgments. In fact, the author may not have chosen to perform the present study if the prediction had been that subjects would simply behave in a rational manner. This question of rational versus irrational choices is an interesting one that should become prominent in future studies of human judgment.

Some other aspects of this study also proved interesting. Besides displaying rationality in judgment, the clinicians also were prudent in their assignment of thought disorder. Such a judgment could have serious life consequences for a real patient because of the well known effects of psychiatric labelling. In this study, the conservative nature of clinicians' judgments was manifested in the tendency both to opt for selecting no thought disorder when the information was conflicted and in the readiness to select no psychiatric disorder, especially in the initial judgment. Thus, for a number of the clinicians, when the patient's status was in doubt, a judgment of serious psychopathology was deferred.

Given that the results speak well of clinical judgment, what can be inferred about the use of psychological reports? In general, the role of the reports in the clinicians' judgments was mixed. Predominantly, the reports were followed. As measured by responses to one piece of

information alone, i.e. cognitive functioning, the reports demonstrated the capacity to alter clinical judgments. On the other hand, the clinicians gave comparatively low ratings for the reports' usefulness. This may have been because half the reports in fact disagreed with a crucial judgment. The low utility rating would have been a means of resolving cognitive dissonance via devaluation of conflicting information. Alternatively, the clinicians may have hoped for the reports to relieve them of the high degree of ambiguity that the cases were presenting. Except for the evaluation of cognitive functioning, the reports were not designed to resolve the diagnostic dilemmas. Some subjects, mostly the psychologists, expressed disapproval of the reports' psychodynamic emphasis. Thus, another reason for disapproval may have been stylistic preference. As already noted, some of the faculty psychologists, whose group valued reports the least, remained consistent with their opinions by not following the reports. Overall, the reports had an impact, with different factors mitigating their influence.

The results of this study raise some questions about the experimental methodology. As was expected, the creation of hypothetical psychiatric patients who would elicit an even split on a crucial clinical judgment was not easy.

The decision to provide clinicians with two cases instead of one proved wise, as one of the two successfully divided clinical opinions almost exactly in half. Any subsequent study requiring ambiguous cases would best include more than one case to enhance the chances that one of the cases would attain a high level of diagnostic difficulty.

Another possible methodological problem was the use of the six-point rating scales for measurement of certainty. Although such scales had been used in other studies, e.g. Cantor et al. (1980) and Eker (1981), there was no prior determination of their reliability or validity. Since Additional Information failed as a convergent measure of certainty, the question remains as to the meaningfulness of these scales.

A more important problem of experimental design was that reports were probably more salient than they might be in normal clinical practice. The reports clearly stood out as a focus of interest in the second session. Even though other typical clinical information was included, the reports had such a prominent place in the study that they may have been too evident as key aspects of the experiment.

A related issue was the question of recency effects. Even though subjects were explicitly informed to consider

information from both sessions and the Recall Notes were provided at the second session, the concrete presence of the psychological report may have exerted an impact beyond that of the previous information. The study attempted to approximate a naturalistic clinical situation. However, a psychotherapist in a hospital setting would probably have more of a personal investment in his or her clinical opinion of a patient. The experimental situation involving two discrete sessions could not fully reflect the commitment of judgment in a real clinical setting. Although a cover story of the therapist's vacation was meant to rationalize the timing of the experiment, it nonetheless could not eliminate the study's inherent artificiality. Thus, the generalizability of the study is uncertain, as is true for most exploratory investigations. This conclusion is not so much a criticism of the method as it is an impetus for replication and further exploration.

With due reference to these several shortcomings, the results of this study still point to the rational judgments made by the participating clinicians. Their pattern of responses stands in contrast to the predominantly cited work in the human judgment literature. These findings would suggest that more naturalistic studies of judgment should be conducted to test further both the alleged biases of clinical judges and the role of psychological reports in psychiatric decision making.

APPENDIX A  
ADMITTING NOTES

Case S: Sherry

Identifying Information: Sherry is a 21-year-old single white female who arrived via ambulance at 1:00 a.m. accompanied by boyfriend. Patient was transferred from Central Hospital ER where she had been treated following reported ingestion of pills. After being medically cleared, she was immediately transferred here.

Chief Complaint: "I don't know what's going on. Nothing makes sense."

History of Present Illness: Informant was boyfriend, a 23-year-old white male, sloppily dressed in jeans, boots, t-shirt and leather jacket; strong body odor. Informant seemed impatient to leave and gave brief answers to questions. Reliability of history is questionable.

Boyfriend stated he found patient yesterday morning lying on the floor in his apartment in a semi-conscious state. Next to her was a spilled bottle of pills which he claimed he could not identify. She had left him a note stating that she had "a new vision" that had told her that she must "go for it alone for you."

Informant claimed to have known patient for past 8 months. He denied any drug abuse, except, "Sure, we have a joint once in a while, like everybody else." He claimed no knowledge of any previous overdoses by patient. When questioned about her background, he stated that he "did not know her too well," but that he remembered her frequently asking questions such as, "Do you think I'm weird?" when they first met. He mentioned that her grandmother had died several months before and that a sister of the patient had recently moved away from the area. Also, two weeks prior to admission the patient had attended an "est" weekend seminar. He claimed that she had been "weird" since returning from the meeting. "She couldn't stop talking and she acted real cool, like all of a sudden she knew all the answers or something, but she was weird." He described the patient as being extremely talkative, restless and had not slept normally for the two weeks following the "est" weekend. She had stopped going to her job as a cocktail waitress. He could not recall her behaving this way before.

The patient denied having previous psychiatric treatment, but murmured something about a "counselor." She refused to discuss family or early personal history. Boyfriend stated that he did not know her family: "They live out of state."

Mental Status Exam: Young adult female dressed casually; unresponsive to questioning, appears sleepy, head down with no eye contact, lethargic movements. Blunted affect, moderately depressed mood, speech slurred, thought blocking. Denies hallucinations, but states, "Voices? If I could translate what they said, who would believe me? No, there's nothing to hear but in here." Patient said nothing further and seemed preoccupied. Judgment poor.

## Case F: Fred

Identifying Information: Fred is a 24-year-old single white male who resides with his mother and younger sister. He was brought to the hospital by the police, accompanied by his mother, after having spent three days alone in his room without eating.

Chief Complaint: "My spirit died." "I am going through the motions."

History of Present Illness: The patient's mother reported that he had been "acting strange" during the previous two weeks. He had withdrawn from social contact and could be heard talking loudly when alone in his room. He refused to eat with his family and stopped receiving any phone calls. His employer, the owner of a local fast food place, called several times. Fred refused to speak with him. He stopped speaking to anyone and began spending all of his time alone in his room or in the garage. On Friday, while working on the family car, the carburetor caught fire. He became very upset, started shouting. Sister put out fire with extinguisher. The patient then became very quiet and "just stared straight ahead." Without helping to clean up the mess, he walked upstairs to his room and locked the door. He would not answer to repeated attempts by mother and sister to get him to open door. Mother and sister said they felt "disgusted" with his behavior, so they "let him sit" in his room on Saturday. By Sunday, however, they "began to worry." When he did not respond to their pounding on the door, they called the police, who came and forced door open. They found him lying in bed staring at ceiling in the same clothes he had worn on Friday. He was unshaven and had soiled clothing. The police recommended that his mother have him committed for emergency observation at the hospital. When arrangements were made, the police carried patient to their car and into the hospital.

Patient's mother described him as always having been a "loner" with few friends. She said that after leaving high school he had "taken up with the wrong crowd" and had often been away from home "for days at a time." Most of his jobs had been at night "because he works better when there are not a lot of people around." His longest job was



eight months as a night janitor. He quit that because he felt "too much pressure from his boss." His mother stated that he "has been unsure about what career he wants. He works well with things rather than people."

His mother could not recall any serious difficulties in patient's childhood. He was "always a quiet, well-behaved boy." When his father died when patient was 13, "he didn't even cry." Mother stated, "I used to tell everyone, 'Isn't that something. He takes his father's death like a real man.'"

Mother reported no previous psych. history for patient. "He never did nothing like this before."

Mental Status: Young adult male, sloppily dressed, poor hygiene, unshaven. Blunted affect, thought blocking, slow motor response, inappropriate staring. Stated that he sees "dark things sometimes," but would not elaborate. Denies auditory hallucinations. Failed simple calculations, repeated 7 digits. Proverbs concrete. Judgment poor.

APPENDIX B  
TAPED INTERVIEW TRANSCRIPTS

Case S: Sherry

"T" = Therapist (interviewer)

"S" = Patient (Sherry)

(Sherry's initial responses are slow, hesitant, with depressed affect. As the interview continues, she seems more anxious, irritable and stressed by the questions. Her speech is in rapid bursts. She gives the impression of a sarcastic, angry little girl who is afraid of losing control if she were to fully express herself).

