PSYCHODIAGNOSIS:
ON THE WORN FABRIC OF AN OLD PARADIGM

CHRI$$ HIGGI$$SON

Submitted in partial fulfilment of the requirements of the degree of Master of Science of the Australian National University, Canberra.

March, 1980
To Jane and Prue
I sincerely thank the following people who have helped me with the preparation of the thesis: Jan Bennett, John Fraser, David O'Sullivan, Liz McMahon, Lorna Myers, Sandi Plummer, Roger Rees, Gareth Smith and Thelma Wasiliew.

Particular thanks: to Gareth who has 'fed my head' for the past five years;

to Alan Jones who demonstrated that the ideals of the 'scientist-practitioner' can be attained; and

to Sally, without whose support and love the thesis would not have been completed.
DECLARATION

I hereby certify that this thesis, submitted in candidature for the degree of Master of Science from the Australian National University Canberra, has not been submitted in substance for any degree and is not being currently submitted for any other degree.

The work is the result of my own enterprise except where otherwise stated.

Candidate: [Signature]
Over the past several years I have been involved in teaching in a unit on the psychological assessment of children and youth to postgraduate students of counselling. This experience has highlighted what was for me a stark and inexplicable schism between research and practice: Namely, that practitioners (in this case school counsellors) persist in using psychological assessment techniques and instruments for which negative results concerning reliability, validity and utility are in the ascendency. Moreover, not only do they use tests of dubious validity but express unshakeable confidence in the resultant clinical judgement based on test data.

In a sense, I was grateful to find that the difference between my perceptions as to the value of psychodiagnostic tests and those of my practitioner colleagues was neither an atypical nor endemic phenomenon. Recent evidence suggests that the popularity of psychological tests, including projective techniques, has persisted in everyday clinical practice. Most clinicians, undaunted by the data, apparently feel comfortable basing their judgements and predictions on inferences shown experimentally to be illusions. Surprisingly, most practitioners know about the negative research, yet themselves continue to use the devices regularly, assumedly because they have seen the signs "work" in their own clinical work. Similarly, even when psychologists are cognizant of the vast literature attending to the fallibility of selection/mental status interviews, it is commonly observed that they not only conduct such
interviews but express considerable confidence in their predictions. There is a puzzling and substantial discrepancy existing between the findings of research on psychological assessment and judgemental fallibility and the clinicians' stubborn confidence in their own diagnostic acumen.

As a result of teaching in the assessment unit, and the concomitant discussion, debate and reading, I have begun to identify some of the possible factors contributing to this schism and the persistence of the illusion of validity. But despite the occasional insight, the intractability of the situation remains daunting. For although I could claim some success in convincing an occasional student of the telling nature of the research data, I have been singularly unsuccessful in 'developing' a practitioner capable of not succumbing to the temptation of the 'offending' assessment procedures and the resultant illusions of validity.

Although the overarching aim of the counselling program is to develop competent 'scientific practitioners', my experience has been that it is inordinately difficult to teach trainee counsellors to harness the power of science in a manner that would promote reliable and rational psychodiagnostic procedures; that is, to develop counsellors who incorporate ways of behaving that maximise the possibility of uncovering invalid and ineffectual approaches to human problems and who in an ongoing, everyday way refine the practice of psychodiagnosis by exposing themselves to 'good', 'hard' data. In contrast unfortunately it appears that most practitioners unwittingly behave in ways - such as relying on subjective impressions
accrued through personal experience with tests to guide usage - which maximise their vulnerability to fall prey to such phenomena as illusory correlates and the 'Barnum effect'; resulting in the widespread presence of 'magical thinking' in psychodiagnosis.

The aim of this essay is to thoroughly explore the antecedents of the research/practice dislocation - in order to uncover some of the plausible reasons for the persistence of what appear to be patently untenable practices and the paradoxical but equally untenable, illusion of validity. Hopefully, in future years my teaching efficacy may be augmented by the inclusion of the recommendations that flow from the critical analysis of the traditional 'attribution' approach to psychodiagnosis.
CONTENTS

ACKNOWLEDGEMENTS ................................................................. i.
DECLARATION ........................................................................ ii.
PREFACE ................................................................................ iii.

CHAPTER I: INTRODUCTION

INTRODUCTION .......................................................................... 1
The Research/Practice Schism
Old Paradigms Never Die - But They Should Fade Away
On the Rise and Fall of Psychodiagnostic Assessment
Psychological Assessment as Alchemy
Stereotyping, Clinical Fallibility and the Attribute Model
A Prolegomenon for Interactional Assessment

CHAPTER II: ON THE DECLINE AND FALL OF PSYCHODIAGNOSTIC ASSESSMENT

INTRODUCTION .......................................................................... 24
FACTORS LEADING TO THE DECLINE OF PSYCHODIAGNOSTIC ASSESSMENT.. 26
1. Negative Research
2. Invocation of Psychopathology
3. Psychiatric Classification
4. Treatment Irrelevance
5. Dehumanising
6. Organismic Error
7. Poor Media Coverage
8. Academic Training
9. Aging Assessment Instruments
10. Cost

PSYCHODIAGNOSIS IS ALIVE AND WELL .......................... 40
(a) Academic Prejudice Against Testing
(b) Time Devoted to Psychodiagnosis by Clinicians
(c) On the Decline of Projective Techniques
CHAPTER III: ON THE RISE OF PSYCHODIAGNOSTIC ASSESSMENT

INTRODUCTION .......................................................... 48
PSYCHODIAGNOSIS AND THE ROLE OF CLINICAL PSYCHOLOGIST ...... 50
Clinical Psychologists as Pseudopsychiatrists
The Psychological Alchemy of Assessment
PSYCHODIAGNOSIS AND MYTH MAKING .................................. 55
Psychological Myth Making
PSYCHODIAGNOSIS AND PSEUDOSCIENCE ............................. 62
What is Science?
What is Pseudoscience?

CHAPTER IV: PSYCHOLOGICAL ALCHEMY - I: PSYCHOMETRIC PSEUDOSCIENCE

INTRODUCTION .......................................................... 74
PSYCHOMETRIC PSEUDOSCIENCE ........................................ 77
Intelligence Testing
The Transformation of Binet's Ideas
Innate Differences in Intelligence
Who Defines?
Intelligence Tests and Social Sanitation
Defining Intelligence
The Validity of the General Intelligence Concept
   (a) Tests and Grades in School
   (b) Test and Occupational Success
SCIENTIFIC PARADIGMS AND SOCIAL VALUES ......................... 101
Within or Between Paradigm Debate

CHAPTER V: PSYCHOLOGICAL ALCHEMY - II: FREUDIAN PSEUDOSCIENCE

INTRODUCTION .......................................................... 105
THE IMPACT AND STATUS OF FREUD'S IDEAS ............................. 108
Obstacles to the Accurate Evaluation of Freud's Contribution
THE PSYCHOANALYTIC MOVEMENT .................................... 113
FREUD'S PSEUDOSCIENCE ................................................. 119
Freud's Assumption to Scientific Status
Psychoanalysis Ought to Become Scientific
PHILOSOPHICAL CONSIDERATIONS ............................................. 127
  Freud's Theories - Progressive and Degenerating Program
  Prediction and Freud's Theories
EMPIRICAL VALIDATION .......................................................... 133
  Charges of Deliberate Scientific Pretence .............................. 137
    The Sources of Freud's Ideas - Observation and Literature
    Free Association - The Methodological Key
ALTERNATIVE EPISTEMOLOGIES ............................................... 145

CHAPTER VI: THE ATTRIBUTE MODEL OF PSYCHODIAGNOSIS

INTRODUCTION ........................................................................... 149

THE ASSumptIVE BASE OF THE ATTRIBUTE MODEL ....................... 151
  Central Assumptions
  Test Responses as Signs or Samples
  The Language of Assessment
  Qualitative Differences Between People
  Reliability and Validity

DISPOSITIONAL EXPLANATIONS OF BEHAVIOUR ............................ 159
  Mischel's Critique of the Attribute Model
  Empirical Evidence for the Consistency Position
  The Categorization of People

THE QUASIMEDICAL MODEL ..................................................... 168
  The Clinical Definition of Normality
    (a) The Pathological Model of Normality
    (b) The Statistical Model of Normality
  The Transformation of Behavioural Differences into Signs of Pathology

FROM SINNER TO SICK DEVIANT .............................................. 175
  The Social Impact of a New Label
  The 'Insane' as Patient or Victim
  The Social Construction of Mental Illness
  The Latent Function of Psychiatric Help
CHAPTER VII: STEREOTYPES AND PSYCHODIAGNOSIS

INTRODUCTION ................................................. 185
THE POLITICS OF PSYCHODIAGNOSTIC LABELLING ............... 188
The Influence of Task Irrelevant Client Characteristics in Psychodiagnosis
IMPRESSION FORMATION AND PSYCHODIAGNOSIS ................. 191
Stereotypes and Implicit Personality Theories
Professional Training and Clinical Judgement
RELIABILITY OF PSYCHIATRIC DIAGNOSIS ..................... 195
Change of Psychiatric Diagnosis in Time and Space
The Influence of Examiner Characteristics
Suggestion and Psychodiagnosis
Diagnostic Decision Making
The Diagnostic Utility of Different Kinds of Data
Barnum Statements in Psychodiagnosis

CHAPTER VIII: THE FALLIBLE PSYCHODIAGNOSTICIAN

INTRODUCTION .................................................. 211
PITFALLS OF INFORMATION IN PSYCHOLOGICAL ASSESSMENT ...... 213
Magical Thinking and the Disinclination to Think Probabilistically
Inferring Causality in Person Perception
Correlation as a 'Non-Intuitive' Concept
Causality and Correlation and Clinical Diagnosis
Probability Learning
Resemblance in Reasoning
CONFIRMATORY BIAS ........................................... 221
Disconfirmation and the Logic of Lay Persons
Logical Reasoning by Scientists
The Clinical Judgemental Task and Disconfirmatory Evidence
Additional Analogue Demonstrations of our Logical Fallback
On the Strong Proclivity of all Humans to Favour Confirmatory Evidence
Confirmation and Self-Fulfilling Prophecies
CHAPTER IX: CONCLUSION

INTRODUCTION ........................................ 246

Influences on Clinician's Testing Practices

WITHIN PARADIGM RECOMMENDATIONS ................. 250

Improving the Clinician's Ability to Learn from Experience
Sensitising the Clinicians to the Importance of Their Own
Cognitive Structures
Scientific Knowledge and Social Values
Perceiver and Perceived Relationship
Giving up the Luxury of Labelling

BEHAVIOURAL ASSESSMENT ................................ 260

The Nature of Behavioural Assessment
Central Aspects of Behaviour Assessment
Rejection of traditional diagnostic systems
Behavioural approach to abnormality
Identification of controlling variables
Situational specificity and temporal consistency
Triple response system

A PROLEGOMENON FOR INTERACTIONAL ASSESSMENT ......... 271

Interactionist Conceptions
The Main Features of Modern Interactionism
1. Reciprocal Causality
2. Intentional and Active Agents
3. Cognitive Personal Variables
4. The Analysis of Environment

Interactional Assessment

BIBLIOGRAPHY ........................................... 289
CHAPTER I: INTRODUCTION

Over a decade ago Walter Mischel (1968) published 'Personality and Assessment', a book that was widely perceived to be an attempt to debunk the traditional dispositional approach to assessment and usher in an alternative paradigm. The book was seen to be a situationalist manifesto aimed at undoing the role of dispositional explanations of human behaviour. However, Mischel (1979) claimed:

"My intentions in writing that book were not to undo personality but to defend individuality and the uniqueness of each person against what I saw as the then prevalent form of clinical hostility: the tendency to use a few behavioural signs to categorise people enduringly into fixed slots on the assessor's favorite nomothetic trait dimensions and to assume that these slot positions were sufficiently informative to predict specific behavior and to make extensive decisions about a person's whole life. My intention in 'Personality and Assessment' was to document the potential hazards to such attributions, of such categorisations often made on the basis of flimsy evidence. My aim was to call attention to the specific reciprocal interactions between person and context and hence to the need to examine those interactions in fine-grained detail. My concern was that clinicians, like other scientists tend to infer, generalise, and predict too much while observing too little. Moreover the judgments of clinicians - like everyone else's judgments - are subject to certain systematic biases that can produce serious distortions and oversimplifications in inferences and predictions." (p.245)

Mischel's intention in 'Personality and Assessment' is shared by the author of this essay and is the central motivating factor for writing the essay. The focus of this essay, like that of Mischel's treatise, is primarily on current psychodiagnostic assessment and classification, an enterprise that seems to have changed little in the last decade and of which a major component...
remains the administering and interpretation of various tests. These tests are many and varied and range from the Rorschach, in which the person being tested responds to a series of inkblots by describing what he or she 'sees', through more objective personality inventories such as the Minnesota Multiphasic Personality Inventory (MMPI) in which the person answers a series of questions about preferences and thoughts, to standard intelligence tests. Some of the more popular tests among psychodiagnosticians have been the Rorschach, the MMPI, the Thematic Apperception Test, various forms of an Incomplete Sentence Test, various intelligence tests, and certain tests that require the examinee to draw things. The model implicit in the administration of such tests is that the only variable is the 'personality' of the examinee. Thus, any differences in the Rorschach protocols, for example, produced by two different people are presumed to reflect differences in their personalities, differences in their preoccupations and concerns.

There have, of course, been other challenges, protests and revolts in addition to Mischel's critique - for example, in contrast to the implicit assumption mentioned above, Rotter (1960) emphasised the situational specificity of test behaviour, and, in Britain particularly, Shapiro (1961, 1966, 1970) inspired an experimental clinical model. In one of his earliest papers, Shapiro (1951) set out clearly the reasons for his dissatisfaction with current clinical psychological assessment procedures, especially those derived from the 'battery' approach of Schaefer (see Rapaport, Gill and Schaefer, 1946). His views are summarised in the following quotation.
"Psychiatric patients suffer from a variety of disorders of affect, cognition, and volition. A large number, if not the majority, of papers published by psychologists do not deal with these phenomena. Instead, they report upon the performance, by psychiatric patients, of a variety of tasks which might be described, without much loss of accuracy, as puzzles and indoor games. Examples are such tests as the pursuit rotor, the mirror drawing test, the block design test, the Rorschach test and the Thematic Apperception Test." (Shapiro and Ravenette, 1959, p.295).

Although one may disagree with Shapiro and Ravenette's evaluation of the various tests cited, their quote does sensitise us to an image of the psychodiagnostician. The traditional image of the psychodiagnostician as tester, the role expectations of other professionals, and the extraordinary faith in assessment devices that lack clinical utility have been hard to shrug off. As Clarke and Clarke (1973) caustically remark:

"[Assessment] has been classically regarded as the main contribution of psychologists. As currently employed it appears more often as an epiphenomenon keeping them busy, stimulating often unprofitable research and leading to a perpetual quest for the philosopher's stone (better and better tests) which will ensure more and more accurate prediction." (p.24)

Dependence on intuitive and invalid projective devices, lack of contact with the methods of experimental psychology and an over-reliance on the routinely administered test battery are all real criticisms of the stereotyped assessment that has characterised clinical psychology too long. For practising clinicians there is a need for assessment that is prescriptive (describing behavioural deficits and assets as the starting point for a treatment programme) and evaluative (providing an objective measure of treatment progress) rather than diagnostic (describing the client
with reference to some comparison population) or prediction (estimating the client's probable status at a later time). Peterson (1968) expressed this earlier in his conjecture that "the only legitimate reason for spending time... in assessment is to generate propositions which are useful in forming decisions of benefit to the person under study" (p.32). As we will see subsequently Shapiro would fully endorse Peterson's statement.

With respect to the above, it should be noted that one may construe people alternatively from many complementary perspectives. Thus if our aim is to seek strategies to induce change in performance the above focus may be most appropriate. Alternatively, construed from the perspective of the theorist interested in how therapy produces its effect it may be more useful to focus on competencies, personal constructs, expectations and other theoretical person variables that may mediate the effects of intervention on behaviour. In short, different goals require different foci and measurement strategies. This point is often ignored, as will be evident in the subsequent discussion. In this essay the status of the attribute model of assessment is oriented primarily to the goals of clinicians interested in assessing the problems of the individual.

The argumentation of both Mischel and Shapiro form the backdrop for the argument to be presented. As we have seen, Mischel's central point is that we are best served by viewing our clients as complex and multifaceted beings that defy easy classification and comparison on any single or simple common dimension; as being multiply influenced in a reciprocal fashion by a host of interacting determinants; as uniquely organised on the basis of prior experience and
future expectations; and yet, as rule guided in systematic, potentially comprehensible ways that are open to study by the methods of science. Hence it becomes difficult, if not meaningless, to attempt to achieve broad, sweeping generalisations about human behaviour: many qualifiers must be appended to our notions about cause-effect relations when dealing with human beings.

To this perspective, Shapiro (1951, 1961, 1970) brings a cogent argument for the application of the 'hypothetico-deductive' method of investigation to the individual case. For thirty years Shapiro has repeatedly argued that the clinical psychologist must be an experimental psychologist, though not of the conventional kind. He saw assessment and therapy as being inextricably interwoven in an on-going hypothesis generation-and-testing process, a position recently endorsed by proponents of behaviour assessment. The discriminativeness or 'specificity', and ideographic organisation in how behaviour is generalised and patterned across situations argued for by Mischel (1977, 1979) need not depress or discourage clinical assessors. On the contrary, more limited, specific, modest assessment goals may be refreshing for a field in which hubris has often exceeded insight. The requirement to qualify our generalisations about our clients does not prevent us from studying their behaviour scientifically. The recent spate of books on the intensive study of the individual case attests to this possibility (see, e.g. Hersen and Barlow, 1976; Kratochwill, 1978; Thoresen, 1979).

However, as was briefly alluded to earlier, psychological assessment has tended to be "... dissociated from the mainstream of contemporary psychology" (Anastasi, 1967, p.290) and that
"... testing today is not adequately assimilating developments from the science of behaviour" (p.300). Mischel (1968) has also pointed out that "... developments in personality assessment have been largely oblivious to advances in our knowledge of the conditions that change and influence human behaviour" (p.2). The principles that emerge from basic research have too often not been seen as directly relevant to the understanding of the determinants of test responses in the clinic.

In the assessment literature there appear to be two independent realms; the abstract and artificial situations of the researcher's laboratory and the 'realities' of life that confront the clinician and demand immediate attention. The resulting schism is graphically evidenced by the occasional proclamation from academic researchers to the effect that the psychodiagnostic enterprise is 'dying' and rather than prolong the agony, they advocate euthanasia by abandoning or at least reducing the time devoted to training graduate students in psychodiagnostic procedures. In stark contrast a comprehensive recent survey concluded that "... clinicians have definitely not abandoned testing" (Wade and Baker, 1977, p.879).

The Research/Practice Schism

Nowhere in psychology does the disparity between practice and the body of empirical research findings appear more glaring than in the domain of psychodiagnostic assessment and psychiatric classification. Although it has been repeatedly claimed that the role of psychodiagnostic testing, both in use and status has plummeted in the past two decades (Lewandowski and Saccuzzo, 1976;
Ivnik, 1977), recent surveys indicate that clinicians of all theoretical persuasions still devote substantial time to testing (Garfield, Kurtz, 1974; Wade and Baker, 1977).

Particularly interesting is the finding that the preferred assessment techniques have changed little since the early 50's (Lubin et al, 1971; Reynolds and Sundberg, 1976). Thus, despite the numerous inveterate critiques (e.g. Meehl, 1954; Mischel, 1968; Peterson, 1968; Stuart, 1970), despite claims that information provided by many objective and projective tests is not reliable or valid and lacks utility (e.g. Bem, 1972; Chapman and Chapman, 1971; Little and Schneidman, 1959), and despite the more recent calls for a totally new approach to testing (e.g. Bersoff, 1973; McClelland, 1973) clinicians have not abandoned psychodiagnostic testing.

Nor have they abandoned the traditional system of classifying problematic behaviour. For although other than at the grossest of levels, neither diagnostic classification per se nor test findings typically lead in themselves to differential treatment (Linder, 1965; Breger, 1968), and notwithstanding the documented unreliability of diagnostic classification and its questionable nosological foundations (Blashfield and Draguns, 1976; Cromwell and Blashfield and Strauss, 1975), psychodiagnostic labels still carry an authority and a stigmatising connotation that has not languished (Langer and Abelson, 1974; Stuart, 1970).

While academics may be convinced by the negative results concerning reliability, validity and utility - so much so that they have advocated the discontinuance of such enterprises - the
practitioner appears unconvinced and largely unaffected having sailed through the storm with almost complete aplomb.

Apparently the intrinsic deficiencies identified are insufficient to offset the information about personality structure they furnish the practitioner. Indifference to reliability and validity was patent in the survey conducted by Wade and Baker (1977). While poor reliability and validity were recognised as distinct disadvantages by most practitioners, these characteristics were not considered particularly important in test usage decisions. Instead personal experience with tests was the primary criterion in test selection. Unfortunately, a reliance on subjective impressions accrued through personal experience renders clinicians prey to such phenomena as 'illusory correlates' (Chapman and Chapman, 1971) and the 'Barnum effect' (Meehl, 1956; Snyder, Schenkel and Lower, 1977). Such effects can provide compelling and tenacious subjective indicators of efficacy and serve to reinforce untenable clinical assessment practice: untenable, at least from the perspective of the researcher. But are 'such phenomena' sufficient to account for the continued dominance of the field by the assessment devices spawned by the 'attributed' model?

How can the contradiction between the considerable evidence concerning the doubtful validity of many popular tests, the lack of reliability and utility of the psychiatric labelling process, and the fallibility of human judgement, be reconciled with the seemingly unshakeable confidence clinicians exhibit in their diagnostic ability? What promotes and maintains the 'illusion of validity'? These are questions addressed in this essay.
Old Paradigms Never Die - But They Should Fade Away

Thomas Kuhn (1970) has suggested that science progresses through the confrontation or clash of paradigms. Once the clash is over, the less-sensitive and less-comprehensive paradigm is not simply thrown on an historical garbage heap, but continues to struggle until its main proponents have died away. But importantly, by that time its task has been fulfilled and it has become a link in the cultural-sociological history. Thus old paradigms never die - they simply fade away.

It appears that in recent years, within the academic literature we have witnessed a sweeping shift, a still continuing dialectic process concerning the relative importance of the person and the environment in the control of human behaviour. According to Kanfer (1979) we are currently witnessing the synthesis stage of this dialectic - a synthesis that should result in the eventual abandonment of both prior positions. The true implementation of an approach that recognises the transactional nature of the person - environment relationship requires that the person be viewed at the psychological level as a component of the complex system of which he or she is a part (see, e.g. Buss, 1977; Howard, 1979; McReynolds, 1978).

However, it is doubtful that this 'revolution' has touched the practising clinician. It appears that s/he may well be operating in a different paradigm. For example, practitioner proponents of psychological testing are not in the main ignorant of the critical research findings in respect of reliability, validity and utility but generally feel that the relevant research is inadequate, that
accurate prediction of behaviour and accurate diagnostic assignments are not important goals of test usage (Holt, 1967; Weiner, 1972), that tests are too subjective and complex to be objectified and examined in an analytic fashion (Blatt, 1975) and as mentioned above, that personal experience with tests is a better basis for evaluation and a guide to practice than empirical data (Wade and Baker, 1977).

Kuhn (1970) further suggested that a paradigm, no matter how flawed and inadequate, is only made obsolete by the occurrence of a more viable alternative. On this point it is interesting to note that Wade and Baker (1977) found that a common argument for the continued use of traditional psychological tests was that they were seen to be the only economical and practical diagnostic tool available to the private clinician or institution. In this sense, their efficacy is proven by default. And neither the behavioural assessment nor interactional assessment paradigms are construed as viable alternatives.

To treat this as a sufficient explanation for the schism would be superficial and unwarranted. As we have seen paradigms arise and are maintained because they serve a purpose - that is to solve particular problems. If what is seen to be an obsolete or defective paradigm by one group fails to 'fade away', an analysis of the relationship between the solutions it provides and the socio-political intentions of a strong power group may prove illuminating. Such an analysis is undertaken in the essay.

Notwithstanding the obvious failure to agree on what constitute suitable criterion for evaluating current psychodiagnostic practices and the dearth of sound data, it is widely assumed that there has
been a marked decline in both testing and dispositional assignment. Numerous papers have been written elucidating the various events and forces that have been identified as antecedent and precipitating variables in the decline (e.g. Cleveland, 1976; Ivnik, 1977; Lewandowski and Saccuzzo, 1976).

On the Rise and Fall of Psychodiagnostic Assessment

As has been indicated, the value of psychodiagnostic endeavours has been hotly debated among psychologists for the last 20 years. Reading the scholarly journals, one could justifiably conclude that psychologists have forsaken - or certainly should have - traditional assessment procedures. However, whether the much debated decline in use and status of psychological testing and categorical assignment has substance remains a moot point. Personal opinion and theoretical polemic are a feature of the debate (e.g. Bersoff, 1973; Mischel, 1968). Rational evaluation based on sound empirical data is difficult as a feature of the scanty research available is its equivocal nature.

Chapter II reviews the relevant literature that addresses the question of the status and use of psychological tests. Surprisingly, in the light of the prior discussion, it concludes that despite a marginal decline in the time spent testing, psychological testing as a diagnostic procedure continues to be a frequent and significant activity of the clinical practitioner. In the light of this conclusion, the delineation of factors that purportedly have acted as antecedents to the decline appears a little misguided or at least mislabelled. It may be better to identify these factors as those
leading to disillusionment regarding psychodiagnostic assessment primarily in groups of non-practitioners.

It is patently obvious that the vast amount of negative evidence is unlikely to precipitate the downfall of the dominant assessment model - what is required is a viable alternative. However, even if a viable alternative did exist a major obstacle to its acceptance in practices would be the role expectations for clinical psychologists and the imperatives of their professional identity within the mental health professions. It is suggested that the current use and status of psychological tests cannot be fully understood independently of the nexus between test administration and the professional identity of the practising clinical psychologist.

Although a cogent argument can be put with respect to the inextricable link between psychodiagnostic assessment and intervention (see, e.g. Frank 1978; Shapiro, 1970, Yates, 1975), historically, the opposite has been the case. For many years clinical psychologists were excluded from performing therapy and were 'left' to do routine assessment work for psychiatrists. In retrospect, it is clear that clinical assessment was the major vehicle for psychologists to gain entry to the medically dominated 'mental health' arena and that the importance of this for subsequent attitudes and practices should not be overlooked.

In Chapter III it will be shown that the rivalry between the professions of psychiatry and psychology and the psychologists' struggle to repudiate the depreciated clinical role, were important determinants of our current assessment attitudes and procedures.
It will be shown that the early exigencies created an environment that promoted a 'proprietary dogma' surrounding testing that almost led to the development of an exclusive guild based on the mystification of testing and scientific pretence. The importance to early clinical psychologists of the justification of their practices by recourse to science is emphasised. It will be argued that many myths and false claims were perpetrated under the guise of being 'scientific'. Because of the authority and high esteem accorded to science in our society the benefits of such claims, although patent, are misguided. It is a central contention of this essay that science is only one of many available ways of gathering knowledge and is, of itself, not intrinsically superior to alternative methods. Each method of knowledge generation should be judged against the criteria selected for evaluating the knowledge produced, rather than against each other.

Thus, when it is argued in this essay that, for example, 'the history of psychometrics is a chronicle of pseudoscience', the label 'pseudoscience' merely conveys the judgement that the knowledge generated did not satisfy somebody's or some group's criteria for ascribing the label 'scientific' but that it aspired to be so labelled. It is not necessarily a pejorative label, although it has taken on a negative value in our society. It is sobering to realise that today's science may well be tomorrow's alchemy.

In a recent article Bersoff (1973) 'accused' the psychological assessment enterprise of being little more than psychological alchemy, little different from the assessment techniques of our predecessors (i.e. horoscopic astrology, physiognomy). It is easy
to disparage unkindly the assessment techniques of bygone days
and to admire our present psychometric sophistication, but can the
two be differentiated against the criteria of scientific knowledge?
Bersoff thinks not and posed the question: "Who is to be held
accountable for this psychological alchemy?" and his answer was:

"two brands of "psychos": psychoanalysts and
psychometricians. Psychoanalysts are to blame
because they have perpetrated a fraudulent
(Freudulent?) theory of personality and have
perpetuated its myth. Psychometrists, the test
constructors, are to blame because they have
forgotten their historical antecedents and have
become overly concerned with psychometric
aesthetics to the neglect of validity."
(Bersoff, 1973, p.892)

Psychoanalytic theory and psychometric theory come together and
are encompassed under the assumptions of the attribute model of
assessment since the model purports to measure both psychodynamic
states and traits. Thus if Bersoff's accusations can be substan-
tiated then grave doubts about the scientific status of the
attribute model must be entertained. For this reason it was
decided to include a chapter evaluating the charge of psychometric
pseudoscience and evaluating the charge of Freudian pseudo-
science.