- T: Good morning, Sherry. I'm Dr. Smith. Could we talk for a few minutes?
- S: About what?
- T: Well, to begin, what brings you here to the hospital?
- S: (sighs) I already told sixteen different interrogators--do you want me to confess one more time? Why are you tormenting me? Haven't I paid enough for my sins already? Maybe I haven't....
- T: Well, I certainly do not want to torment you. You're feeling tormented?
- S: (softly) Yes.
- T: What are you feeling tormented about?
- S: Oh, it's a long story. Why don't you read my chart.
- T: I prefer to hear directly from you.
- S: You'll never get it. You have to go through it to get it.
- T: Get what?
- S: Never mind.
- T: Have you ever felt this way before?
- S: I don't know. How am I supposed to remember?  
(pause)
- T: Perhaps you can recall a time when you did not feel this badly.
- S: (pause) Well....I once thought I was Ms. Superwoman--like, wow, undomitable, incredible, abominable, you know, like they couldn't stop me. On the street they knew me and everybody said, "Hey, Sher, what's happenin'!" and we would strut our bad stuff. (pause) But, what's the use of this good-old-days crap. I don't want to talk about it.
- T: It's hard remembering a time when you felt better?
- S: Really. And then it always crashes, like what's the use of opening the window at six-thirty in the morning?

- T: Are you saying that you have had times of feeling high and then times of feeling very down?
- S: Yeah. It's in my chart. You've got me all figured out, Doc. Put me in the hole with the other pigeons. Or canaries. Or ravens--I'm a raven.
- T: You're smiling. A raven?
- S: Yeah, ravin' mad. That's why I'm here, right? What's the matter? Can't you laugh at a joke?
- T: I understand that you came here following a drug overdose. Why did you take the pills?
- S: Please.....I've already confessed. Do you want it signed in blood? I'll say seventy-seven Hail Mary's. Anything.
- T: Have you ever tried to hurt yourself before?
- S: Not really. I scratched my wrists once. It was no big deal. I was just a kid. A lamb, a sacrificial lamb.
- T: Have you ever seen a psychiatrist, psychologist, or any other counselor before?
- S: My mother tried to make me go to a "psycho-man." I went once and he was weird. He just sat there. He tried to read my mind, but I wouldn't let him.
- T: Could you tell me more about your family?
- S: My grandmother died six months ago. Then my older sister moved to California.
- T: What about your parents?
- S: I've had it with my mother. Please don't make me talk about her. I never knew my old man--if he ever existed.
- T: What do you mean?
- S: I don't remember.....my father left when I was 18 months--at least that's what my mother said. She's the sick one; she should get the inquisition. She remarried when I was five or something. I've got two half-sisters and one real sister. (speaking more rapidly) If you subtract me, that's one-and-three quarters. Mother always said she would give two bits for somebody to take me away. Now the dream came true. Hey, you don't have to write..... hey, come on, don't write that down. I'll bet you didn't know I went to college...(pause)...(in a soft voice)..Oh, God help me.....
- T: Sherry, how do you think we can help you?
- S: (pause) How can I get it if you haven't got it? Oh, God...(sound of chair moving, footsteps, door slamming).

## Case F: Fred

"T" = Therapist (interviewer)

"F" = Patient (Fred)

(Fred's responses are characterized by long pauses, monotone voice.)

T: Could you tell me what brings you to the hospital?  
F: (pause) The police brought me.  
T: The police brought you? Why did they do that?  
F: My mother.....told them....  
T: Yes?  
F: Yes.  
T: Well, what were you doing that would prompt your mother to call the police?  
F: Nothin'.....I feel tired.....  
T: You do seem to be sleepy. How have you been sleeping?  
F: Not good.....Last night I saw black skeletons.....Can I go to my room?  
T: Black skeletons?  
F: Yeah.  
T: Could you tell me more about them?  
F: (pause) They were.....black.....(longer pause)  
T: Well, what were they doing?  
F: Nothin'.  
T: Have you ever seen anything like them before?  
F: No.  
T: Were you awake or were you dreaming?  
F: They were at the end of the bed.  
T: Can you remember whether you were asleep?  
F: I saw them. Maybe it was a dream.  
T: How did you feel when you saw them?  
F: Scared.  
T: What do they make you think of?  
F: Maybe I did something...I feel tired.....  
T: It seems to be hard for you to talk about the skeletons.  
F: There's nothin' to talk about. It's all over.  
T: All over?  
F: Me. I'm done.  
T: What do you mean?  
F: That's right.....I'm mean.  
T: I'm not sure that I understand.  
F: I was keen. Now I'm mean.

T: What makes you believe that you are mean?  
F: The evidence is all there. Case closed. It's no use in talking about it.  
T: It sounds like it is difficult for you to talk about these things. Do I hear you saying that you are feeling guilty, as though you think you've done something wrong?  
F: It's no use in talking about it. You are just wasting your time on me.  
T: Well, we don't have to talk about that now. Would you mind if I just asked a couple more questions?  
F: Your time must be valuable. I don't want you to waste it on me.  
T: Have you ever seen a therapist or counselor for any other problems before?  
F: No.  
T: When did you start feeling the way you do now?  
F: I lose track of time. It seems like I've always been this way.  
T: What is today's date?  
F: Like I said, I lose track of time.  
T: Could you tell me where we are?  
F: In between these walls in this room.  
T: In this room? What building is this room in?  
F: This is a hospital, isn't it?...Can I go now?  
T: You're shaking your head.  
F: What's the use. You probably think I'm nuts or somethin'. I just can't make it at anything.  
T: What makes you feel that you can't make it?  
F: Well, what have I ever done? I can't even fix a car without burning it up.  
T: Did anyone ever teach you how to work on a car?  
F: No. Not really. But every guy should be able to at least tune it up. I can't concentrate.  
T: Well, how can you learn unless someone teaches you?  
F: Oh, I don't know. You're just trying to make me feel better.  
T: Fred, how do you think we can help you?  
F: Too late. Like in the song: "Don't talk about Fred. 'cause Fred is dead."

APPENDIX C  
NURSING REPORT SUMMARIES

Nursing Report Summary (Case S)

Sherry has been kept off medication by Dr. Jones, who has been covering the case. Substance-induced organic mental disorder has been ruled out, although extent of previous substance abuse is still questionable. Family history has not been obtained; staff were unable to contact the patient's relatives. Patient has stated, "Keep my family out of this, please. My insurance will pay for this."

Nursing staff noted that the patient was preoccupied, kept to herself during the first several days of hospitalization. On the ward she was seen both as depressed and irritable, with brief expansive moods. At times her speech did not make sense. She had a tearful discussion with one of the psychiatric aides who worked the night shift. She revealed that her boyfriend had actually been spending most of his time away from her for the past two months and that she thought he might be with another girl. She seemed to imply that this "is the same old story" for her. However, she appeared elated when the boyfriend visited unexpectedly one evening a week after her admission. Later she became tearful and withdrawn. The boyfriend had not since returned; the patient had not made any apparent attempt to contact him, either. By the second week she was noted to be spending much time with a young, withdrawn male patient. Nursing staff noted that the patient appeared anxious prior to psychological testing and agitated afterwards. Her participation in group activities had been minimal, but she had not been seen as a behavior management problem by the staff.

## Nursing Report Summary (Case F)

Fred had to be persuaded to participate in activities. He left group therapy and other activities several times after only several minutes of participation. Spent much time alone in room. When staff locked the door to prevent him from doing this he alternated from slowly walking the hallways to sitting in front of the TV, frequently smoking cigarettes. He was observed to eat sparingly at meals. Evening staff reported that he would eat large amounts of snacks. When other patients complained, staff set limits on his snacking.

Fred's mother visited every night, usually bringing a large milkshake and cookies. Patient did not appear to interact with mother. He was observed eating the food she brought after she would leave. Efforts by staff to talk with patient brought little response. He usually avoided eye contact and at times would appear as though he were about to cry, although he was never observed to express feelings directly. Other patients have not interacted with patient. Some have complained of his poor hygiene.

APPENDIX D  
PSYCHOLOGICAL REPORTS

I. Thought disorder present

Psychological Report

Patient: Sherry J.

Dates of Testing: August 3 and 4, 1983

Reason for Referral: Sherry was referred for a general evaluation of cognitive and personality functioning as part of the team diagnostic workup.

Tests Administered: Bender Gestalt, WAIS-R, Human Figure Drawings, TAT, Rorschach.