Psychological Assessment as Alchemy

The charges of pseudoscience directed at contemporary psycho-
metric testing can be understood only by knowing how the discipline
developed its concepts, vocabulary and experimental procedures.
For this reason Chapter IV focuses particularly on the early
development of intelligence tests and the relationship with the
eugenics movement. Of particular interest is the transformation
of the concept of intelligence perpetrated at the hands of eugenicists. It will be argued that much of the work of the scientific testing movement was fallacy ridden principally because it was conducted within a defective paradigm. That is, the psychometricians operated within a paradigm that specified questions without having any available method of scientifically validating their answers. It will be shown that the history of intelligence testing is largely a chronicle of 'pseudoscience' with both defective argument and false conclusions becoming institutionalised primarily because of its pretence to 'science'. Although worthless as 'science', much of the resultant research did perform an important ideological service to 'the ruling elite'. Thus it will be argued that the psychometric intelligentsia participated in 'Lysenkoism' on a scale to rival the notorious example in Soviet agricultural science in the 30's.

Bersoff's (1973) charge against psychoanalysis was that it promoted a theory of behaviour that emphasised intraorganismic causal factors to the exclusion of situational determinants. According to this model, personality is perceived as a set of needs, drives, repressed impulses, trans-situational traits, etc. that initiate and guide behaviour. These dispositions are considered to characterise the individual and become translated into classificatory description. Because these traits and states are seen as underlying overt behaviour, special assessment devices, primarily projective tests, are designed to elicit responses (signs and symbols) that facilitate the accurate assessor inferences.
In Chapter V it will be shown that the extraordinary social and professional penetration of Freud's ideas was, to a considerable extent, promoted under the guise of 'being scientific'. It will be shown that against various criteria Freud's theories do not meet the requirements for the ascription of the label 'science', whereas there are considerable grounds to justify the labelling of the enterprise 'pseudoscience'.

Stereotyping, Clinician Fallibility and the Attribute Model

In the light of the conclusions of Chapters IV and V, Bersoff's claim that psychoanalytic and psychometric theory and practice do not provide a sound scientific foundation for the practice of clinical assessment does seem justifiable. For this reason Chapter VI is devoted to analysing the now dominant attribute model of psychodiagnosis. The Chapter is designed to undertake an analysis of the underlying assumptions implicit in the conceptual model and to evaluate the empirical support for these assumptions. This will lead to an examination of the model of causality that is intrinsic to the model and guides much professional clinical practice.

It will be shown that the 'quasi-medical' model of abnormal behaviour represents the logical extension of the attribute model to the explanation of grossly deviant behaviour. The assumptions implicit in the quasi-medical model and their implications for our daily clinical practice will be discussed. Lastly, the way in which the model defines normality/abnormality will be elucidated and the resultant tendency to medicalise social problems will be discussed.
The suggestion of a latent, socio-political function of the clinician's covert social mandate, raised in the previous chapter, will be developed in Chapter VII. It will be argued that a picture emerges which supports the assertion that diagnosis represents a transformation of stereotyped thinking and social prejudices into clinical descriptions that then serve as a basis for diagnosis. In this way the 'attribute model' converts undesired differences into deficits or symptoms. This is, psychodiagnostic categories may be construed as little more than scientific sounding social stereotypes that serve a distinct sociopolitical purpose. Empirical data to support this conjecture will be presented.

It is possible, however, to suggest that psychodiagnosis is a form of stereotyping in a different sense, namely, that the psychological processes in both everyday and social stereotyping and psychodiagnosis are similar and amenable to similar situational and information processing sources of bias and distortion. And that psychodiagnosis - being an interpersonal encounter - is an inevitably fallible endeavour. Empirical data attesting to the lack of reliability in psychological assessment and classification will be discussed.

The last chapter offers evidence corroborating the conjecture that trained professional clinical psychologists are generally as fallible in their inferential judgements as intuitive psychologists. Additionally, it will be shown that trained professionals are equally vulnerable to the temptations of magical thinking. As we have seen stereotyping and psychodiagnosis are in principle no different and thus we would expect magical thinking to be a pervasive characteristic of the psychodiagnostic enterprise.
The aim of Chapter VIII is to reveal the sources of biases, distortions, preconceptions and expectations in psychodiagnosis that interfere with the accuracy of clinical judgements. The evidence presented in this chapter casts severe doubts on the ability of a practitioner operating within the assumptive base of the attribute model to accurately code and process assessment information in a way that does not lead to systematic error and generate illusions of validity.

The central aim of this essay was not only to highlight the substantial disparity between psychodiagnostic practice and the body of empirical data evidencing its shortcomings, but to delineate plausible explanations for the failure of clinical psychologists to adapt themselves to the imperatives of the literature. Notwithstanding research discussed above, the failure of clinical psychologists to adapt themselves to this situation remains an outstanding anomaly. Clinicians to this day continue to try to confine their assessment operations to general traits, and in so doing they concern themselves therefore with only a relatively small part of the reliable variance of behaviour and experience. When one places the great variety of human phenomena in the context of relative specificity, one is led to expect a unique pattern of dysfunctions and related variables in each individual client. However, a conventional psychometric approach which aims to measure common variance only, and not specific variance, does not acknowledge this and cannot therefore be expected to meet our clinical needs and inadvertently serves to maintain what Mischel (1979) identified as the 'prevalent form of clinical hostility' (p.245).
In the earlier chapters several of the more important research/practice anomalies were discussed along with a delineation of the factors and processes that maintain what appears to be untenable practices. However, before considering what may be done to change the situation, if that is desirable, it was decided to discuss briefly, factors that may influence the clinician's initial approach to psychodiagnostic assessment.

In their recent survey, Wade and Baker (1977), found that 70.7% of their respondents first learned to use assessment instruments in graduate training, 27.7% in undergraduate training and only 1.6% in postgraduate employment. These figures strongly suggest that initial training instructions are to be held at least partly responsible for the current situation. This is augmented by other findings, for example, only 20.6% of respondents indicated that they systematically collected and analysed data regarding their own testing practice. Additionally, although 54.7% indicated that they read at least several articles relevant to testing every six months, only 25% stated that the studies critical of tests seemed accurate. The majority of respondents indicated that these studies employed inappropriate criteria or questionable methods, that they over-generalised or that they reported conflicting findings.

It will be argued in Chapter IX that the initial training of clinicians in Universities and Colleges appears to fail in at least two important respects. Firstly, they continue to teach students to administer and interpret tests of dubious validity, and secondly, they fail to develop behaviours compatible with the ideals embodied in the concept of 'scientist-practitioner'. It is beyond the scope
of this essay to undertake an analysis of the factors that could explain this situation but it should be reiterated that negative evidence may not prove sufficient to dissuade neophytes from using the 'offending' tests, especially when faced with the role expectations of mental health colleagues. A necessary but probably insufficient requirement will be the provision of a viable alternative.

A Prolegomenon for Interactional Assessment

In the light of the importance of initial training of clinicians for later assessment practices, the conclusion (Chapter IX) will focus specifically on the implications of the earlier chapters for the teaching of clinicians. The specific aim is: Firstly, to recommend ways of improving clinical assessment conducted within the dominant approach. This will be intentionally brief as it is the author's view that the paradigm is not the 'best' available. It has been argued that the model is not only conceptually and empirically inadequate but that it facilitates the medicalisation of deviance and supports a 'blame-the-victim' ideology (Ryan, 1971). Although the essay has been largely negative it will be argued that it is possible to convert some of the critical findings into recommendations that may help clinicians avoid some of the shortcomings. Although this may amount to little more than 'mending a thread bare' paradigm, it cannot be disputed that the attribute model remains the dominant assessment model and as such, until a viable alternative is accepted, the onus is on trainers and researchers to develop strategies that diminish the errors and biases in diagnostic assessment conducted within the paradigm. Thus in this section several recommendations
for inclusion in the training course for students in psychodiagnostic assessment will be made. They include suggestions on how to improve the clinician's ability to learn from experience, on how to sensitise clinicians to the importance of the cognitive structure and strategies they bring to the assessment situation, on how to sensitise the clinicians to the influence of their values and biases in clinical judgement, on how to give up the luxury of global labelling, and so on.

Secondly, an already available alternative assessment paradigm, namely, behavioural assessment, will be briefly introduced. Again its presentation will be brief as it is considered to be predicated on an equally limited and truncated image of humans. However, it will be argued that many of the problems associated with the attribute approach and delineated in the earlier chapters, are avoided when the behavioural model is applied. Thus it will be argued that it is a preferable approach.

Although at the theoretical level social learning theorists have emphasised the reciprocal interaction of the person and the situation, rarely has this commitment manifested itself in practice. In practice a rather static, unidirectional model of causality where the individual is construed as relatively passive seem to dominate. This is particularly the case with much of the operant assessment therapy. It will be argued that the behavioural assessment model does not represent 'true' interactionalism, rather it adopts what Buss (1977) referred to as a 'mechanistic' paradigm and Overton and Reese (1973) referred to as a reactive model. That is, behavioural assessment assumes that 'independent variables' affect 'dependent variables' whereas modern interactionalism is concerned with
reciprocal interaction between environmental events and behaviour.

Lastly, an approach to assessment predicated on a recognition of the transactional nature of the person-environment relationship that requires that the person be viewed at the psychological level as a component of the complex-system of which he/she is a part, will be advocated. This interactionist perspective does not represent a rapprochement between the opposing theoretical camps represented in the two alternatives listed above (i.e. personologism and situationalism) but a synthesis of the dialectic. It represents a dialectic synthesis that should result in the eventual abandonment of both prior positions. The recognition that complex human behaviour tends to be influenced by many determinants and reflects the almost inseparable and continuous interaction of a host of variables both in the person and in the environment, has deep implications for psychodiagnostic assessment.

Although it is widely accepted among psychologists that the behaviour of a given person in a particular environment is a joint function of the relevant behavioural determinants within the person and those within the environment, few researchers appear to operate within the paradigm. However, it will be argued that when examined more concretely, commitment to this paradigm has deeper implications for psychology in general, as well as specifically for psychodiagnosis. For, if it is useful to construe behaviour as a joint function of person and situation determinants,
then it follows that psychological assessment should systematically take account of both person and situational variables in their transactions. It need hardly be said that this is rarely the case in contemporary assessment practice.
CHAPTER II: ON THE DECLINE AND FALL OF PSYCHODIAGNOSTIC ASSESSMENT

We are, according to McReynolds (1971) at the beginning of what, in retrospect, will someday be seen to be a revolution in psychological assessment. The revolution - or rapid evolution - has been precipitated and fueled by numerous influences, prepotent among the influences being:

(a) the widespread disillusionment with the 'traditional' approach to psychodiagnostic assessment, personified by the battery approach advocated by Shafter and Rapaport; (Rapaport, Gill and Schafer, 1946)

(b) the trenchant criticism of both the conceptual and empirical adequacy of the theory of human behaviour that stresses intraorganism determinants to the exclusion or expense of situational determinants (e.g. Bandura, 1969; Mischel, 1968); and

(c) the dramatic proliferation of behaviour therapy (Benassi and Lanson, 1972; Kazdin and Wilson, 1978) and the, albeit belated, elucidation of a technology of behavioural assessment (Ciminero et al 1977; Cone and Hawkins, 1977; Hersen and Bellack, 1976).

Concomitant with the attacks on psychological assessment, the traditional Kraepelinian system for classifying problematic behaviour has undergone increasingly intense scrutiny. Numerous proposals have emerged for revising, reformulating or abolishing the system of classifying psychological disturbances. Generally criticism has been stimulated by three considerations, namely:
(a) the empirical adequacy of the current psychiatric nosology reliability, validity and utility (Zigler and Phillips, 1961; Zubin, 1967);

(b) the relevance of the psychiatric labels for traditional treatment (Meehl, 1960; Mintz, 1968) and particularly for the newer modes of psychotherapy, for example behaviour therapy or humanistically inspired therapy (Kanfer and Saslow, 1969; Rogers 1973); and

(c) the assumptive base and undesirable consequences of equating problematic behaviour with 'mental illness' and designating some people 'mentally ill'. (Leifer, 1964; Sarbin, 1967; Szasz, 1960).

Overwhelmingly the attack on the traditional psychodiagnostic endeavours of clinical psychologists has been mounted by academic psychologists and sociologists. A survey of the literature reveals that some academic clinicians believe the psychodiagnostic enterprise to be 'dying' and rather than prolong the agony, they advocate euthanasia by abandoning or at least reducing the time devoted to training graduate students in psychological assessment and diagnosis (Levy and Fox, 1975). Meanwhile, at the work face, clinical practitioners are faced with the realities of persons coming to clinics with a myriad of different complaints and problems and in all cases some clinical decision must be made and communicated to colleagues. Despite the veracity of the critiques and inveteracy of the deficiencies, psychological tests and dispositional labels are often found to be helpful in the decision-making process (Maloney and Ward, 1976; Wade and Baker, 1977). Clinicians are apparently blissfully ignorant of the hypothetical illness that besets their procedures and have failed to note the marked progressive deterioration over time.
An important consequence of the co-existence of disparate views on the value of psychodiagnosis has been the development of an obfuscatory schism between the academic/researcher and the clinical practitioner. The present chapter explores the two divergent points of view as to the value of traditional psychodiagnosis. It also delineates some of the factors identified as precipitating the much vaunted, and occasionally mourned, demise of such endeavours. Subsequently it will be argued that rather than 'dying' psychodiagnosis is 'alive and well' (Levy and Fox, 1975) and unquestionably the dominant assessment system used by practising clinicians irrespective of their therapeutic persuasion (Wade and Baker, 1977). It will be shown that projective and objective testing still constitutes a significant part of the clinician's professional activity.

FACTORS LEADING TO THE DECLINE OF PSYCHODIAGNOSTIC ASSESSMENT

Obviously some of the forces purported to have precipitated the decline arise from changes in the sociopolitical milieu and are so pervasive and ephemeral that their identification is difficult. Others however are more definitive and yield more easily to identification. Although the purported demise of psychodiagnosis has not yet been corroborated, authors have readily delineated the factors that have contributed to the decline. Some of the more frequently cited factors will be introduced in a summary form below. Most of them will be discussed in much more detail and critically appraised in later chapters. Among the factors identified are:

1. the negative research literature;
2. the invocation of psychopathology;
3. the stigmatising dispositional nosology;
4. the treatment irrelevance of psychodiagnosis;
5. the rise and popular acceptance of humanistic/existential perspectives;
6. the exaggerated emphasis on intraorganismic determinants;
7. the vilification of tests in the media;
8. the poor academic training in psychodiagnostic assessment;
9. the antiquated nature of many instruments; and
10. the cost of the enterprise.