Behavioral Observations: Sherry presented as a thin, fair-skinned young woman, wearing a plain blouse, jeans and hospital slippers. She sat quietly in her chair throughout the testing sessions, responding only when directly prompted. She smoked several cigarettes during each session and seemed content to have the ashes fall onto the floor until the examiner offered her an ashtray. She displayed little affect during testing. Twice she smiled while observing TAT pictures, but did not offer any explanation as to why she grinned. The only noticeable change in her demeanor occurred after testing had been completed, when the examiner asked her about her experience with "est." Her eyes widened and she appeared startled. She quickly asked, "Are you a trainer?" Apparently this was a reference to the leaders of "est" workshops. At that point she did not respond further, despite assurances by the examiner that he had no affiliation with "est."

Test Results: Sherry's performance on cognitive measures places her in the average range of intellectual functioning (Verbal IQ = 101; Performance IQ = 112; Full Scale IQ = 107). Closer examination of her response style suggests that this may be a conservative estimate of her intellectual potential. Her lowest performance occurred on subtests known to be sensitive to decrements in attention and



concentration. Also, she responded hastily to a number of items. When queried further, she often was able to provide correct answers.

Sherry's responses to personality measures reveal a highly guarded personality style. She attempts to mask the excessive amount of tension and painful affect that dominate her experience. As she is unable to relieve these internal pressures through appropriate emotional expression, she attempts to manage this tension in a cognitive manner. She is likely to be constantly striving to cope with stress in her own mind, with much active ideation. At times, she tries to distance herself from her painful feelings; her responses suggest that she resorts to reaction formation which succeeds only temporarily, with a subsequent lapse into feelings of despair.

The structure and content of her responses to projective measures suggest that Sherry is struggling with basic issues of separation. She feels caught in a symbiotic state such that efforts to distinguish herself from a maternal object representation are tantamount to an experience of being "torn apart." Among her most pressing feelings are those of vulnerability and a wish to be cared for, which become intolerable for her to face. Her pattern of responses suggests that she feels enraged for having these unmet needs. These aggressive impulses are also intolerable for her. Test responses suggest that she either introjects her aggression back upon herself in a punitive manner or she impulsively discharges her feelings, probably in a way that is ultimately self-destructive.

This affective turmoil impinges upon Sherry's cognitive processes. Her test responses suggest incoherence and distortion of reality. Idiosyncratic thinking and loose associations intrude upon her processes. Some responses suggest that this may occur in the context of dependent feelings that are evoked by sexual stimuli. She lacks the capacity to recover adequately; however, when not under stress, she is unlikely to present evidence of these cognitive problems.

Sherry's test responses reflect the constrained manner of her observed affect. Although indications of depression are in her test record, these alone do not constitute enough evidence for a diagnosis of severe depression. Continued observation of her behavior will clarify this assessment.

Summary and Recommendations: Test results and observed behavior suggest that Sherry is an intensely troubled young woman who cannot tolerate the painful experiences of basic unmet needs. She is likely to react to these impulses in a self-punitive manner and is at high risk for impulsive behavior. Although she did not seem suicidal or severely depressed at the time of testing, her weak defensive structure and self-punitive tendencies place her at risk for self-destructive behavior if subjected to moderate stress. Her test responses suggest that such stress would impair her normal cognitive functioning. She is in need of a structured therapeutic environment to protect her and help her cope with the feelings related to separation and individuation. It is recommended that Sherry be involved in long-term treatment to address these problems.

## I. Thought disorder present

## Psychological Report

Patient: Fred B.

Dates of Testing: August 1 and 2, 1983

Reason for Referral: Fred was referred for an evaluation of cognitive and personality functioning as part of the psychiatric team diagnostic workup.

Tests Administered: WAIS-R, Rorschach, TAT, Figure Drawings, Bender Gestalt.

Behavioral Observations: Pt. presented as a young male, dressed in faded t-shirt, jeans, slippers and hospital robe. He was unshaven and hair appeared uncombed. He was expressionless and moved slowly, as if fatigued. His first comment was, "Why are you doing this." Pt. yawned frequently during the assessment and asked several times if we were finished. He cooperated in a passive manner, at times mumbling his responses. Several times he said, "I'm sorry, man. I'm a sorry man."

Results: Pt.'s intellectual performance is at the lower end of the Dull Normal range. However, intellectual assessment at this time would not likely appropriately represent the pt.'s abilities. In particular, he did poorly on tasks which require speed and visual motor dexterity. He seemed to have trouble maintaining concentration to perform calculations and other cognitive tasks requiring a consistent train of thought.

Responses to personality measures indicate a regressed state of psychosocial functioning. Evidence of severe anxiety and depression dominates the test protocols. This emotional disturbance is occurring within the context of a personality style characterized by marked ambivalence and unexpressed inner conflict.

The patient's attempts at coping with his anxiety are likely to involve such defenses as constriction, denial and regression. These defenses offer an inadequate solution to his pervasive inner stress. Test responses suggest that he is likely to fantasize about acting out his aggressive feelings, but then despairs at being unable to do so. The ideation surrounding the feelings is primitive and intolerable to his self-concept. He is at a loss for directing his rage in a manner that would relieve him of his inner pain.

Under conditions of stress, his cognitive functioning is likely to deteriorate. This would include the tendency to misinterpret reality and an inability to maintain cognitive focus. At times his reasoning processes could be characterized by autistic logic and idiosyncratic thought. The content of his projective responses suggests that stress would especially occur in situations where he perceives that his competence is being observed by others. This would likely relate to his feelings of low self-esteem as a young man who is concerned about relating with women. His attitude towards women is one of fearful ambivalence. On the one hand he feels the strong need to be nurtured and cared for, while on the other hand he views females as objects not worthy of trust. He takes a passive approach to this conflict. His projective responses suggest that he resigns himself to a "wait and see" style of coping with his environment. Unfortunately, there is little evidence in the test data to suggest that he has the capacity or ego strength to take appropriate action once he has pondered his decisions. Thus, he is likely to feel psychologically paralyzed when placed in stressful social situations.

Summary and Recommendations: This troubled young man is still in the early stages of recovery from what he describes as his first "nervous breakdown." As he appears to be still in crisis and is demonstrating very little capacity to observe himself, he is probably not a good candidate for insight-oriented psychotherapy at this time. He is in need of a supportive therapeutic relationship within the safe confines of a structured environment. Continued stay in an inpatient setting at this time would be therapeutic in providing predictable structure to his life. Supportive therapy should give him an opportunity to test reality and reconstitute his cognitive functioning. Over time, this therapy may evolve into a more insight-oriented approach, depending upon his response to treatment.

## II. Thought disorder not present

## Psychological Report

Patient: Sherry J.

Dates of Testing: August 3 and 4, 1983

Reason for Referral: Sherry was referred for a general evaluation of cognitive and personality functioning as part of the team diagnostic workup.

Tests Administered: Bender Gestalt, WAIS-R, Human Figure Drawings, TAT, Rorschach.

Behavioral Observations: Sherry presented as a thin, fair-skinned young woman, wearing a plain blouse, jeans and hospital slippers. She sat quietly in her chair throughout the testing sessions, responding only when directly prompted. She smoked several cigarettes during each session and seemed content to have the ashes fall onto the floor until the examiner offered her an ashtray. She displayed little affect during testing. Twice she smiled while observing TAT pictures, but did not offer any explanation as to why she grinned. The only noticeable change in her demeanor occurred after testing had been completed, when the examiner asked her about her experience with "est." Her eyes widened and she appeared startled. She quickly asked, "Are you a trainer?" Apparently this was a reference to the leaders of "est" workshops. At that point she did not respond further, despite assurances by the examiner that he had no affiliation with "est."

Test Results: Sherry's performance on cognitive measures places her in the average range of intellectual functioning (Verbal IQ = 101; Performance IQ = 112; Full Scale IQ = 107). Closer examination of her response style suggests that this may be a conservative estimate of her intellectual potential. Her lowest performance occurred on subtests known to be sensitive to decrements in attention and concentration. Also, she responded hastily to a number of items. When queried further, she often was able to provide correct answers.

Sherry's responses to personality measures reveal a highly guarded personality style. She attempts to mask the excessive amount of tension and painful affect that dominate her experience. As she is unable to relieve these internal pressures through appropriate emotional expression, she attempts to manage this tension in a cognitive manner. She is likely to be constantly striving to cope with stress in her own mind, with much active ideation. At times, she tries to distance herself from her painful feelings; her responses suggest that she resorts to reaction formation which succeeds only temporarily, with a subsequent lapse into feelings of despair.

The structure and content of her responses to projective measures suggest that Sherry is struggling with basic issues of separation. She feels caught in a symbiotic state such that efforts to distinguish herself from a maternal object representation are tantamount to an experience of being "torn apart." Among her most pressing feelings are those of vulnerability and a wish to be cared for, which become intolerable for her to face. Her pattern of responses suggests that she feels enraged for having these unmet needs. These aggressive impulses are also intolerable for her. Test responses suggest that she either introjects her aggression back upon herself in a punitive manner or she impulsively discharges her feelings, probably in a way that is ultimately self-destructive.