1. Negative Research

Allusion has already been made to the negative results concerning reliability, validity, and utility of psychological tests and dispositional labelling. It would be conservative to conclude that the traditional methods of assessment and diagnosis have been convincingly criticised on both conceptual and empirical grounds (see Mischel, 1968; Peterson, 1968; Stuart, 1970). Additionally, overwhelming evidence for the superiority of actuarial prediction over clinical prediction has been offered by Meehl (1960) and Sawyer (1966) for nearly 20 years, yet clinicians continue to ignore the former and use the latter.

Proponents of psychodiagnosis argue that, although generally negative, the evidence is equivocal, being ambiguous, indecisive and methodologically flawed (Blatt, 1975; Lewandowski and Saccuzzo, 1976; Weiner, 1972). They have argued that acceptance of the negative results reflects more the current zeitgeist than the compelling nature of the data. For example Blatt (1975) admits that a simple tally of studies would indicate:
"little, if any, support for the validity of projective procedures. But why has there been so much uncritical acceptance of the research findings? Could it be that the same social-psychological forces which affected the attitudes of many psychologists towards psychodiagnostic assessment may also be the reason for the eager and uncritical acceptance of negative research findings. Results are far from disheartening ... In fact, a large proportion of these studies offer considerable support for many of the interpretative assumptions of projective techniques." (p.333)

No doubt Blatt is correct in asserting the importance of the dominant zeitgeist in data interpretation (see, e.g. Chalmers, 1976; Kuhn, 1970; Mahoney, 1976). As Mannheim pointed out several decades ago, protagonists in the present controversy "speak as if their differences were confined to the specific question at issue around which the present disagreement crystallised. [i.e. the validity of projective tests] They overlook the fact that their antagonists differ from them in his [her] whole outlook and not merely in his [her] opinion about the point under discussion". (Mannheim, 1936, p.251). Because the controversy is fundamentally a confrontation between competing conceptual systems or paradigms (Kuhn, 1970), empirical critiques are largely inefficacious. In his autobiography, Planck (1950) remarked that "new scientific truth does not triumph by convincing its opponents and making them see the light, but rather because its opponents eventually die, and a new generation grows up that is familiar with it." (pp.33-34) If Planck is right, reasoned argument and evidence do not play a large part in changing scientific paradigms and knowledge. Thus, in this case, rather than the negative research being responsible for precipitating change, it may be that the generation of psychologists committed to psychodiagnosis are dying off and a new generation of clinical psychologists is growing up committed to an alternative paradigm. But as we have seen some paradigms won't 'fade away'.

2. Invocation of Psychopathology

A second variable identified as contributing to the devaluation of psychodiagnosis has to do with the psychopathological focus of much testing. Test and test reports tend to emphasise maladaption and dysfunction at the expense of the client assets, personal resources and 'healthy' characteristics (Kanfer and Saslow, 1969). Numerous studies are available (see, e.g. Soskin, 1954; Samuel, 1965) demonstrating that judges exposed only to subjects' Rorschach and Thematic Apperception Test (TAT) test protocols invoked significantly more psychopathology in their evaluations than did judges who were familiar with the subjects social history or who personally interviewed them.

Bersoff (1973) argues that some projective tests, such as the TAT typically elicit symptomatic responses because of their composition, being grey and sombre and portraying sad scenes. To the question: "What happens to a person who is referred for some behavioural difficulties and is administered the TAT?" Bersoff answers:

"Well, like the women accused of being witches in older times she is not likely to come out of the assessment experience too healthily. During the sixteenth century, an alleged witch's guilt was tested in an ingenious way ... The examination involved dumping her in deep water restrained by ropes. If she sank, she was innocent; if she floated, she was guilty - in either case she was dead. And like the witch, an individual referred for projective testing often ends up in the same position. Given cards that normally evoke themes of death or sadness or aggression s/he is normally going to tell stories that involve these themes. Ignoring the fact that his/her behaviour is the result of the interaction of what s/he brings to the test and the nature of the test itself, in 'essence' terminology the testee is likely to be labelled 'depressed', 'aggressive', etc. and it will be assumed that this is the way s/he is in most, if not all situations." (1973, p.896)
If the goals of psychotherapy are essentially reconstructive (Lorr and McNair, 1966), then it would seem imperative that diagnostic testing shows the strengths upon which improved functioning must be premised (Martin, 1966; Scarbrough, 1966). Certainly with the increasing popularity of behaviour therapy the need for assessment strategies capable of identifying the environmental variables - internal as well as external, self-imposed as well as imposed by others - that are currently maintaining the individual's maladaptive thoughts, feelings and behaviour are required rather than dispositional diagnosis (Dickson, 1975; Mash and Terdal, 1974; Ross, 1974). Thus, as a tighter link is forged between assessment and intervention procedures the focus on dispositional labelling will become less useful.

3. Psychiatric Classification

It has been suggested that many students nowadays enter postgraduate training in clinical psychology with negative attitudes to diagnostic classification as a result of the popularisation of the societal reaction 'theory' of deviance (Becker, 1963; Matza, 1969; Schur, 1971) and the application of this perspective to the concept of mental illness (Sarbin, 1967; Scheff, 1966; Szasz, 1961). Their negative attitudes are often seen to be reinforced by the misuse that institutions and the public make of diagnostic labels, i.e. to stereotype and stigmatise (Farina et al, 1966; Hobbs, 1975; Jones, 1972). The widespread publicity given to Rosenhan's (1973) pseudopatient study, when coupled with movies like 'One Flew Over the Cuckoo's Nest', 'Family Life', and 'Outrages', serves to augment negative attitudes: as does the bemusing decision by the American Psychiatric Association to exclude homosexuality from the list of
mental illnesses by majority vote in 1974 and the equally bemusing removal of the mental illness of 'sexual sadism' from their latest guide book for diagnosis of mental disorders (DMS-III) following a successful feminist lobby (Goleman, 1978).

The issue takes on a more philosophical bent when one considers Sartre's assertion that labelling violates the integrity of persons and Szasz' claim that the ultimate freedom of all individuals is not to be labelled against their will. Rather than calling people names, individuals should be experienced and related to intuitively - according to the rhetoric of some humanistic/ existential therapists (Dana and Leech, 1974; Strupp, 1976). In his article 'Clinical Psychology Irrationalism, and the Erosion of Excellence' Strupp (1976) makes the point that the notion that diagnosis and evaluation are intrinsically antithetical to the proper conduct of psychotherapy and thereby harmful to individuals is "one of the greatest pseudo-issues besetting our field" (p.564). Rather than be a rigid and static assignment to a dispositional category, Strupp suggests that assessment should be a process of forming hypotheses and decision-making (Strupp, 1976). However, the traditional separation of assessment from psychotherapy militates against this experimental approach to diagnostic assessment so cogently advocated by the little known Shapiro (1951, 1961, 1966).

Psychologists, much like other scientists have shown a strong penchant for categorisation (Wallace, 1966). While cognisant of the arbitrary boundaries which, as a matter of convenience, separate one diagnostic category from another, the ever present danger is that categories are used 'as if' they are mutually exclusive and non-
overlapping. Granted that each individual is unique, then her/his interactions with the clinical assessor provides a set of unique data which has to be condensed and categorised. Following the logic of scientific enterprise propounded by Popper (1963) no one classification can logically be said to equal the unique original data. Szasz (1961) highlights the naivety of the view that there exists a finite set of classifications which can successfully encompass an infinite variability of client patterning. It would seem important to have more than one set of categories at one's disposal and importantly, to avoid confusing any one category with the unique data (Feyerabend, 1970).

Psychologists have not acknowledged, or at least not promulgated the notion that psychiatric nosology is pragmatic, provisional and should be explicitly open to change in response to experience (Phillips, Dragans and Bartlett, 1975). As a consequence, the concept of 'mental illness' and the language system used to describe particular kinds of 'mental illness' have become legitimised as scientific 'facts': their metaphoric nature being lost. Given the fractious and acrimonious debate which rages over the legitimacy of these 'facts', it is hardly surprising that there have even been attempts to abolish the concept of 'mental illness' altogether and replace it with some alternative, non-pathological construct such as deviance, problems of living, or community disorder. It is easy to see how this debate could contribute to the decline of psychodiagnostic assessment and classification.

4. Treatment Irrelevance

Another factor contributing to the growth of negative connotations surrounding psychodiagnosis is the apparent irrelevance of assessment decision for therapy. Traditionally psychotherapy and diagnosis were
relatively independent endeavours, being performed by psychiatrists and clinical psychologists respectively. And certainly, for most of us, helping a person change for the better is more gratifying than doing a psychological assessment. Under these circumstances it is not surprising that there is a differential valuing of the two processes. However, if assessment and diagnosis are conceived to be prescriptions for action rather than descriptions of the client (Linder, 1965) then it is appropriate to expect that psychodiagnosis will precede psychotherapy and to a large degree determine its form.

Generally it can be shown that dispositional diagnosis does not predict which treatment will be undertaken (Coles and Magnusson, 1966; Kenfer and Saslow, 1969). For example, in an early study Meehl (1960) revealed that while 17 per cent of the clinicians studied reported that prior knowledge of the client was important in effective therapy, 80 per cent felt the therapists' characteristics are more important determinants of the nature of therapy and its outcome. Similarly Moore et al (1968) who questioned psychiatrists about the function and value of psychological diagnosis found that the majority of respondents indicated that psychological evaluation:

(a) is requested less than 20 per cent of the time;
(b) is not viewed as an essential part of the client work-up;
(c) is not used in making treatment and diagnostic decisions;
(d) would not be used in treatment planning even if changes were made to improve the assessment reporting system.

Although these findings appear to be telling, the assumption is that psychiatrists' opinions about the value of the psychologist's assessment report are a valid criterion for assessing the report; this would probably be challenged by most clinical psychologists.
Nevertheless, that psychodiagnosis is not used as to prescribe 'treatment', the value of the enterprise vis-a-vis psychotherapy is understandable and the implications of this to the hypothesised decline appear patent.

5. Dehumanising

Humanistically oriented psychologists have accused assessment processes of being discriminatory, stigmatising and dehumanising. Test proponents would claim that this criticism is more properly addressed to the abuse and inappropriately administered tests and that testing per se need not be an aversive event for the client. However, this is not the case according to Dana and Leech (1974) who argue that "the easy recourse for the assessor to power (omnipotence) and implicit magic have historically diminished the humanity of the person who is the client and rendered assessors unwitting professional purveyors of dehumanising experience" (p.431). In a similar vein Schafer (1954) has labelled the assessment process as voyeuristic, oracular and autocratic.

With the advent of client-centred therapy, the notion that psychodiagnosis is intrinsically antithetical to the establishment of an appropriate client therapist relationship gained currency. The rise of the human potential movement, humanistic psychology and existential therapy extended this view and gave rise to "the contemporary trend toward anti-intellectualism, antiscientism, and antiprofessionalism" (Strupp, 1976, p.561) and the concomitant eschewing of psychodiagnostic assessment.
Strupp's (1976) diatribe notwithstanding, empirical data does corroborate the sexist nature of cross-sex assessment (Chasen and Weinberg, 1975; Harris and Masling, 1970; Masling and Harris, 1969), the racist nature of some assessment instruments (Mercer, 1973; Williams, 1974), the impression management that invariably is present in assessment situations (Braginsky, 1970; Goffman, 1961), and eisegesis (Dana, 1966). The humanist critique is well summarised by May (1958). Speaking for existential analysis May states:

"Existential analysis is a way of understanding human existence, and its representatives believe that one of the chief (if not the chief) blocks to the understanding of human beings in Western culture is precisely the over-emphasis on techniques, an emphasis which goes along with the tendency to see the human being as an object to be calculated, managed, 'analysed'." (1958, p.76)

No doubt the widespread influence of the existential/humanistic critique has done much to promote the rejection of traditional approaches to psychodiagnosis in some quarters.

6. Organism Error

The sixth factor often identified as contributing to the disillusionment is the persistent de-emphasis of situational determinants of behaviour (Mishel, 1968; Peterson, 1968). Psychologists are not alone in favouring intraorganismic variables in explaining behaviour; lay 'intuitive' psychologists display a similar persistent propensity to understate the influence of contextual factors. So pervasive is this proclivity that Ross (1977) refers to it as the 'fundamental attribute error'. Earlier Wallace (1966) had suggested a similar point of view in his suggestion that the dominant attribute model of assessment construes personality as 'essence'; an idiosyncratic character intrinsic to the individual
which stands behind actions and motivates them. The 'essence' concept of personality rests on a 'within-skin' model of causality and promotes an assessment focus upon a host of psychodynamic forces - for example, repressed impulses, energised traits, latent tendencies and needs - all hypothesised to be capable of initiating, and guiding behaviour somewhat independently of the current situation.

There now exists a substantial body of knowledge attesting to the situational specificity of much behaviour and conversely, a lack of large transitiational consistencies in behaviour (see e.g. Bem and Allen, 1974; Mischel, 1968, 1973; Schweder, 1973). Thus responses emitted in assessment situations (as has been alluded to earlier with respect to impression management) are not only a function of the characteristics of the individual but also of the contextual influence (e.g. sex and race of assessor, temporal and physical etc.) of the test situation (Bersoff, 1971). Accordingly, Wallace (1966) asks "What is the relationship between the ability to make certain verbal responses in a make-believe setting and the ability to respond in a similar fashion in overt behaviour?" (p.135). There may be very little.

Holland and Richard (1965), for example, report that test information correlates poorly with real life performance and McClelland has recently critically analysed - what for a long time seemed sacrosanct - the relationship between high IQ and 'success' in society. He concluded that 'the correlation between intelligence test scores and job success often may be an artifact" (McClelland, 1973, p.51) rather than an example of causal covariance.

It is now widely acknowledged that assessment does not occur in an 'organismic vacuum'. One has only to read the specific instruction to testers designed to facilitate the establishment of rapport and high
motivation in order to promote optimal performance during administration of intelligence tests (e.g. Terman and Merrill, 1960, p.50) to evidence the test makers' sensitivity to this point. Consequently if we are genuinely interested in gathering data about behaviour that is predictive and relevant to the current functioning of an individual outside the test situation then we must, according to Bersoff (1973), discard the tests that are flagrantly inappropriate to the purpose, i.e. those that search for trans-situational intraorganismic factors. If the traditional approach to assessment is to stem the tide of discontent then it must aim at 'contextualising' assessment and discover what Fischer (1969, 1973) called the 'when/when not' of specific behaviour, or in Mischel's (1979) words to discover the "discriminativeness or 'specificity' and ideographic organisation in how behaviour is generalised and patterned across situations". (p.742)

7. Poor Media Coverage

Much of the mystique of testing and assessment endeavours has been set aside by critical exposes in the popular media, an example being the disclosure of accusations of data manufacture and 'cooking' directed at Sir Cyril Burt, in the press (e.g. Age 1979; Encounter, 1977; New Stateman, 1978). Additionally several popular books have also carried the message that tests are fallible, unreliable and amenable to faking and socially dangerous (e.g. Whyte's (1966) 'The Organisation Man'; Black's (1962) 'Thou Shall Not Pass'; and Banesh's (1962) 'The Tyranny of Testing'). The spate of recent films mentioned earlier also served to cast doubt on the ability of psychologists to accurately discriminate between the 'sane' and 'insane' and to cast doubts on the humanity of institutional 'help' (e.g. 'One Flew Over the Cuckoo's Next'; 'Family Life' and 'Outrages').
return the public have accused psychologists of everything from "reading our minds and invading our privacy" to "imposing a yoke of conformity on employees and stifling originality in school children, driven by morbid sexual curiosity and cynical pursuit of money" (Rapaport et al, 1968, p.23).

Public concern over testing is also evidenced by the legal challenges to testing in schools and university. For example, in the U.S.A., in what promises to be a landmark decision after a seven year court battle, San Francisco's chief district court judge has ruled that the use of standardised intelligence tests to categorise children as mentally retarded discriminates against minorities and is thus unconstitutional. The immediate effect will be to make permanent a state-wide injunction against these tests. Several other states in the U.S.A. have similar and occasionally more wide-ranging laws restricting the administration of intelligence tests to school children. A petition requesting a moratorium on all standardised testing was recently submitted to the American Psychological Association by the Association of Black Psychologists (Mercer, 1978-79). Such negative attitudes within the profession can only further exacerbate the declining popularity of traditional psychodiagnostic assessment procedures.

8. Academic Training

The eighth factor identified as contributing to the declining status of psychodiagnostic assessment is the poor quality of academic training of clinical psychologists. Several articles have recently voiced the dissatisfaction of staff of interim agencies with the academic preparation of students in the area of psychodiagnostic
assessment (e.g. Garfield and Kurtz, 1974; Shemberg and Keeley, 1970). Students were perceived as having overly critical attitudes to assessment and insufficient experience in practical settings. University instructors were often accused of, at the best, half-hearted endorsement and at the worst downright cynicism about tests. Accordingly it is thought not surprising that new practitioners approach practicums and internships with little enthusiasm or confidence about diagnostic assessment procedures (Russ, 1978). The accuracy of this conjecture will be evaluated a little later in this chapter.

9. Aging Assessment Instruments

Another factor identified as contributing to the decline of testing involves the age of some of the most used tests. Every one of the ten most frequently used psychological tests in Lubin, Wallis and Paine's (1971) list was at least twenty years old. Several were very much older, for example the Rorschach is now approaching 60 years in print, the T.A.T. 50 years and the Minnesota Multiphasic Personality Inventory (M.M.P.I.) 40 years. Although revised, some of the vocabulary items in the Stanford-Binet date back to 1908 (see Popplestone and McPherson, 1974). Many test items are almost totally irrelevant in the contemporary context and their anachronistic qualities serve as a source of humour rather than penetrative insight into the unconscious of the client.

10. Cost

A last factor identified as contributing to the declining status and use of psychoassessment procedure is the cost. Cleveland (1976) suggests that inflation has led to an escalation of clinical psychologists' hourly fee resulting in a situation where private practitioners
can ill afford to spend the time necessary to administer, score, interpret and write up a psychological test battery. He suggested that there is a real possibility that clinical psychologists may price themselves out of the market and inadvertently accelerate the already declining status and usage of their assessment services.

Without recourse to sound data however, it becomes almost impossible to evaluate the different accounts of the status of psychodiagnosis and to identify accurately the influential precursors to the decline and fall of the enterprise. Much of the foregoing discussion with respect to antecedents of the purported 'decline' appears to be based on personal opinion, theoretical polemic and indirect data. And unfortunately, as is often the case, once a belief is adopted (i.e. that the value and usage of tests is declining) it can become almost impervious to numerous disconfirming instances. Popper's (1963) terms such a belief a 'prescientific dogmatic attitude'. Many of the antagonists in the current debate appear to evidence this attitude as the data indicates that the reports of the death of psychodiagnostic assessment are greatly exaggerated as we shall see below.

PSYCHODIAGNOSIS IS ALIVE AND WELL

Fortunately, there exists a body of research literature concerning the use and status of psychological tests among clinical psychologists. This research can be arranged according to the question that different investigators attempted to answer. These questions are:
(a) do psychologists in academic settings view psychological tests differently from their colleagues in the professional clinical field?

(b) has the amount of time that professional clinicians devote to psychodiagnostic activities declined?

(c) has there been a movement away from using the most frequently and trenchantly criticised projective techniques?

(a) Academic Prejudice Against Testing

Several authors (Garfield and Kurtz, 1973; McCully, 1965; Shemberg and Kelley, 1970; Thelen, Varble and Johnson, 1968) have compared the evaluation of testing by clinical psychologists working in university graduate training programmes with those made by their colleagues in internship centres. There was a high degree of consensus - university psychologists hold a more critical or negative view of psychological tests, especially the projective tests, than do their professional counterparts. Interestingly, younger faculty members (under 36 years) were more pessimistic and more inclined to discard projective techniques altogether, than were members over 45 years (Thelen, Varble and Johnson, 1968). Parenthetically, this may fit the evolutionary change process described earlier by Planck (1950). Because of the relatively low status of psychological assessment in universities it is usually the younger and least experienced faculty member who teaches such courses (Cleveland, 1976). This practice may inadvertently be the most important sociohistorical factor in the ensuing paradigm choice and use, and thus be the most salient agent of change.
The dissatisfaction of internship directors with the preparation of clinical psychologist trainees' skills in test administration and interpretation was first documented by McCully (1965). The situation had changed little when Levitt (1973) repeated the McCully survey. The trade off between innovation and established practice is a difficult choice for training institutions to make and will probably always be a source of theory/practice tension.

It does seem clear that academic clinical psychologists, particularly younger ones, are more negative about the claimed value of psychological tests, particularly projective techniques. It would be naive to argue that this attitude is totally shaped by their sensitivity to empirical data and that this sensitivity is due to their scientific posture - a posture that according to the 'story book model' requires an active testing of beliefs, acceptance of negative evidence, and a willingness to abandon previously held ideas (Merton, 1967; Mitroff, 1974). It would be equally naive and misguided, to accept Rapaport's et al (1968) conjecture that:

"Non-clinical psychologists who saw their own hegemony in the science and profession threatened by the far greater growth rate of their applied colleagues often reacted with suspicion, hostility and rejection. Being in positions of power in university departments, they put the young clinical faculty members on the defensive, to which many reacted by identification with the aggressor ..." (p.25)

A more tenable explanation is forthcoming if one:

(a) examines the psychosocial influence operating on the clinical psychologist in the university and applied setting; and
(b) applies to insights provided by Thomas Kuhn (1970) and others to an analysis of scientific belief and behaviour.

This will be undertaken in Chapter III. For the present, one can conclude that the conjecture that academic psychologists are more negative in their attitudes than their professional colleagues can be readily corroborated.

(b) Time Devoted to Psychodiagnosis by Clinicians

A cursory examination of published surveys at first suggests that practitioners are spending increasingly less time doing diagnostic assessment. Garfield and Kurtz (1974) surveyed one third of the clinical psychologists who were members of the American Psychological Association and found that diagnostic assessment accounted for approximately 10 per cent of the profession time. This activity ranked fourth among all activities; exceeded only by individual psychotherapy (25 per cent), teaching (14 per cent) and administration (13 per cent). But although only 10 per cent of the practitioners' total time was devoted to assessment this represents 24 per cent of the psychologist's direct clinical services. The figures of Garfield and Kurtz (1974) are a little below the more recent figures provided by Wade and Baker (1977). They surveyed a similar population but additionally asked respondents to indicate the amount of time devoted to objective and projective testing separately. Interestingly, there were no significant differences between the percentage of time devoted to administering both. Clinicians report spending 7 per cent of their time administering projectives and 7 per cent administering objectives, giving a total of 14 per cent spent on diagnostic assessment.
Although comparisons with previous surveys are difficult, the recent research cited above is consistent with the prior conclusions (see e.g. Thelen, Varble and Johnson, 1968) that diagnostic activities have diminished in relative importance for the clinical psychologist since the early 50's. The decline evidenced above rather than being attributable to general disillusionment with assessment techniques may be the result of clinicians devoting extra time to the more glamorous and prestigious activity of psychotherapy. In spite of the decline in relative importance, assessment activities continue to constitute a significant proportion of the psychologists' professional labours.

In contrast, Holt's (1967) paper clearly attests to the decline in time devoted to assessment by more research oriented clinical psychologists. Using the annual A.P.A. conference (Division 12 - Clinical Psychology) Holt found that in 1955 63 per cent of the annual conference programme was devoted to diagnostic assessment whereas the figure had shrunk to 35 per cent by 1965. By 1975 only 7 per cent of the A.P.A. conference programme in Division 12 was devoted to psychodiagnosis. However, Reynolds and Sundberg (1976) using the number of references cited in the bibliography of Buros' 'Mental Measurement Yearbook' found that research publications continue at an extraordinary rate. For example the Rorschach has had an average of one reference every four days for almost fifty years; the M.M.P.I. after 30 years has averaged one reference every 2½ days.

This high level of research activity, when coupled with the significant amount of clinical time devoted to assessment, evidences the 'aliveness' of psychodiagnosis, both in the applied setting and to a lesser extent, in the university setting.
(c) On the Decline of Projective Techniques

Allusion has already been made to the marked discrepancy between the reliance placed on projective techniques, particularly the Rorschach, in many clinical settings and the frequently negative attitudes expressed about these procedures by research oriented psychologists. It now appears that a discernible shift away from the use of projective techniques in clinical settings has taken place (Lubin, Wallis and Paine, 1971; Sundberg, 1961). Although their data indicates a decline in usage, the Rorschach, the T.A.T., the Draw-a-Person, and the House-Tree-Person continue to maintain positions of prominence. For example, Brown and McQuire (1976) surveyed the use of different kinds of tests and found some shift away from projective techniques, but in general the list of the ten most used tests had not changed much in the past three decades. As indicated above, in the most recently published survey, Wade and Baker (1977) found that equal time was given to projective and objective techniques.

In the same survey Wade and Baker asked respondent clinicians to list the tests they would recommend that clinical students learn. Consistent with previous research, projective techniques were recommended with higher frequency than objective tests, the Rorschach and the T.A.T. being ranked first and second on the list of test trainees should master.

From the brief survey of the literature that addresses the question of the status of psychodiagnostic assessment it appears that despite a marginal decline in the time spent testing, assessment continues to be a frequent activity of clinical practitioners. Further, projective techniques have maintained their centrality in the clinician
assessment armamentarium. Far from being dead, reports of the death of psychological assessment are grossly exaggerated. The data clearly indicates that Directors of intern settings expect trainees to be skilled in test administration and interpretation, as do prospective employers. This was shown in a study by Levy and Fox (1975) the results of which clearly indicated that testing skills are still considered a necessary requirement for more than 90 per cent of the clinical positions advertised in the A.P.A. Employment Bulletin over the calendar year 1971-72. Again 84 per cent of these positions required that testing skills include the use of projective techniques.

In sum then, psychodiagnostic assessment is alive and well and constitutes a significant aspect of the clinical practitioner's professional activity. Has the research demonstrating the doubtful aspects of the assessment enterprise been in vain? We know that practitioners are fairly cognizant of the evidence which casts doubts on the reliability and validity of tests (particularly projectives (Wade and Baker, 1977)) and yet they continue to choose to use the techniques. Was Bertrand Russell wrong in his assertion that there is "...nothing so subversive as a new idea, that neither empire nor institution can indefinitely prevail against them?" Certainly many of our psychodiagnostic practices have prevailed against an onslaught of negative research data.

The crucial objective appears to be not so much to try and convince practitioners of the compelling nature of the data but instead to furnish them with a viable alternative model of assessment and a set of efficacious assessment techniques. A common argument for the continued use of traditional psychological tests is that they are the only economical and practical diagnostic tools available to the
private clinician and institution. In that sense, their efficacy is proven by default. In a different context, but addressing a similar problem, Kuhn (1970), has shown that no amount of negative evidence will precipitate the downfall of the dominant scientific paradigm unless a viable alternative exists and is readily available. This conjecture appears to capture the situation vis a vis diagnostic assessment.

However, even if a viable alternative did exist (many believe the behavioural assessment model does represent such an alternative — see e.g. Ciminero et al, 1977; Cone and Hawkins, 1977; Hersen and Bellack, 1976) the major obstacle to its introduction in the practical arena would be the role expectation for clinical psychologists. The nexus between test administration and the professional identity of the practising clinical psychologist cannot be over-estimated. The importance of this nexus for the rise and current status of psycho-diagnosis will be explored in the next chapter.
CHAPTER III: ON THE RISE OF PSYCHODIAGNOSTIC ASSESSMENT

From the discussion in the first chapter it is apparent that the rise and current status of psychodiagnostic procedures is not predicated on empirical data alone, if at all. Authors on each side of the dispute (e.g. Breger, 1968; Holt, 1967; Rapaport et al, 1968; Rosenwald, 1963) agree that historically the rise of psychodiagnostic assessment has been inextricably enmeshed with the emergence of clinical psychology as a profession and the concomitant struggle to establish professional identity. They point out that clinical assessment was a major vehicle for psychologists to gain entry to the medically dominated 'mental health' arena.