Despite this affective turmoil, Sherry is able to maintain her normal cognitive processes. Even when stressed in the testing situation, she was able to respond in a coherent manner without any significant distortion of reality. Some responses suggest that she does experience confusion in the context of dependent feelings that are evoked by sexual stimuli. However, she demonstrates a capacity to recover such that her vulnerability is not serious enough to be considered a dysfunction in cognitive processing.

Sherry's test responses reflect the constrained manner of her observed affect. Although indications of depression are in her test record, these alone do not constitute enough evidence for a diagnosis of severe depression. Continued observation of her behavior will clarify this assessment.

Summary and Recommendations: Test results and observed behavior suggest that Sherry is an intensely troubled young woman who cannot tolerate the painful experiences

of basic unmet needs. She is likely to react to these impulses in a self-punitive manner and is at high risk for impulsive behavior. Although she did not seem suicidal or severely depressed at the time of testing, her weak defensive structure and self-punitive tendencies place her at risk for self-destructive behavior if subjected to moderate stress. Her test responses suggest that her cognitive functioning is intact; however, she is in need of a structured therapeutic environment to protect her and help her cope with the feelings related to separation and individuation. It is recommended that Sherry be involved in long-term treatment to address these problems.

## II. Thought disorder not present

## Psychological Report

Patient: Fred B.

Dates of Testing: August 1 and 2, 1983

Reason for Referral: Fred was referred for an evaluation of cognitive and personality functioning as part of the psychiatric team diagnostic workup.

Tests Administered: WAIS-R, Rorschach, TAT, Figure Drawings, Bender Gestalt.

Behavioral Observations: Pt. presented as a young male, dressed in faded t-shirt, jeans, slippers and hospital robe. He was unshaven and hair appeared uncombed. He was expressionless and moved slowly, as if fatigued. His first comment was, "Why are you doing this?" Pt. yawned frequently during the assessment and asked several times if we were finished. He cooperated in a passive manner, at times mumbling his responses. Several times he said, "I'm sorry, man. I'm a sorry man."

Results: Pt.'s intellectual performance is at the lower end of the Dull Normal range. However, intellectual assessment at this time would not likely appropriately represent the pt.'s abilities. In particular, he did poorly on tasks which require speed and visual motor dexterity. He seemed to have trouble maintaining concentration to perform calculations and other cognitive tasks requiring a consistent train of thought.

Responses to personality measures indicate a regressed state of psychological functioning. Evidence of severe anxiety and depression dominates the test protocols. This emotional disturbance is occurring within the context of a personality style characterized by marked ambivalence and unrepressed inner conflict.



The patient's attempts at coping with his anxiety are likely to involve such defenses as constriction, denial and regression. These defenses offer an inadequate solution to his pervasive inner stress. Test responses suggest that he is likely to fantasize about acting out his aggressive feelings, but then despairs at being unable to do so. The ideation surrounding the feelings is primitive and intolerable to his self-concept. He is at a loss for directing his rage in a manner that would relieve him of his inner pain.

Under common conditions of stress, his cognitive functioning is likely to remain intact. This would include the consistent ability to test reality and maintain appropriately focussed reasoning processes. The content of his projective responses suggests that stress would especially occur in situations where he perceives that his competence is being observed by others. This would likely relate to his feelings of low self-esteem as a young man who is concerned about relating with women. His attitude towards women is one of fearful ambivalence. On the one hand he feels the strong need to be nurtured and cared for, while on the other hand he views females as objects not worthy of trust. He takes a passive approach to this conflict. His projective responses suggest that he resigns himself to a "wait and see" style of coping with his environment. Unfortunately, there is little evidence in the test data to suggest that he has the capacity or ego strength to take appropriate action once he has pondered his decisions. Thus, he is likely to feel psychologically paralyzed when placed in stressful social situations.

Summary and Recommendations: This troubled young man is still in the early stages of recovery from what he describes as his first "nervous breakdown." As he appears to be still in crisis and is demonstrating very little capacity to observe himself, he is probably not a good candidate for insight-oriented psychotherapy at this time. Although there is no evidence that he has an impairment in cognitive functioning, he is in need of a supportive therapeutic relationship within the safe confines of a structured environment. Continued stay in an inpatient setting at this time would be therapeutic in providing predictable structure to his daily life. Over time, this therapy may evolve into a more insight-oriented approach, depending upon his response to treatment.

APPENDIX E  
BRIEF PSYCHIATRIC RATING SCALE

Please rate the patient for each dimension. It may be difficult to rate some of the items because of the limited information on the patient. Please give your best estimate.

"0" = Not present  
"6" = Extremely severe  
"NA" = Not applicable

- |                               |  |                  |
|-------------------------------|--|------------------|
| 1. Somatic Concern            | Degree of concern over present bodily health. Rate the degree to which physical health is perceived as a problem by the patient, whether complaints have a realistic basis or not.   | 0 1 2 3 4 5 6 NA |
| 2. Anxiety                    | Worry, fear, or over-concern for present or future. Rate solely on the basis of verbal report of patient's own subjective experiences. Do not infer anxiety from physical signs or from neurotic defense mechanisms.         | 0 1 2 3 4 5 6 NA |
| 3. Emotional Withdrawal       | Deficiency in relating to the interviewer and to the interview situation. Rate only the degree to which the patient gives the impression of failing to be in emotional contact with other people in the interview situation. | 0 1 2 3 4 5 6 NA |
| 4. Conceptual disorganization | Degree to which the thought processes are confused, disconnected or disorganized. Rate on the basis of integration of the verbal   | 0 1 2 3 4 5 6 NA |

products of the patient: do not rate on the basis of patient's subjective impression of his own level of functioning.

5. Guilt Feelings Over-concern or remorse for past behavior. Rate on the basis of the patient's subjective experiences of guilt as evidenced by verbal report with appropriate affect, do not infer guilt feelings from depression, anxiety or neurotic defenses. 0 1 2 3 4 5 6 NA
6. Tension Physical and motor manifestations of tension, "nervousness," and heightened activation level. Tension should be rated solely on the basis of physical signs and motor behavior and not on the basis of subjective experiences of tension reported by the patient. 0 1 2 3 4 5 6 NA
7. Mannerisms and Posturing Unusual and unnatural behavior which causes certain mental patients to stand out in a crowd of normal people. Rate only abnormality of movements; do not rate simple heightened motor activity here. 0 1 2 3 4 5 6 NA
8. Grandiosity Exaggerated self-opinion, conviction of unusual ability or powers. Rate only on the basis of patient's statements about himself or self-in-relation-to-others, not on the basis of his demeanor in the interview situation. 0 1 2 3 4 5 6 NA

9. Depressive Mood      Despondency in mood, sadness. Rate only degree of despondency; do not rate on the basis of inferences concerning depression based upon general retardation and somatic complaints.      0 1 2 3 4 5 6 NA
10. Hostility      Animosity, contempt, belligerence, disdain for other people outside the interview situation. Rate solely on the basis of the verbal report of feelings and actions of the patient toward others; do not infer hostility from neurotic defenses, anxiety nor somatic complaints. (Rate attitude toward interviewer under "uncooperativeness").      0 1 2 3 4 5 6 NA
11. Suspiciousness      Belief (delusional or otherwise) that others have now, or have had in the past, malicious or discriminatory intent toward the patient. On the basis of verbal report, rate only those suspicions which are currently held whether they concern past or present circumstances.      0 1 2 3 4 5 6 NA
12. Hallucination      Perceptions without normal external stimulus correspondence. Rate only those experiences which are described as distinctly different from the thought and imagery processes of normal people.      0 1 2 3 4 5 6 NA

- |     |                               |   |   |   |   |   |   |   |   |    |
|-----|-------------------------------|---|---|---|---|---|---|---|---|----|
| 13. | Motor<br>Retardation          | Reduction in energy level evidenced in slow movements. Rate on the basis of observed behavior of the patient only; do not rate on the basis of patient's subjective impression of own energy level. | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA |
| 14. | Uncooperative-<br>ness        | Evidence of resistance, unfriendliness, resentment, and lack of readiness to cooperate.   | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA |
| 15. | Unusual<br>Thought<br>Content | Unusual, odd, strange, or bizarre thought content. Rate here the degree of disorganization of thought processes.  | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA |
| 16. | Blunted<br>Affect             | Reduced emotional tone, apparent lack of normal feeling or involvement.   | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA |
| 17. | Excitement                    | Heightened emotional tone, agitation, increased reactivity.   | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA |
| 18. | Disorientation                | Confusion or lack of proper association of person, place or time.   | 0 | 1 | 2 | 3 | 4 | 5 | 6 | NA |

Note: For use in this study, this scale photographically reduced to fit on one page.

Note: This scale is a revised version of that reported by Overall and Gorham (1962).

APPENDIX F  
SYMPTOM CHECKLIST

Mark "Yes" or "No" for whether each of the following symptoms is present for this patient (Case S: Sherry)\*. Please make one check for each item and then circle the degree of certainty that you have for your choice.

<u>Symptom</u>	<u>Yes</u>	<u>No</u>	<u>Degree of Certainty</u>					
			Uncertain			Certain		
			1	2	3	4	5	6
Anhedonia	___	___	1	2	3	4	5	6
Aphasia	___	___	1	2	3	4	5	6
Auditory hallucinations	___	___	1	2	3	4	5	6
Bizarre ideation	___	___	1	2	3	4	5	6
Dysphoric mood	___	___	1	2	3	4	5	6
Euphoric mood	___	___	1	2	3	4	5	6
Feelings of inadequacy	___	___	1	2	3	4	5	6
Flight of ideas	___	___	1	2	3	4	5	6
Inappropriate affect	___	___	1	2	3	4	5	6
Severe anxiety	___	___	1	2	3	4	5	6
Social isolation	___	___	1	2	3	4	5	6
Thought broadcasting	___	___	1	2	3	4	5	6
Thought disorder	___	___	1	2	3	4	5	6
Visual hallucinations	___	___	1	2	3	4	5	6
Word salad	___	___	1	2	3	4	5	6

\*Or Case F: Fred.

APPENDIX G  
DIAGNOSTIC IMPRESSION

On this page, please give your diagnostic impression of this patient (Case S: Sherry)\*. For classification and standardization, please include ICD-9 (DSM-III) category numbers.

Please indicate your degree of certainty in this judgment by circling an appropriate value on the following scale.

Uncertain      1    2    3    4    5    6      Certain

\*Or Case F: Fred.





## Diagnostic Selection: Axis I (Fred)

From the following list of disorders, please select DSM-III Axis I diagnoses for this patient (Case F: Fred) using the following two methods.

- a. Atypical Psychosis
- b. Brief Reactive Psychosis
- c. Depersonalization Disorder
- d. Factitious Disorder with Psychological Symptoms
- e. Major Depressive Episode with Melancholia
- f. Major Depressive Episode with Psychotic Features
- g. Schizoaffective Disorder
- h. Schizophrenic Disorder, Undifferentiated Type
- i. Schizophreniform Disorder
- j. No Axis I disorder

## A. Multiple choice

(Choose one of the disorders listed above) \_\_\_\_\_

Please indicate your degree of uncertainty in this choice by circling an appropriate value on the following scale:

Uncertain    1    2    3    4    5    6    Certain

## B. Rank Order

(List as many of the disorders listed above as you believe are possible diagnoses for this patient. List the primary diagnosis first, followed by your second choice and any others which you believe are appropriate.)

(Primary)	1. _____	4. _____	7. _____
	2. _____	5. _____	8. _____
	3. _____	6. _____	9. _____

Please indicate your degree of certainty in this ranking by circling an appropriate value on the following scale:

Uncertain    1    2    3    4    5    6    Certain

## Diagnostic Selection: Axis II (Sherry)

From the following list of disorders, please select DSM-III Axis II diagnoses for this patient (Case S: Sherry) using the following two methods.

- a. Antisocial Personality Disorder
- b. Atypical Personality Disorder
- c. Avoidant Personality Disorder
- d. Borderline Personality Disorder
- e. Dependent Personality Disorder
- f. Histrionic Personality Disorder
- g. Mixed Personality Disorder
- h. Narcissistic Personality Disorder
- i. Schizotypal Personality Disorder
- j. No Axis II disorder

## A. Multiple choice

(Choose one of the disorders listed above) \_\_\_\_\_

Please indicate your degree of certainty in this choice by circling an appropriate value on the following scale:

Uncertain      1    2    3    4    5    6      Certain

## B. Rank Order

(List as many of the disorders listed above as you believe are possible diagnoses for this patient. List the primary diagnosis first, followed by your second choice and any others which you believe are appropriate.)

(Primary)	1. _____	4. _____	7. _____
	2. _____	5. _____	8. _____
	3. _____	6. _____	9. _____

Please indicate your degree of certainty in this ranking by circling an appropriate value on the following scale:

Uncertain      1    2    3    4    5    6      Certain

## Diagnostic Selection: Axis II (Fred)

From the following list of disorders, please select DSM-III Axis II diagnoses for this patient (Case F: Fred) using the following two methods.

- a. Atypical Personality Disorder
- b. Avoidant Personality Disorder
- c. Borderline Personality Disorder
- d. Compulsive Personality Disorder
- e. Dependent Personality Disorder
- f. Mixed Personality Disorder
- g. Passive-Aggressive Disorder
- h. Schizoid Personality Disorder
- i. Schizotypal Personality Disorder
- j. No Axis II disorder

## A. Multiple choice

(Choose one of the disorders listed above) \_\_\_\_\_

Please indicate your degree of certainty in this choice by circling an appropriate value on the following scale:

Uncertain      1      2      3      4      5      6      Certain

## B. Rank order

(List as many of the disorders listed above as you believe are possible diagnoses for this patient. List the primary diagnosis first, followed by your second choice and any others which you believe are appropriate.)

(Primary)    1. \_\_\_\_\_                      4. \_\_\_\_\_                      7. \_\_\_\_\_  
                   2. \_\_\_\_\_                      5. \_\_\_\_\_                      8. \_\_\_\_\_  
                   3. \_\_\_\_\_                      6. \_\_\_\_\_                      9. \_\_\_\_\_

Please indicate your degree of certainty in this ranking by circling an appropriate value on the following scale:

Uncertain      1      2      3      4      5      Certain

APPENDIX I  
ADDITIONAL INFORMATION

Please select the following pieces of information that you would require for you to be sufficiently certain in making a diagnosis and recommending treatment for this patient (Case S: Sherry)\*. Rank order the items, beginning with "1" for the information that you would request first. In selecting any information, please take into account reasonableness in terms of cost and efficiency, as you would normally do in such a clinical situation. Do not simply request "as much information as possible." Please assume that each piece of information represents a significant expense and that you have not otherwise received it as yet.

Additional Information Needed (rank order only items that you would request)

- None (current information is sufficient) \_\_\_\_\_
- CAT Scan \_\_\_\_\_
- Criminal record \_\_\_\_\_
- Description of subculture, area of residence \_\_\_\_\_
- Dexamethasone Suppression Test (DST) \_\_\_\_\_
- Educational records \_\_\_\_\_
- EEG \_\_\_\_\_
- Family history report \_\_\_\_\_
- Food intake, weight change \_\_\_\_\_
- History of drug use \_\_\_\_\_
- History of psychosexual development \_\_\_\_\_
- Lab results (routine admission tests) \_\_\_\_\_
- Medical records \_\_\_\_\_
- Neurological examination \_\_\_\_\_
- Neuropsychological assessment \_\_\_\_\_
- Nursing report of behavior since admission\*\* \_\_\_\_\_

\*Or Case F: Fred.

Occupational history	_____
Personal interview with family	_____
Physical examination	_____
Psychological report**	_____
Quantitative assessment of problem behaviors	_____
Religion	_____
Sleep record	_____

\*\*Nursing report and psychological report items are omitted when the scale is administered at the second session. The subjects will have just received those reports.

APPENDIX J  
RECALL NOTES

In approximately two weeks, you will receive additional information about this patient (Case S: Sherry)\*. You will also be given this personal reminder sheet. Please write whatever you would like in the space below to help you recall the information given thus far and your current impression of this patient. Feel free to use whatever format you prefer and to use the back of this sheet if necessary.

\*Or Case F: Fred.



APPENDIX L  
STATEMENT OF INFORMED CONSENT

The purpose of this study is to examine how clinicians use information in making psychiatric assessments. Structured and unstructured measures of diagnostic and symptom evaluation will be generated from your participation.

Materials will be given to you for two sessions, approximately two weeks apart. You will be able to complete the measures at your convenience. Each session is designed to require approximately 30 to 45 minutes of your time. For the first session you will need a cassette tapeplayer to listen to a brief audiotape. Descriptions of two hypothetical patients will be presented. You will be asked to make clinical judgments based upon this information.

There is no anticipated risk in participation. This study is not an assessment of clinical aptitude, as there are no established right or wrong answers to the questions. There will be no compensation for your participation. Confidentiality will exist to the extent provided by law. Participants will be assigned random code numbers to keep identity confidential.

The potential benefits of the study are: 1) that you may profit from thoughtful attention to different sources of clinical judgments; and 2) that the study may yield insights into how different types of clinical information are used and to how specific judgments of symptoms related to formal diagnostic judgments.