Regrettably, then as now, assessment was the poor relation to psychotherapy. Since it was possible to engage in one of these activities to the exclusion of the other and since psychotherapy was seen to be the exclusive domain of the psychiatrist, diagnostic assessment was 'left' to the psychologist. The reluctance of psychiatrists in the late 40's to allow psychologists to function as therapists left psychologists to perform the less prestigious auxiliary function. The second rate status of this role was probably exacerbated by the fact that the clinical psychologists' reports rarely influenced psychiatrists' treatment plans (Mintz, 1968; Moore, Bobbitt and Wildman, 1968). The current ambivalence about diagnostic assessment cannot be properly considered independently of the clinical psychologists' early attempts to repudiate and redefine the depreciated role and the historical rivalry between the professions of psychiatry and psychology (Franks, 1978; Yates, 1970).
The aim of this chapter is:

(a) to explore further the historically important psychosocial influences that set the occasion for and promoted the rise and widespread application of psychodiagnostic assessment. It will be argued that the early attempts by clinical psychologists to redefine the depreciated ascribed role led to the development of a 'proprietary dogma' (to use a George Kelly expression) that spawned an exclusive guild based largely on the mystification of testing and scientific pretense;

(b) to argue that early clinical psychologists inadvertantly exploited the ubiquitous dangers inherent in testing and measurement, i.e. reifying the variables that the tests measure. This resulted in their participation in an elaborate process of 'metaphor-to-myth' transformation (Sarbin, 1967). Although once established, myths characteristically tend to perpetuate themselves, the myths created during this period were sponsored and supported by claims of 'scientific' status; and

(c) to argue that the rise of psychodiagnostic assessment was not only characterised by myth generation which was reinforced by the false persuasion of scientific pretense, but that the associated 'pseudoscience' was of a magnitude to rival the notorious 'Lysenkoism' of Soviet genetics and agricultural science of the early 30's. This is a serious claim; for this reason Chapters IV and V are
devoted to examining the accusation that traditional psychodiagnostic assessment rests on two important sources of pseudoscience; namely, psychometric pseudoscience and Freudian pseudoscience.

PSYCHODIAGNOSIS AND THE ROLE OF CLINICAL PSYCHOLOGIST

The period of triumph for the psychodiagnostic enterprise covered the first decade following World War II. It was during this period that clinical psychology emerged as an occupational specialisation and universities adopted the scientist-professional, or Boulder model of training - leading to the PhD in clinical psychology (Rainey, 1950). The professional identity of clinical psychology was to a considerable extent rooted in assessment procedures that differentiated the psychologist from other 'mental health' professionals, notably psychiatrists. The centrality of testing in the clinical psychologist's professional activities is evidenced by Watson's (1951) description of work they customarily performed: mental testing (e.g. the assessment of intelligence); educational assessment (e.g. backwardness in scholastic attainment); personality assessment (particularly by the use of projective techniques); diagnostic testing (the differential assessment between schizophrenia and hysteria, for example); the assessment of intellectual deficits or deterioration of brain damage; vocational guidance and selection; psychotherapy and research.

The testing defined a unique domain of professional functioning wherein the psychologist was the unquestioned expert. According to Cleveland (1976) the clinical trainees' crowning ambition was to administer and interpret psychological tests, especially projective tests and more particularly the Rorschach. Projective tests and the
Rorschach above all supposedly provided the ultimate in penetrative insights into the person. The Rorschach inkblots were once thought (and apparently still are used by some) to constitute a 'royal road to the unconscious' eliciting the pattern of internal organisation, comparable with X-rays of the person's mind. They were accorded a hallowed status (see e.g. Frank, 1939). During this period the clinical psychologists report were to case conferences accorded the same deference as the pathologist's report in some case conferences, however, as we have seen, the reports rarely had an impact on treatment decisions.

**Clinical Psychologists as Pseudopsychiatrists**

The uncritical adoption of the role of diagnostician, with the attendant acceptance of the established tests without questioning the purpose of the exercise, when coupled with the need to communicate meaningfully with psychiatrists by means of a supposedly common language, led to assessment practices in line with the tenets of the medical model. Because of this Yates (1970) disparagingly refers to the role played by early clinical psychologists as pseudopsychiatrists. Be this as it may, the implications for psychodiagnosis of the extrapolation of the medical model which utilised the model but not its content, cannot be overemphasised. The implications of the adoption of the intraorganismic or quasi-medical model (Bandura, 1969) for psychodiagnostic processes will be fully discussed in a forthcoming Chapter (Chapter VI).

Initially, clinical psychologists used what tests were available to them; sometimes without the data to support their decisions. For example, the configuration analysis of the test data, as practised,
was not readily amenable to research corroboration, and the 'art' of psychological test interpretation overran its scientific verification (Ivnik, 1977). As a result of this and similar practices, clinical psychologists came perilously close to establishing a guild or magicians' union centred on the mystique of testing, particularly the use of the Rorschach. According to Albee (1970):

"...what could have been clinical psychologists' most mysterious and powerful magic - the Rorschach. Here was a procedure that fulfilled most of the criteria establishing the power of an independent profession. The Rorschach belonged to psychology. It had enormous face validity and credibility, and it required long and arduous supervised training before one could be expert in its mysteries and its mystique. Its powers were so great that all sorts of controls had to be placed around the preparation of those who would be using it. Twenty years ago, most of us believed in the art of the Rorschach. The public was fascinated by the magic we owned. Other professions viewed our magic with respect. But along came the hardnosed scientists, the measurement people, with their questions about reliability and validity, with their split-half technique and their demands for public demonstration of the value of our magic under strictly controlled conditions. As a consequence, the Rorschach is quietly disappearing from the professional psychology scene because of our professional sensitivity to the claims of science." (p.1075)

Even if Albee was correct in his conjecture that the 'Rorschach' is 'quietly disappearing', the attendant mystique of testing persists. The widely accepted notion that a psychological test is a device that the psychologist carries around with him/her with which s/he gathers certain data that are then interpreted in terms of special knowledge, remains as a central myth of psychodiagnostic practise.
The Psychological Alchemy of Assessment

The special mystique of tests is related, at least in a genealogical sense, to the techniques of horoscopic astrology and physiognomy (McReynolds, 1975), to the divination and fortune telling practiced by medieval alchemists (Bersoff, 1973), and to 'the decision by ordeal' as sponsored by the medical clergy (Mackay, 1932). It is easy to disparage unkindly the assessment technique of our predecessors and admire our psychometric sophistication. Today we consider that assessment techniques such as physiognomy, astrology and phrenology are in principle all totally invalid, although we may condescendingly concede that in the hands of insightful, discerning and high-status practitioners they may, at least on occasions, have been more efficacious than generally given credit for. William James (1890) recognised this and, using phrenology as an example, suggested that "there seems to be no doubt that phrenology, however little it satisfies our scientific curiosity about the function of different parts of the brain, may still be, in the hands of intelligent practitioners, a useful help in the art of reading character". (p.28)

What will the critics in the future conclude when they look back at our times? Will they agree with Bersoff's (1973) conjecture that the traditional approach to psychodiagnosis is little more than psychological alchemy? Will they endorse Blum's (1978) contention that the central ideas in psychological testing are fundamentally erroneous and that much of what has been accepted because of its apparent scientific 'verification' is little more than psuedoscience? Will they be equally condescending and suggest that "in the hands of an insightful ... etc."? The illusion of validity and the complex reciprocal influences exerted by the examiner and examinee that may
result in invalid procedures accurately predicting the future will be discussed in Chapter VIII with particular reference being made to illusory correlates, the 'Barnum effect' and self-fulfilling prophecies. Wagar's conjecture that "The ultimate function of prophecy is not to tell the future, but to make it" (1963, p.66) appears to apply equally to the ultimate function of most psychodiagnostic assessment.

Several contemporary critics have already reached their conclusion with respect to the 'hallowed' enterprise of intelligence testing. Scott (1977) for example, suggested that:

"In 25 years time psychologists will look back on IQ's as one of the quaintest beliefs of scientific history - like the belief that the earth was flat ... IQ has more to do with magical beliefs of primitive people than with science ..." (p.15)

and Kantor (1971) concurs, stating that "the concept of 'intelligence' was not generated by empirical observations but by theology and philosophy". (p.165) As we will see later, the metaphysical origin of a concept does not make it intrinsically unscientific. But we will also see that psychometricians have been more concerned with psychometric aesthetics of intelligence tests than in scientifically validating the concept of intelligence.

In sum, it should be apparent from the foregoing discussion that the early rivalry between psychologists and psychiatrists and the depreciated role ascribed to clinical psychologist created exigencies that did not favour the conservative scientific evaluation of their professional efficacy. Rather it promoted the uncritical acceptance and promotion of assessment instruments that enabled a distinct
proprietary area of expertise to be established. Face validity, credibility and exclusivity appeared to be more important than established scientific reliability, validity and utility. Most assessment instruments were validated by endorsement.

**PSYCHODIAGNOSIS AND MYTH MAKING**

The manner in which tests were used and interpreted by early clinical psychologists exploited an inherent danger in any system of measurement, namely, the danger of reifying the variable that the test measures. To reify a concept or function is 'to give it a name and presently to consider that the name represents a thing, and finally to believe that the thing so named somehow explains the performance of the function' (Hull, 1943, p.28). In effect Thorndike's dictum that if something exists it can be measured was easily and illegitimately transformed to read that if a measure is taken, something exists. With reference to intelligence testing this tendency was identified early by the vocal antagonist of testing, Lippmann:

"Because the results are expressed in numbers, it is easy to make the mistake of thinking that the intelligence test is a measure like a foot rule or a pair of scales ... But intelligence is not an abstraction like length and weight: it is an exceedingly complicated notion which nobody has yet succeeded in defining ... If the impression takes root that these tests really measure intelligence, that they constitute a sort of last judgement on the child's capacity, that they reveal scientifically his predetermined ability, then it would be a thousand times better if all the intelligence testers and all their questionnaires were sunk without warning in the Sargasso Sea." (Lippmann, 1922, 1976)

Another important aspect of the myth making process is the propensity of many psychologists to assume the validity of the concept under study, while setting out to demonstrate it. That is, rather
than asking first, whether mental illness or schizophrenia, or intelligence is a meaningful or real entity, many psychologists ask instead: 'How can we demonstrate that the mentally ill/schizophrenic/retardates are different from normals?' An illicit conclusion, therefore, is entertained (that mental illness/schizophrenia/retardation exists) before it has been adequately defined or empirically demonstrated.

**Psychological Myth Making**

Skinner (1953) and Sarbin (1967) describe the process that allows the professional as well as laypersons to accept the realness of something for which there may be no referent. This phenomenon is particularly common with respect to psychological 'characteristics' as Skinner (1953) has pointed out:

"Trait names visually begin as adjectives - 'intelligent', 'aggressive', 'disorganised', 'angry', 'introverted' and so on, but the almost inevitable linguistic result is that adjectives give birth to nouns. The things to which these nouns refer are then taken to be the active cause of the aspects. We begin with "intelligent behaviour", pass first to "behaviour which shows intelligence" and then to "behaviour which is the effect of intelligence" ... but at no point in such a series do we make contact with an event outside of the behaviour itself which justifies the claim of a causal connection." (p.202)

Psychologists in general and test makers in particular show a strong proclivity to give names or labels to the variables on which their instruments yield scores. The result is a plethora of psychological concepts afforded the mantle of reality. Even such well accepted concepts as "intelligence", "adjustment" and "anxiety" have been given, through the medium of psychological tests, a much stronger aura of fundamentality than they merit (McReynolds, 1971). Thus, despite the widespread lack of agreement on the meaning of concepts
like "mental retardation", "mental illness", "schizophrenia", etc. The myth created represents an almost universally shared conception about the nature of mental illness, etc.

Myths, being widely held beliefs that are not true (i.e. readily falsified) are major factors behind the inefficiency of therapy, breakdown of communication, and failure to cope with many modern problems in humane and efficacious ways. In psychological thinking and practice, they create a continuous barrier to innovation and change. In clinical psychology, the early adoption of the role of diagnostician and the struggle by psychologists to differentiate and legitimate their professional identity led to the continuance of the myth making and psychological alchemy of our ancestors (Bersoff, 1973; McReynolds, 1975a). Test constructors and administrators apparently forgot their historical antecedents and became more concerned with things like psychometric aesthetics. This concern often resulted in the neglect of validity. They were convinced by statistical norms and charts, by the statistical sophistication of their procedures that their tests were the instruments of intrapsychic detective work. By optimistically assuming the validity of the concept being measured they unwittingly facilitated the process of "metaphor-to-myth" transformation (Sarbin, 1967).

It should be restated that the test constructors never directly observe a measure intelligence, introversion or schizophrenia, or for that matter, intelligent behaviour, introverted behaviour or schizophrenic behaviour. The terms are simply metaphors chosen to connote certain assumed qualities of putative, invisible mental processes. More specifically, they infer that behaviour appears "as if" it were intelligent, introverted or schizophrenic. Too often, Sarbin (1967) suggests, the "as if" condition is forgotten and we are left with
intelligent behaviour etc. As a consequence, intelligence, introversion or schizophrenia become accepted as concrete intrapsychic states instead of the metaphors that they are.

It is at this point that Sarbin (1967) would say that the concept has progressed through a "metaphor-to-myth" transformation. This cognitive strategy takes the following form:

"Every metaphor contains a wealth of connotations, each connotation has the potential for manifold implications, and each implication is a directive to action. While metaphors are ordinarily used by people to facilitate communication, the peril is always at hand that people may be used by metaphors (Turbayne, 1960). Such a peril is actuated when the user of a metaphor ignores, forgets, or purposely drops syntactical modifiers, such as "as if", that denote the metaphor, and instead employs the word in literal fashion. To say "Jones is a saint" carries one set of implications, if we supply the tacit modifier ("it is as if Jones is a saint"), the sentence carries a radically different set of implications if the predicate is treated as literal. The effects of permanently ignoring the metaphoric property of a word, that is, of dropping the expressed or tacit modifiers, is to hypostatize an entity. Such hypostatization sets the stage for myth-making." (Sarbin, 1967, p.447)

Every generation has been victim of its own myths. When people believed that the liver was the primary "mental" organ (somewhat in the way that we think of the brain), hepatoscopy was patently sensible. When they believed that behavioural anomalies were external manifestations of evil spirit possession then the treatment was directed accordingly toward exorcising demons. All systems of assessment and treatment, be they astrological horoscopy or Rorschach interpretation, exorcism or psychoanalysis, imply an underlying theory of personality or model of causality. The methods of assessing and modifying
psychological phenomena therefore cannot be understood independently of the personality theory upon which they are based and which provide a supportive rationale for professional behaviour.

The pervading, and now classical assessment model—represented in the works of Galton (1907), Cattell (1890), Kelley (1923), Hull (1928) and Gulliksen (1950)—conceptualises the assessment procedure as a measuring enterprise and stresses the importance of accuracy of measurement. Essentially the aim of this approach is to measure a given attribute—intelligence, anxiety, introversion, and the like—of given individuals. Assessment consists in the main, of either:

(a) assigning the individual to given categories, or
(b) placing individuals at a specific point on a continuum.

Naturally, psychologists choose to measure attributes that they suspect will help them to understand and predict 'important' behaviour.