I have fully explained to \_\_\_\_\_ the nature and purpose of the above described procedure and the risks that are involved in its performance. I have answered and will answer all questions to the best of my ability.

\_\_\_\_\_  
Signature of person obtaining consent

I have been fully informed of the above described procedure with its possible benefits and risks and have received a copy of this description. I have given my permission for my participation in this study.



I understand that I am free to withdraw this consent and discontinue participation in this project at any time without prejudice.

\_\_\_\_\_  
Signature of subject

\_\_\_\_\_  
Witness to signature

\_\_\_\_\_  
Date

\_\_\_\_\_  
Principal Investigator

\_\_\_\_\_  
Date

## REFERENCES

- Adams, J. (1972). The contribution of the psychological evaluation to psychiatric diagnosis. Journal of Personality Assessment, 36, 561-566.
- Adams, J. K., & Adams, P. A. (1961). Realism of confidence judgments. Psychological Review, 68, 33-45.
- Affleck, D. C., & Strider, F. D. (1971). Contributions of psychological reports to patient management. Journal of Consulting and Clinical Psychology, 37, 177-179.
- Allison, J., Blatt, S., & Zimet, C. (1968). The interpretation of psychological tests. New York: Harper & Row.
- Anderson, N. H. (1965). Averaging versus adding as a stimulus-combination rule in impression formation. Journal of Experimental Psychology, 70, 394-400.
- Anderson, N. H. (1974). Cognitive algebra: Integration theory applied to social attribution. In L. Berkowitz (Ed.), Advances in experimental social psychology (Vol. 7). New York: Academic Press.
- Appelbaum, S. A. (1970). Science and persuasion in the psychological test report. Journal of Consulting and Clinical Psychology, 35, 349-355.
- Appelbaum, S. A. (1976). Objections to diagnosis and diagnostic testing diagnosed. Bulletin of the Menninger Clinic, 40, 559-564.
- Appelbaum, S. A. (1977). The anatomy of change. New York: Plenum.
- Arkes, H. R. (1981). Impediments to accurate clinical judgment and possible ways to minimize their impact. Journal of Consulting and Clinical Psychology, 49, 323-330.

- Arthur, A. Z. (1969). Diagnostic testing and the new alternatives. Psychological Bulletin, 72, 183-192.
- Ausubel, D. P. (1956). Relationships between psychology and psychiatry: The hidden issues. American Psychologist, 11, 99-105.
- Bellak, L. (1959). Psychological test reporting: A problem in communication between psychologists and psychiatrists. Journal of Nervous and Mental Disease, 129, 76-77.
- Beutler, L. E. (1973). The therapy dyad: Yet another look at diagnostic assessment. Journal of Personality Assessment, 37, 303-308.
- Bieri, J., Atkins, A. L., Briar, S., Leaman, R. L., Miller, H., & Tripoli, T. (1966). Clinical and social judgment: The discrimination of behavioral information. New York: Wiley.
- Blaser, A. (1978). Components of the clinician's confidence. Perceptual and Motor Skills, 47, 1275-1280.
- Blashfield, R. K., & Sprock, J. (1983). A replication study on the use of prototypes for understanding psychiatric classification. Unpublished manuscript, University of Florida, Gainesville.
- Blatt, S. J. (1975). The validity of projective techniques and their research and clinical contribution. Journal of Personality Assessment, 39, 327-343.
- Breger, L. (1968). Psychological testing: Treatment and research implications. Journal of Consulting and Clinical Psychology, 32, 176-182.
- Brodsky, S. L. (1972). Shared results and open files with the client. Professional Psychology, 3, 362-364.
- Brody, E. B. (1956). Interprofessional relations or psychologists and psychiatrists are human too, only more so. American Psychologist, 11, 105-111.
- Brody, E. G. (1959). The psychiatrist's decision to refer a patient for psychological testing. Journal of Nervous and Mental Disease, 129, 88-91.
- Brown, J. F., & Rapaport, D. (1941). The role of the psychologist in the psychiatric clinic. Bulletin of the Menninger Clinic, 5, 75-84.

- Bruner, J. S., Goodnow, J. J., & Austin, G. A. (1956). A study of thinking. New York: Wiley.
- Brunswik, E. (1952). The conceptual framework of psychology. Chicago: University of Chicago Press.
- Brunswik, E. (1956). Perception and the representative design of experiments. Berkeley: University of California Press.
- Cantor, N., Smith, E. E., de Sales, R., & Mezzich, J. (1980). Psychiatric diagnosis as prototype classification. Journal of Abnormal Psychology, 89, 181-193.
- Cantril, H., Ames, A., Jr., Hastorf, A. H., & Ittelson, W. H. (1949). Psychology and scientific research. I. The values of scientific inquiry and scientific method. Science, 110, 461-464.
- Carr, A. C. (1980). Psychological testing of personality. In H. I. Kaplan, A. M. Freidman & B. J. Sadock (Eds.), Comprehensive textbook of psychiatry/III (3rd ed.) (Vol. 1). Baltimore: Williams & Wilkins.
- Carson, R. C. (1958) The status of diagnostic testing. American Psychologist, 13, 79.
- Cassirer, E. (1944). An essay on man: An introduction to a philosophy of human culture. New Haven: Yale University Press.
- Christensen-Szalanski, J. J. J., & Beach, L. R. (1984). The citation bias: Fad and fashion in the judgment and decision literature. American Psychologist, 39, 75-78.
- Clavelle, P. R., & Turner, A. D. (1980). Clinical decision-making among professionals and paraprofessionals. Journal of Clinical Psychology, 36, 833-838.
- Cleveland, S. E. (1976). Reflections on the rise and fall of psychodiagnosis. Professional Psychology, 7, 309-318.
- Cole, J. K., & Magnussen, M. G. (1966). Where the action is, Journal of Consulting Psychology, 30, 539-543.
- Craddick, R. A. (1972). Humanistic assessment: A reply to Brown. Psychotherapy: Theory, Research and Practice, 9, 107-110.

- Craddick, R. A. (1975). Sharing oneself in the assessment procedure. Professional Psychology, 6, 279-282.
- Crary, W. G., & Steger, H. G. (1972). Prescriptive and consultative approaches to psychological evaluation. Professional Psychology, 3, 105-110.
- Crenshaw, D. A., Bohn, S., Hoffman, M. R., Matheus, J. M., & Offenbach, S. G. (1968). The use of projective methods in research; 1947-1965. Journal of Projective Techniques & Personality Assessment, 32, 3-9.
- Cuadra, C. A., & Albaugh, W. P. (1956). Sources of ambiguity in psychological reports. Journal of Clinical Psychology, 12, 109-115.
- Dailey, C. A. (1953). The practical utility of the clinical report. Journal of Consulting Psychology, 17, 297-302.
- Diethelm, O., & Knehr, C. A. (1951). The diagnostic use of psychological tests from the psychiatrist's standpoint. In P. H. Hoch & J. Zubin (Eds.), Relation of psychological tests to psychiatry. New York: Grune & Stratton.
- Einhorn, H. J. (1980). Overconfidence in judgment. New directions for methodology of social and behavioral science, 4, 1-16.
- Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. Psychological Review, 85, 395-416.
- Einhorn, H. J., & Hogarth, R. M. (1982). A theory of diagnostic inferences: I. Imagination and the psychophysics of evidence. Unpublished manuscript, University of Chicago, Graduate School of Business, Center for Decision Research.
- Einhorn, H. J., & Hogarth, R. M. (1983). Diagnostic inference and casual judgment: A decision making framework. Unpublished manuscript, University of Chicago, Graduate School of Business, Center for Decision Research.
- Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. Psychological Review, 86, 465-485.

- Eker, D. (1981) Diagnosis and conceptions about mental disorders among Turkish and American clinicians (Doctoral dissertation, University of Florida, 1981). Dissertation Abstracts International, 42, 2052B.
- Erikson, E. (1964). The nature of clinical evidence. In E. Erikson, Insight and responsibility. New York: Norton.
- Fischer, C. T. (1970). The testee as co-evaluator. Journal of Counseling Psychology, 17, 70-76.
- Fischer, C. T. (1972). Paradigm changes which allow sharing of the results. Professional Psychology, 3, 364-365.
- Freedheim, D. K. (1972). Beyond diagnosis and intervention. Professional Psychology, 3, 104.
- Fulkerson, S. C. (1965). Some implications of the new cognitive theory for projective tests. Journal of Consulting Psychology, 29, 191-197.
- Garfield, S. L., Heine, R. W., & Leventhal, M. (1954). An evaluation of psychological reports in a clinical setting. Journal of Consulting Psychology, 18, 281-286.
- Garfield, S. L., & Kurtz, R. M. (1974). A survey of clinical psychologists: Characteristics, activities, and orientations. The Clinical Psychologist, 28, 7-10.
- Gauron, E. F., & Dickinson, J. K. (1966). Diagnostic decision making in psychiatry: 1. Information usage. Archives of General Psychiatry, 14, 225-232.
- Gauron, E. F., & Dickinson, J. K. (1969). The influence of seeking the patient first on diagnostic decision making in psychiatry. American Journal of Psychiatry, 126, 199-205.
- Gough, H. G. (1971). Some reflections on the meaning of psychodiagnosis. American Psychologist, 26, 160-167.
- Green, D. M., & Swets, J. A. (1966). Signal detection theory and psychophysics. New York: Wiley.

- Griffith, B. C., Small, H. G., Stonehill, J. A., & Dye, S. (1974). The structure of scientific literatures. 2. Toward a macro-microstructure for science. Science Studies, 4, 339-365.
- Hammond, K. R., Hursch, C., & Todd, F. J. (1964). Analyzing the components of clinical inference. Psychological Review, 71, 438-456.
- Hammond, K. R., Stewart, T. R., Brehmer, B., & Steinmann, D. O. (1975). Social judgment theory. In M. F. Kaplan and S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press.
- Harlow, R. G., & Salzman, L. F. (1958). Toward the effective use of the psychological consultation. American Journal of Psychiatry, 115, 228-231.
- Hartlage, L. C., Freeman, W., Horine, L., & Walton, C. (1968). Decisional utility of psychological reports. Journal of Clinical Psychology, 24, 481-483.
- Hartlage, L. C., & Merck, K. H. (1971). Increasing the relevance of psychological reports. Journal of Clinical Psychology, 27, 459-460.
- Hertz, M. R. (1970). Projective techniques in crisis. Journal of Projective Techniques & Personality Assessment, 34, 449-467.
- Hinkle, J. E., Nelson, S. E., & Miller, D. (1968). Psychological test usage by psychologist psychotherapists in private practice. Psychotherapy: Theory, Research, and Practice, 5, 210-213.
- Hoffman, P. J. (1960). The paramorphic representation of clinical judgment. Psychological Bulletin, 47, 116-131.
- Hogarth, R. M. (1980). Judgment and choice: The psychology of decision. Chicester, England: Wiley.
- Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. Psychological Bulletin, 90, 197-217.
- Hogarth, R. M., & Makridakis, S. (1981). Forecasting and planning: An evaluation. Management Science, 27, 115-138.

- Holt, R. R. (1967). Diagnostic testing: Present situation and future prospects. Journal of Nervous and Mental Disease, 144, 444-465.
- Holt, R. R. (1970). Yet another look at clinical and statistical prediction: Or, is clinical psychology worthwhile? American Psychologist, 25, 337-349.
- Howe, H. E., Jr. (1981). Description and application of an evaluation scheme for assessment from a decision making perspective. Journal of Clinical Psychology, 37, 110-117.
- Howes, R. J. (1981). The Rorschach: Does it have a future? Journal of Personality Assessment, 45, 339-351.
- Ivnik, R. J. (1977). Uncertain status of psychological tests in clinical psychology. Professional Psychology, 8, 206-213.
- Jenkins, J. J. (1974). Remember that old theory of memory? Well, forget it! American Psychologist, 29, 785-795.
- Johnson, D. M. (1972). A systematic introduction to the psychology of thinking. New York: Harper and Row.
- Kahn, E. (1965). Psychology and psychiatry: A book review and some general considerations. American Journal of Psychiatry, 121, 1023-1025.
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. Psychological Review, 80, 237-251.
- Kaplan, M. F. (1975). Information integration in social judgment: Interaction of judge and informational components. In M. F. Kaplan and S. Schwartz (Eds.), Human judgment and decision processes. New York: Academic Press.
- Kendall, R. E. (1973). Psychiatric diagnoses: A study of how they are made. British Journal Psychiatry, 122, 437-445.
- Kleinmuntz, B. (1968). The processing of clinical information by man and machine. In B. Kleinmuntz (Ed.), Formal representation of human judgment. New York: Wiley.
- Klopfer, W. G. (1959). The psychological report as a problem in interdisciplinary communication. Journal of Nervous and Mental Disease, 129, 86-88.



- Klopper, W. G. (1962). The role of diagnostic evaluation in clinical psychology. Journal of Projective Techniques, 26, 295-298.
- Klopper, W. G. (1964). The blind leading the blind: Psychotherapy without assessment. Journal of Projective Techniques & Personality Assessment, 28, 387-392.
- Klopper, W. G. (1968). The metamorphosis of projective methods Journal of Projective Techniques & Personality Assessment, 32, 402-404.
- Korchin, S. J., & Schuldberg, D. (1981). The future of clinical assessment. American Psychologist, 36, 1147-1158.
- Kubis, J. F. (1951). Part two: Discussion II. In P. H. Hoch & J. Zubin (Eds.), Relation of psychological tests to psychiatry. New York: Grune & Stratton.
- Lacey, H. M., & Ross, A. O. (1964). Multidisciplinary views on psychological reports in child guidance clinics. Journal of Clinical Psychology, 20, 522-526.
- Lakin, M., & Lieberman, M. A. (1965). Diagnostic information and psychotherapeutic conceptualization. Journal of Clinical Psychology, 21, 385-388.
- Lanyon, R. I. (1972). Technological approach to the improvement of decision making in mental health services. Journal of Consulting & Clinical Psychology, 39, 43-48.
- Leach, C. (1974). The importance of instructions in assessing sequential effects in impression formation. British Journal of Social & Clinical Psychology, 13, 151-156.
- Levy, M. R., & Fox, H. M. (1975). Psychological testing is alive and well. Professional Psychology, 6, 420-424.
- Lewandowski, D. G., & Saccuzzo, D. P. (1976). The decline of psychological testing. Professional Psychology, 7, 177-184.
- Lubin, B., Wallis, R. R., & Paine, C. (1971). Patterns of psychological test usage in the United States: 1935-1969. Professional Psychology, 2, 70-74.

- Luchins, A. S. (1957). Primacy-recency in impression formation. In C. I. Hovland (Ed.), The order of presentation in persuasion. New Haven: Yale University Press.
- Luchins, A. S. (1958). Definitiveness of impression and primacy-recency in communications. Journal of Social Psychology, 48, 275-290.
- MacRae, D. (1969). Growth and decay curves in scientific citations. American Sociological Review, 34, 631-635.
- Malpass, R. S., & Devine, P. G. (1981). Eyewitness identification: Lineup instructions and the absence of the offender. Journal of Applied Psychology, 66, 482-489.
- Meehl, P. E. (1954). Clinical versus statistical prediction. Minneapolis: University of Minnesota Press.
- Meehl, P. E. (1977). Some ruminations on the validation of clinical procedures. In P. E. Meehl, Psychodiagnosis: Selected papers. New York: W. W. Norton. (Reprinted from Canadian Journal of Psychology, 1959, 13).
- Millon, T. (1982). Millon clinical multiaxial inventory. Minneapolis: National Computer Systems.
- Mintz, J. (1968). Survey of student therapists' attitudes toward psychodiagnostic reports. Journal of Consulting & Clinical Psychology, 32, 500.
- Moore, C. H., Boblitt, W. E., & Widman, R. W. (1968). Psychiatric impressions of psychological reports. Journal of Clinical Psychology, 24, 373-376.
- Mosak, H. H., & Gushurst, R. S. (1972). Some therapeutic uses of psychological testing. American Journal of Psychotherapy, 26, 539-546.
- Neuringer, C. (1967). Art and science in projective techniques. Journal of Consulting Psychology, 31, 85-88.
- Newell, A. (1968). Judgment and its representation: An introduction. In B. Kleinmuntz (Ed.), Formal representation in human judgment. New York: Wiley.

- Olive, H. (1972). Psychoanalysts' opinions of psychologists' reports: 1952 and 1970. Journal of Clinical Psychology, 28, 50-54.
- Oskamp, S. O. (1965). Overconfidence in case-study judgments. Journal of Consulting Psychology, 29, 261-265.
- Overall, J. E., & Gorham, D. R. (1962). The brief psychiatric rating scale. Psychological Reports, 10, 799-812.
- Oxford English Dictionary (Compact ed.). (1971). Oxford: Oxford University Press.
- Paul, G. L. (1967). Strategy of outcome research in psychotherapy. Journal of Consulting Psychology, 31, 109-119.
- Payne, R. W. (1958). Diagnostic and personality testing in clinical psychology. American Journal of Psychiatry, 115, 25-29.
- Petzelt, J. T., & Craddick, R. (1978). Present meaning of assessment in psychology. Professional Psychology, 9, 587-591.
- Piotrowski, C., & Keller, J. W. (1978). Psychological test usage in Southeastern outpatient mental health facilities in 1975. Professional Psychology, 9, 63-67.
- Polanyi, M. (1962). Personal knowledge. Chicago: University of Chicago Press.
- Price, D. J. D. (1963). Little science, big science. New York: Columbia University Press.
- Rapoport, L., & Summers, D. (1973). Origins, aims and scope: A general introduction. In L. Rappoport & D. Summers (Eds.), Human judgment and social interaction. New York: Holt, Rinehart and Winston.
- Reynolds, W. M. (1979). Psychological test: Clinical usage versus psychometric quality. Professional Psychology, 10, 324-329.
- Rhodes, W. C. (1974). Principles and practices of consultation. Professional Psychology, 5, 287-292.