It is not surprising that by far the most significant class of individual differences addressed by developers of assessment scales, inventories and procedures includes those traits and characteristics which purportedly distinguish the better "adjusted" members of our society from their "neurotic", psychotic, "mentally retarded", or "criminal" neighbour (Goldberg, 1971). The pressure on psychologists to devise indicators of "mental illness", "mental retardation" or potential emotional breakdown came not only directly from the military and industrial organisations (see e.g. Watson, 1978) but indirectly from the growing "medicalisation" of deviance in Western societies. In this therapeutic-meliorist view of society and social deviance, the "mentally ill/retarded" (i.e. social deviants) form a special class of 'victims' who must, both for their own good and for the interests
of the community, be 'helped' - coercively and against their will, if necessary - by the healthy, and especially by 'scientifically' qualified 'helpers'. Thus the assessment focus was upon the individual as the entity to be diagnosed, studied and treated. Dysfunctional mental behaviour of individuals was to be observed, diagnosed, and categorised according to a classificatory system that was assumed to be descriptive of specific behaviour and prescriptive of a course of "treatment".

Invariably the resultant assessment practices and instruments ignored the individual-situation dialectic and socioeconomic factors, and focused almost exclusively upon uncovering the intraorganismic determinant of the problematic behaviour. The bulk of assessment research stemmed from this "essence" concept of personality or causality. That is, the idea that individuals have an essence, a soul, or a personality which stands behind actions and motivates them (Bakker, 1975). Thus assessment was in fact the search for one or other of the host of inimical psychodynamic forces, for example, energising tracts, prepotent needs, repressed wishes, latent tendencies, unresolved conflicts, etc. that purportedly cause misbehaviour. For according to the quasi-medical model, these hypothetical internal agents underlie "mental illness" and the deviant or maladjusted behaviour is construed as a symptom. Psychological "diseases", while not considered to be necessarily real entities, function analogously to the systematic and traumatic disorders in the medical model adopted by modern medicine (Buss, 1966; Clare, 1976). Assessment and treatment to be other than superficial and dangerous, must focus on the underlying psychic states or traits. Traditionally, the focus has been on identifying disorders and 'symptom' patterns that go together, along the lines practised in medicine (Draguns and
Phillips, 1971). Once the disorders were clearly described, it was assumed that their etiology would be apparent.

Since the turn of this century the dominant conceptions of the nature, cause and treatment of mental illness have been products of this quasi-medical framework. From the perspective of this model, mental illness is something that a person has, like cancer or heart disease. A person with mental illness exhibits peculiar, unusual, or problematic behaviour which is related to some underlying causes that reside in the human organism. The language and concepts of the quasi-medical model are unmistakably analogous to that of physical medicine, a fact which significantly influences the manner in which those afflicted with "mental disease" are "treated". Here lies the conceptual foundation and accompanying rhetoric for the establishment of the myth of mental illness (Foucault, 1965; Sarbin, 1967; Scheff, 1966; Spanos, 1978; Szasz, 1960).

The basic assumptions operative have been that the deviant individuals, because of their illness, are impaired or defective and thus ineffectual or incapable of exercising options. This assumption has led inevitably to an institutional device (the mental hospital) to provide compensatory help to inflicted persons. Once the deviant is hospitalised, this assumption has further dictated a set of practices within the institution that have become the subject of much recent criticism because of their authoritarianism (Holzberg, 1960), degradation (Sarbin, 1967), dehumanisation (Goffman, 1961) and illness-maintenance (Schwartz, 1960). The construct of "mental-illness" is not only a "myth" but, because of its intrinsic assumptions, is a vehicle for the devaluing of the humanness of those so labelled (Laing, 1967; Mitchell, 1973). It may thus facilitate the social
sanitation process and reinforces the normative structure by demonstrating to other participants in the system the negative consequences of ignoring the normative boundaries (an obvious example of the 'social sanitation process' will be given in Chapter IV with respect to the sterilisation of the "mental retarded") However the process is generally more covert and pervasive and not readily identified for what it is. The confusion of 'cure' with 'control' will be discussed further in Chapter VI.

PSYCHODIAGNOSIS AND PSEUDOSCIENCE

Test developers and administrators alike, failed to exercise an appropriate degree of conservatism in the evaluation of their evidence: caution and tentativeness (hallmarks of legitimate science) gave way to confidence, proselytism and false persuasion by scientific pretence (Braginsky and Braginsky, 1974; Blum, 1978). When one examines the history of psychometrics and psychological assessment one finds a succession of recognised leading scientists (e.g. Burt, 1948; Eysenck, 1971; Galton, 1907; Goddard, 1913; Jensen, 1969; Terman, 1916) boldly professing to having empirical verification for their ideas. Upon close inspection however, it turns out that the supporting evidence and argumentation contain errors, ambiguities and untenable implicit assumptions of such magnitude that no claims of scientific confirmation could possibly be warranted. Thus much of the persuasion of the psychometric/assessment movement was effected by establishing a pretence of scientific discovery and data. As Braginsky and Braginsky (1974) have indicated "In the scientific enterprise idolatry may be transformed into 'methodolatry', the worship of methods, 'quartopheria', 'testomania' and 'the bribe', publications of results leading to promotion, prestige, and so on." (1974, p.29)
With respect to psychological assessment this description appears apt (as will be shown in Chapter IV). However, this is not to suggest a conspiratorial plot, no doubt the proposers were, for the most part, sincere and believed their work to be genuinely 'scientific' but unfortunately were working within a 'defective' paradigm.

In making the claim that 'the period of triumph' and before was a period of myth-making and psuedoscience we should not harbour illusions about the nature and limits of science. Certainly we should remain sensitive to George Orwell's (1946) warning that science is a word used in variable ways and in most cases more or less dishonestly. The authority and high esteem accorded to science, should be analysed and the 'story book' model of science (see e.g. Barber, 1961; Merton, 1969; Mitroff, 1974) be rejected outright. Briefly, the story book model contends that:

(a) scientific knowledge is proven knowledge;
(b) scientific theories are rigorously derived from the facts; and
(c) scientific facts are based on what scientists see, hear, and touch etc. and are essentially value free and neutral.

Such a view is not only seriously mistaken (Mahoney, 1976; Weimer, 1976) but is not the way of science (Kuhn, 1970; Lakatos and Musgrove, 1970) and serves a politically conservative function (Andreski, 1972; Rose and Rose, 1971). The human element in science makes all science an inevitably fallible endeavour (Mahoney, 1976; Mitroff, 1974).
What is Science?

To even begin to explore the question of 'what is science?' as Ziman (1968) points out: "is almost as presumptuous as to try to state the meaning of life itself. The question has long been debated. Famous books have been devoted to it. It has been the theme of whole schools of philosophy. To give an account of all the answers with all the variations, would require a history of western thought. It is a daunting subject." (p.1-2) Notwithstanding Ziman's warning, it is important that we hold a tenable concept of what science is if we are to seriously evaluate the charges of pseudoscience directed at psychodiagnostic assessment. However, this is not an easy task and unfortunately only the briefest of discussions is possible here. It is worth noting that even the recent emergence of several philosophers of science of great stature (e.g. Bartley, 1962; Kuhn, 1962; Lakatos, 1970; Popper, 1959, 1963, 1972) has not led to a resolution of the 'demarcation issue'. That is, at this point in time no acceptable criterion has been established for the demarcation of science from non-science. Attempts to set science apart from non-science have been numerous and unanimously unsuccessful (Chalmers, 1976; Lakatos and Musgrove, 1970; Weimer, 1976). For example, in the early 30's the principle of verification had hardly been proposed (see e.g. Ayer, 1946; Carnap, 1939) before its inadequacies were recognised. Sir Karl Popper, who did so much to highlight the shortcomings of the verification criterion, proposed the notion of 'testability' as the salient demarcation criterion (Popper, 1959). For Popper the essential aspect of testability is the possibility of falsification. Counterintuitive as it may seem, negative results and
predictive failures have far-reaching logical implications, and positive results (successful prediction) have comparatively little information content. This is contrary to the popular practice of selectively publishing positive result manuscripts and emphasising these successes far more heavily in literature reviews (Meehl, 1967). The epistemological costs of this practice are themselves quite distressing (cf. Mahoney, 1977; Smart, 1964). Unfortunately, although the logic of falsification is very different from that of verification and Popper's philosophy has much intuitive and logical appeal, falsification as we shall see later (i.e. Chapter 4) also contains critical flaws (Chalmers, 1976; Latatos and Musgrove, 1970; Weimer, 1979).

A rather dramatically different approach to the demarcation issue is provided by the internationally reputed, epistemological anarchist Paul Feyerabend (1970). Feyerabend takes an iconoclastic stance by insisting, "when the mood takes him", that there is in fact nothing special about science and that it should be seen as one ideology or religion among many. That is, science is not intrinsically different from religion and, according to Feyerabend, has a hold over modern people, akin to the hold that Christianity had over earlier Western societies. Similarly, Barnes (1973, 1977) who takes Kuhn's (1970) ideas to their logical conclusion also claims that there are no independent criteria for distinguishing scientific beliefs from other belief systems. Scientific knowledge according to Barnes is nothing more than beliefs accepted by the scientific community and their acceptance does not necessarily make them 'correct' beliefs. This close relationship between religion and science has been identified by many writers (e.g. Bakan, 1967; Braginsky and Braginsky, 1973, Myrdul, 1969). Similarly, the close relationship between the role
played by religious leaders in days gone by, and that played by 'scientific' psychiatrists and psychologists in contemporary society has been noted by Thomas Szasz (1971) and others (Foucault, 1973; Leifer, 1971). According to the Braginsky's:

"The priests' and rabbis' functions have largely been taken over by therapists, and those of the new prophets and latter-day saints by the researchers and theorists in psychology ... It is no accident, then, that psychology has flourished in these countries that abandoned either a national church or God, especially in times of social and political chaos ... Thus, psychology moved into the vacuum created by the absence of a religious institution that could be counted upon to effectively foster the interests of the state. Even in countries with a national church, historical accounts (see Foucault, 1965, Sarbin 1967, Merton, 1967) show that as the church weakened in its usefulness, the science of psychology and its ideological ally, psychiatry, gained in strength" (1973, p.30)

If it is difficult to erect even a single criterion for demarcating science from non-science, then obviously it is equally difficult to specify what constitutes legitimate scientific research which happens to have flaws, and research so deeply flawed that it becomes tainted by pseudoscience. The recent controversy surrounding the work of Sir Cyril Burt attests to this difficulty (see Eysenck, 1977; Gillie, 1978; Jensen, 1974; Kamin, 1974).

Scholars in nearly every field of endeavour (e.g. Kuhn, 1970, Myrdal, 1969; Polanyi, 1958) have begun to appreciate the role of the investigator's values, goals and beliefs as well as his/her creativity, in the acquisition of knowledge. We now know that perception does not mirror the external world or in Hanson's (1958) words "There is more to seeing than meets the eye ball" (p.7). Perception is necessarily an interpretation based on sensory information. Thus, there is no such thing as an ultimate sense datum.
Physical perception of the world is contingent: it depends on experience and context. An excellent example of this is provided by Michael Polanyi's description of the changes in a medical student's perceptual experience when s/he is taught to make a diagnosis by inspecting an x-ray picture.

"Think of a medical student attending a course in the x-ray diagnosis of pulmonary diseases. He watches, in a darkened room, shadowy traces on a fluorescent screen placed against a patient's chest, and hears the radiologist commenting to his assistants, in technical language, on the significant features of these shadows. At first, the student is completely puzzled. For he can see in the x-ray picture of a chest only the shadows of the heart and ribs, with a few spidery blotches between them. The experts seem to be romancing about figments of their imagination; he can see nothing that they are talking about. Then, as he goes on listening for a few weeks, looking carefully at ever-new pictures of different cases, a tentative understanding will dawn on him; he will gradually forget about the ribs and begin to see the lungs. And eventually, if he perseveres intelligently, a rich panorama of significant details will be revealed to him: of physiological variations and pathological changes, of scars, of chronic infections and signs of acute disease. He has entered a new world. He still sees only a fraction of what the experts can see, but the pictures are definitely making sense now and so do most of the comments made on them." (Polanyi, 1958; p.85)

It is now widely acknowledged that all scientific theorising and research is predicated on a priori assumption and, contrary to the popular adage in introductory psychology text books, our theories generate rather than reflect our data. That is, theories determine data, not visa versa: Indeed there would be no data without prior theoretical specification of what, in the flux of experience, constitutes significant observation. Facts, far from being the data upon which theories rest, are end products of theories (see Popper, 1959; Weimer, 1979). That experience is saturated with theory raises questions about Sir Francis Bacon's recommendation that "if you want to understand nature - look at nature, not at the works of Aristotle".
The dependence of perception on experience and context applies by analogy most directly to questions of the stability of scientific facts, data and evidence. Scientific facts are not independent of the interpretative schema by which they are apprehended. The status, implications and significance of a scientific fact may change dramatically when it is interpreted from a different viewpoint. Therefore the existence of a scientific fact necessarily implies a theoretical orientation; the difference between fact and theory is one of degree rather than of kind (Martin, 1979).

What is Pseudoscience?

Since all science is imbued with values and can be understood in its particular sociopolitical context, the label 'pseudoscience' becomes pertinent only when the biases displayed by scientists reach such extraordinary proportions that their relentless pursuit of verification leads to the commitment of major errors of reasoning and distortion of the consensually validated data. Paradoxically, assertions that the truth value of a proposition or hypothesis has been strengthened, proven or otherwise augmented by research is unfortunate and misleading since no empirical hypothesis can ever be scientifically confirmed (Mahoney, 1976; Weimer, 1976).

Sir Karl Popper has noted this for years, and yet scientists persist in their claim of confirmation (cf. Popper, 1959, 1963, 1972) rather than claiming corroboration (loosely speaking, corroboration refers to the act of having survived our most energetic and stringent attempts at falsification). This claim of empirical verification, proof or truth should be rejected on logical grounds and we should be wary of 'scientists' who make such claims.
Tautological fallacies, refusal to consider obvious alternatives, untestable assertions, and absolutistic and unconditional claims may be other hallmarks of pseudoscience. In contrast, philosophers of science suggest we should show considerable tolerance of statements conveying relativity and tentativeness, and that we should proceed with humility about our understanding of what it is that constitutes 'good scientific research'. Thus although we may acknowledge the difficulty of demarcation and the intrinsic fallibility of science, Gardner (1957) argues that the concept of pseudoscience remains useful. Writing in the defence of the concept Gardner states that "The fact that black shades into white through many shades of grey does not mean that the distinction between black and white is difficult" (1957, p.7). But as with the concepts of black and white, an understanding of what is science is culturally specific and, probably more importantly, today's science may well be tomorrow's alchemy (Mahoney, 1979).