- Richman, J. (1967). Reporting diagnostic test results to patients and their families. Journal of Projective Techniques & Personality Assessment, 31, 62-70.
- Rosenwald, G. S. (1963). Psychodiagnostics and its discontents. Psychiatry, 26, 222-240.
- Ryback, D. (1967). Confidence and accuracy as a function of experience in judgment-making in the absence of systematic feedback. Perceptual and Motor Skills, 24, 331-334.
- Sandifer, M. G., Hordun, A., & Green, L. M. (1970). The psychiatric interview: The impact of the first three minutes. American Journal of Psychiatry, 126, 968-973.
- Sargent, H. (1951). Psychological test reporting: An experiment in communication. Bulletin of the Menninger Clinic, 15, 175-186.
- Schafer, R. (1954). Psychoanalytic interpretation in Rorschach testing. New York: Grune & Stratton.
- Schinka, J. (1976). Confidence in personality assessment. Psychological Reports, 38, 62.
- Schneider, D. J. (1973). Implicit personality theory, Psychological Bulletin, 79, 294-309.
- Schwartz, F., & Lazar, J. (1979). The scientific status of the Rorschach. Journal of Personality Assessment, 43, 3-11.
- Shakow, D. (1949). Psychology and psychiatry: A dialogue. American Journal of Orthopsychiatry, 19, 191-208, 381-396.
- Shectman, F., & Harty, M. K. (1982). Mental health disciplines in conflict: The patient pays the price. Bulletin of the Menninger Clinic, 46, 458-464.
- Shemberg, K., & Keeley, S. (1970). Psychodiagnostic training in the academic setting: Past and present. Journal of Consulting and Clinical Psychology, 34, 205-211.

- Sloves, R. E., Docherty, E. M., Jr., Schneider, K. C. (1979). A scientific problem-solving model of psychological assessment. Professional Psychology, 10, 28-35.
- Slovic, P. Cue consistency and cue utilization in judgment. American Journal of Psychology, 1966, 79, 427-434.
- Slovic, P., & Lichtenstein, S. (1973). Comparison of Bayesian and regression approaches to the study of information processing in judgment. In L. Rappoport & D. Summers, (Eds.), Human judgment and social interaction. New York: Holt, Rinehart & Winston.
- Small, H. G., & Griffith, B. C. (1974). The structure of scientific literatures. 1. Identifying and graphing specialties. Science Studies, 4, 17-40.
- Smyth, R., & Reznikoff, M. (1971). Attitudes of psychiatrists toward the usefulness of psychodiagnostic reports. Professional Psychology, 2, 283-288.
- Snyder, S., Pitts, W. M., Goodpaster, W.A., Sajadi, C., & Gustin, Q. (1982). MMPI profile of DSM-III borderline personality. American Journal of Psychiatry, 139, 1046-1048.
- Sterling, M. E. (1982). Must psychology lose its soul? Professional Psychology, 13, 789-796.
- Sundberg, N. D. (1961). The practice of psychological testing in clinical services in the United States. American Psychologist, 16, 79-83.
- Swensen, C. H. (1974). The process of psychological consultation: Introduction. Professional Psychology, 5, 286-187.
- Tallent, N. (1956). An approach to the improvement of clinical psychological reports. Journal of Clinical Psychology, 12, 103-109.
- Tallent, N. (1965). Clinical psychological testing: A review of premises, practices and promises. Journal of Projective Techniques & Personality Assessment, 29, 418-435.

- Tallent, N. (1983). Psychological report writing (2nd ed.). Englewood Cliffs, N.J.: Prentice-Hall.
- Tallent, N., & Reiss, W. J. (1959a). Multidisciplinary views on the preparation of written clinical psychological reports: I. Spontaneous suggestions for content. Journal of Clinical Psychology, 15, 218-221.
- Tallent, N., & Reiss, W. J. (1959b). Multidisciplinary views on the preparation of written clinical psychological reports: II. Acceptability of certain common content variables and styles of expression. Journal of Clinical Psychology, 15, 273-274.
- Tallent, N., & Reiss, W. J. (1959c). Multidisciplinary views on the preparation of written clinical psychological reports: III. The trouble with psychological reports. Journal of Clinical Psychology, 15, 444-446.
- Thelen, M. H., & Ewing, D. R. (1970). Roles, functions, and training in clinical psychology: A survey of academic clinicians. American Psychologist, 25, 550-554.
- Thelen, M. H., Varble, D. L., & Johnson, J. (1968). Attitudes of academic clinical psychologists toward projective techniques. American Psychologist, 23, 517-521.
- Tolor, A. (1973). Diagnosing the state of the diagnostic function: An analysis of the literature. Journal of Clinical Psychology, 29, 338-342.
- Towbin, A. P. (1964). Psychological testing from end to means. Journal of Projective Techniques & Personality Assessment, 28, 86-91.
- Valentine, C. W. (1929). The relative reliability of men and women in intuitive judgments of character. British Journal of Psychology, 19, 213-228.
- Wade, T. C., & Baker, T. B. (1977). Opinions and use of psychological tests: A survey of clinical psychologists. American Psychologist, 32, 874-882.
- Watson, R. I. (1953). A brief history of clinical psychology. Psychological Bulletin, 50, 321-346.

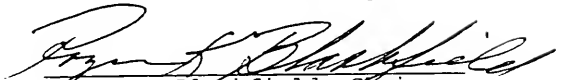
- Webster's third international dictionary. (1976).  
Springfield, Mass.: G. & C. Merriam.
- Wegner, D. M., & Vallacher, R. R. (1977). Implicit psychology: An introduction to social cognition.  
New York: Oxford University Press.
- Weiner, I. D. (1972). Does psychodiagnosis have a future? Journal of Personality Assessment, 36, 534-546.
- Weissman, M. S. (1980). Decisiveness. In R. H. Woody (Ed.), Encyclopedia of clinical assessment (Vol. I).  
San Francisco: Jossey-Bass.
- Wiedemann, C., & Mintz, J. (1974). Student therapists' assessment of diagnostic testing. Journal of Personality Assessment, 38, 203-214.
- Wiggins, N. (1973). Individual differences in human judgments: A multivariate approach. In L. Rappoport & D. A. Summers (Eds.), Human judgment and social interaction. New York: Holt, Rinehart & Winston.
- Zubin, J. (1951). Part two: Discussion I. In P. H. Hoch & J. Zubin (Eds.), Relation of psychological tests to psychiatry. New York: Grune & Stratton.

## BIOGRAPHICAL SKETCH

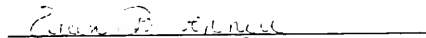
James Martin Siwy was born on April 19, 1951, in Chicago, Illinois. He spent his childhood and adolescence in Glenview, Illinois, where he graduated from Glenbrook South High School in June, 1969. He later attended Harvard College in Cambridge, Massachusetts, graduating with the degree of Bachelor of Arts cum laude in psychology and social relations, in March, 1977. He enrolled in the Department of Clinical Psychology of the University of Florida, where he received the Master of Arts degree in December, 1979. In 1981-82, he served his predoctoral internship at Hall-Brooke Hospital in Westport, Connecticut. Upon receiving the Doctor of Philosophy degree, he plans to enter into professional practice of clinical psychology.




I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
Roger K. Blashfield, Chairman  
Associate Professor of  
Clinical Psychology

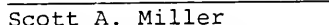
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
Eileen B. Fennell  
Associate Professor of  
Clinical Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
Mary H. McCaulley  
Assistant Professor of  
Clinical Psychology

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

  
Scott A. Miller  
Associate Professor of  
Psychology

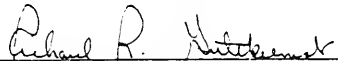
I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

---

Austin B. Creel  
Professor of Religion

This dissertation was submitted to the Graduate Faculty of the College of Health Related Professions and to the Graduate School, and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1984

---

Dean, College of Health  
Related Professions

---

Dean for Graduate Studies and  
Research

UNIVERSITY OF FLORIDA



3 1262 08285 243 4