A precise definition of pseudoscience is difficult because there are no clear criteria for demarcating between 'actual' pseudoscience and normal cases of fallibility by scientists. However 'pseudoscience' may be characterised as "a sustained process of false persuasion transacted by simulation or distortion of scientific inquiry and hypothesis testing" (Blum, 197 , p.145). According to Blum there are four parts to the definition:

1. essentially incorrect results are generated;
2. these are successfully and persuasively disseminated to a substantial audience;
3. dissemination occurs by a process of convincing the audience that results are bona fide scientific conclusions; and
4. the normal processes of error correction in science are retarded or prevented from functioning, so that the incorrect (readily falsifiable) beliefs generated are sustained over time.

If a precise definition of pseudoscience is difficult because there is no sharp dividing line between actual 'pseudoscience' and 'science' that is unfortunately and unintentionally flawed, cases that are sufficiently flagrant are apparently readily discriminable. Most identified instances of pseudoscience could be called 'petit' (Blum, 1978) in that the independent originators who profess eccentric theories do so without the benefit of significant institutional support. Proponents usually work in almost total isolation from their colleagues and are excluded from the scientific journals and societies which define them as quacks and cranks. A number of instances of this type are described in Martin Gardner's 'Fads and Fallacies in the Name of Science' (1957) and MacDougall's Hoaxes (1968). Although some practitioners of 'petit' pseudoscience have built up substantial cults, aside from their followers they are held in disrepute by the 'scientific establishment'. The contemporary state of psychotherapy seems replete with petit pseudoscience (Frank, 1973; Fish, 1974). Frank (1973) has provided an absorbing account of the common ground between faith healing, persuasion and so-called scientific psychotherapy.

Generally considered to be far more unusual, but perhaps more significant, is 'grand' pseudoscience (Blum, 1978). Here the domain of interest, rather than being clearly limited to a band of devotees, may stretch to an entire nation or group of nations, and the false beliefs are openly professed as valid science by respected authorities. If asked to point to an instance of pseudoscience on a grand scale no
doubt most Western Scientists would point to the dominance of Michuririst biology and the neo-Lamarckian genetics that dominated genetics and agricultural science in the Soviet Union in the early 30's. Referred to as Lysenkoism, this belief system was held to be sacrosanct for three decades. The chief results of Lysenkoism have been: an acceptance of neo-Lamarckian theory which posits the inheritance of acquired adaptive characteristics; persistent faith in a variety of bogus agricultural remedies; and a belief in the transformation of species. The compatibility of Lysenko's brand of genetics and Soviet political ideology need not be emphasised.

In the light of the earlier discussion of the role of the scientist's values, goals and beliefs in the acquisition of knowledge (e.g. Kuhn, 1970; Myrdal, 1969; Polanyi, 1958) it would be the height of naivety to think that 'grand' pseudoscience is not something found in 'free democratic' societies. In a paper entitled 'The Lysenko Syndrome in Western Social Science' Alex Carey (1977) argues that in fact this is demonstrably the case in both the natural and social sciences. Carey (1977) systematically demonstrates that several of our most cited 'classical' studies (e.g. the Hawthorne studies, Roethlisberger and Dickson, 1939; The Leadership Style Study, Lewin Lippitt and White, 1938; The Democratic Consultation Study, Coch and French, 1948) are so methodologically flawed as to render them uninterpretable and scientifically worthless. The popularity of these studies has more to do with the comfortable compatibility of their findings with the dominant cultural ideology (i.e. free enterprise capitalism) than with their research worth.

Carey points out that pseudoscience is most likely to develop in areas of pressing human concern where traditional and powerful institutions have explicit interest in maintaining the status quo:
For it is here that the possibility of offending against power and authority creates a special climate and set of contingencies that makes independent research difficult. When faced with the imminent prospects of clashing with powerful institutions and orthodoxy, Carey suggests, many sound scientists opt to work within the dominant ideology. Working with the implicit assumption that the broad lines of the status quo are sacrosanct, many social scientists abdicate their subversive role (i.e. creatively inventing and empirically testing new ideas) and become what Baritz (1958) refers to as a 'servant of power' - thus serving a politically conservative function by helping people adjust to the current social milieu.

Although Carey's (1977) catalogue of grand pseudoscience is not meant to be all inclusive, he fails to identify two important areas of pseudoscience that are of concern to us here. Recall Bersoff's answer to the question of who is to be held accountable for the psychological alchemy of assessment: '... the answer is two brands of "psychos", psychoanalysts and psychometricians. Psychoanalysts are to blame because they have perpetuated a fraudulent (Freudulent?) theory of personality and have perpetuated its myth. Psychometrists, the test constructors, are to blame because they have ... become overly concerned with psychometric aesthetics to the neglect of validity' (1973, p.892). In the next two chapters Bersoff's accusation will be evaluated in detail. In Chapter IV it will be argued that the 'scientific' testing movement operated within a fundamentally defective paradigm. The inevitable result was the generation of pseudoscience - in this case on a grand scale since much of the resultant research performed an important ideological service to the 'ruling elite' and thus won their patronage. It will be asserted that the psychometric 'intelligentsia' participated in pseudoscience on a scale that rivals
the notorious Lysenko affair of Soviet biology. Subsequently, in Chapter V it will be argued that Freudian theories also constitute an example of grand pseudoscience. Although, as a thinker and therapist Freud made an indelible impact on the minds of twentieth century people, it will be shown that pretence to science was a major contributing influence in the popular acceptance of his doctrine. The process of false persuasion by scientific pretence constitutes sufficient ground for labelling Freud's theories examples of grand pseudoscience.
Earlier it was pointed out that empirical science, being the systematic invention and testing of new ideas, ought to be incorrigibly subversive (Carey, 1972; Russell, 1953). Certainly the revolutions sparked off by Copernicus and Darwin attest to this characteristic of science. However, when it comes to the social sciences Carey (1977) for one, is skeptical about the fulfilment of this function. He claims the 'social scientists in the West (and also in Russia, for that matter) are, in fact, under very effective ideological control' (p.28). According to Carey (1972, 1977) and others (see e.g. Andreski, 1972; Mahoney, 1976; Rose and Rose, 1972; Winkler, 1975) the continually present possibility of asking questions that may offend against power and privilege creates a special environment around the social scientist that is not generally present for natural scientists. This environment encourages and reinforces role abdication.

Carey (1972) suggests that the risk of asking offensive/subversive questions is actively minimised by scientists adopting one of the following two strategies:

1. By confining their research interests to topics so remote, minute and esoteric that the possibility of results conflicting with the interests of powerful groups is effectively removed;

2. Or, by studying questions of social significance but from a vantage point which assumes, at least tacitly, that the broadlines of the existing order are given and sacrosanct.
If the second option is adopted, rather than becoming a potential subversive, the scientist becomes a 'trouble shooter' for the interests of orthodoxy; 'plumbers to the system' (Baritz, 1960; Whyte, 1956) or in the Braginsky's terms 'high priests of the ruling class' (Braginsky and Braginsky, 1973). Too close identification with the status quo presents relevant, penetrating questions from being asked and produces results that fit comfortably with the interests of the power elite (Andreski, 1972; Braginsky and Braginsky, 1974; Carey, 1972; Kaplan, 1964; Young, 1971, 1973).

In all societies, there are persons - Mannheim calls them the 'intelligentsia' - who take the status quo for granted and who provide the ideology which is the interpretation of the world from the perspective of the politically dominant groups. If the dominance of the elite groups in society is unchallenged, the intelligentsia producing 'knowledge' for those powerful groups enjoying complete control in moulding the society's world view. When a uniformity of viewpoint bounds intellectual discussions, there is a tendency for intellectual controversies to become increasingly 'scholastic' and esoteric - arising from the need for 'systematisation' of the belief system of the intelligentsia. (Mannheim, 1936, p.11) With respect to the 'scientific' testing movement Mannheim's observation seems apt. Up until twenty years ago a uniformity of viewpoint predominated and the literature evidenced scholasticism and a focus on statistical sophistication at the expense of addressing fundamental issues, such as the validity of the concept of 'general intelligence'. For example Eysenck describes those who challenge the 'genetic' explanation of IQ differences as being ignorant "alike of psychometric techniques of intelligence testing and biometric techniques of genetic analysis"
and finds it unreasonable to "discuss the problem or write about it, when one cannot tell the difference between epistasis and meiosis, reduce a Hessenberg matrix, or determine an Eigen-value" (Eysenck, 1972, p.6).

As we have seen in Chapter III, scientific knowledge incorporates values (Feyerabend, 1975; Young, 1971), and hence, is always selectively useful to certain groups (Martin, 1979). Attempts to develop value-free scientific knowledge are not only misguided, but results in the obscuring or ignoring of the values implicit in the knowledge. Rather than argue for, or strive to develop, value-free science the object should be to develop and promote scientific knowledge that incorporates preferred values, values which serve to promote desirable social and political structures. The natural science epistemology adopted by the 'scientific' testing movement obscured this important feature of science. Being portrayed as essentially ahistorical and acontextual this natural science paradigm (see Sampson, 1978) purportedly was used to describe the way things really 'were'.

Although all scientific knowledge is ostensibly concerned with the way things 'are' or could be, as opposed to the way they 'ought to be', Martin (1979) contends that "...to say something is is to say it ought to be" (p.79). In other words, every scientific and pseudoscientific statement is implicitly an imperative. That this is the case in the research literature on intelligence and intelligence testing will be evidenced in this chapter.

More specifically, this chapter is devoted to exploring the questions asked by the 'scientific' testing movement, particularly with respect to intelligence testing, and to supporting the
conjecture that the world view or paradigm within which the research was conducted, and which sanctioned the answers produced, was seriously defective. The result of adopting a defective paradigm it will be argued, is the inevitable generation of what we have labelled pseudoscience. This is the case because the central problems initially specified are themselves unsolvable given the available methods for seeking a solution. It will be shown that the history of intelligence testing is largely a chronicle of 'pseudoscience', with both defective arguments and false conclusions becoming institutionalised. Although worthless as 'science', much of the resultant research did perform an important ideological service to the ruling elite. It will be argued that the psychometric intelligentsia participated in 'Lysenkoism' on a grand scale.

**PSYCHOMETRIC PSEUDOSCIENCE**

Every mature scientific field has an accepted paradigm - a consensually agreed upon *modus operandi* that specifies certain questions as worthy of study and suggests methods of investigation which can lead to appropriate answers (Kuhn, 1970). Most scientific research does not attempt to challenge those basic assumptions underlying the speciality or discipline, but works within the assumptions, elaborating the paradigm to cover more detailed and more diverse evidence. Within any paradigm there will inevitably be a certain amount of evidence apparently incompatible with, or unexplained by, the paradigm. These anomalies do not cause the paradigm to be rejected and superseded immediately. Rather they usually stimulate more detailed research. At other times, they are assumed to be incorrect or irrelevant and are therefore ignored (Kuhn, 1970).
It was argued earlier that scientific knowledge is never unique, that there is always a large number of ways of explaining any feature of reality. The actual way of explaining reality that is chosen by a person or group will be suited more for some purposes and less for others. Thus all scientific paradigms are necessarily limited, but additionally they may be defective. That is, they specify questions without having available methods for scientifically validating the answers generated. Defective paradigms differ from superseded ones in that they never give rise to important scientific discoveries in their central area of concern. The superseded paradigm, while no longer seen as correct, at least at one time directed scientists to the solution of problems they viewed as important. Generally a superseded paradigm does not give rise to what would be labelled pseudoscience. In contrast, a defective paradigm is almost certain to generate pseudoscience because, to recapitulate, the central problems initially specified are themselves unsolvable given the available methods for seeking a solution. Consequently, the answers obtained are inherently suspect. When a defective paradigm is adopted by a community of scientists, individuals who subscribe to it continually face the dilemma of whether to admit failure and abandon the paradigm, or to bend the rules and adopt means normally prohibited in scientific inquiry. The more cautious scientists in this situation do not become eminent because they fail to provide solutions to the problems regarded as central. It was pointed out in the last chapter that many eminent spokespersons for the scientific testing movement boldly professed to having empirical verification for the ideas.
It is the contention of the author that the psychometric enterprise that has been directed at measuring differences between mental abilities, culminating in the widespread use of intelligence tests, is based on a seriously defective paradigm. Thus the leading scientific spokespersons for this enterprise have been in the business of promulgating a number of false solutions to major theoretical problems. Although their work has been criticized early in the piece (Haldane, 1938; Lippmann, 1976; Simon, 1953), social conditions were sufficiently auspicious for them to retain important academic posts, attract expensive funding and recruit a sizeable number of disciples, assistants and successors to maintain the paradigm. (These assertions will be supported below). Their work, while almost worthless as science, did perform important ideological services for the ruling elite, and this was sufficient to ensure the field's survival and growth (Henderson, 1976; Karier, 1972; Simon, 1971).

Intelligence Testing

Contemporary pseudoscience in psychometrics (e.g. Eysenck, 1971; Herrnstein, 1971; Jensen, 1972; Shockley, 1972) can only be fully understood by examining the historical development of our understanding and attempts to measure mental ability - specifically intelligence. Intelligence is selected because of its almost unconditional acceptance and success, if one judges success by customary criteria of size, influence and profitability of the enterprise. Intelligence tests have a tremendous and pervasive power over the lives of young people; directly by stamping some of them 'qualified' and others 'less qualified' for tertiary study, and indirectly by acting to control social mobility through defining as worthy of high esteem (i.e. to
define as intelligent) those skills that are held by one group in comparison with the skills held by others (Chomsky, 1972; Henderson, 1976; Rose and Hambley and Haywood, 1973).

As students of psychology know the first usable intelligence test was developed in France by Alfred Binet and Theodore Simon in 1905. Binet had been asked by the French Commissioner for Instruction to develop a testing procedure that would identify those students whose academic aptitude was so low that they might benefit from placement in special schools. What is not widely known is that Binet and his associates' approach to the task was essentially trial-and-error, to see what worked. They began by considering the kinds of anthropometric measures used earlier by Galton (1883) and Cattell (1890). When these proved unpredictive they next examined an extensive collection of photographs to determine whether facial features revealed probable success or failure in school. Again the results were negative; so were the results of the Parisian palmist hired to examine the hands of a hundred boys. What Binet and his colleagues eventually found to be the most predictive was performance on simple tasks like those they encountered in school. The performance on these test items could predict with some accuracy how children would do in school. Tasks appropriate for children aged from three to twelve were compiled, and these composed the Binet-Simon intelligence test scale published in 1908.

This brief history, besides evidencing the complete absence of an explicit 'scientific' theory of intelligence underpinning Binet's research, highlights the presence of important a priori notions concerning 'backwardness', 'feeblemindedness', 'mental retardation'. This critical aspect of test development is rarely acknowledged or the implications discussed, yet it is crucial to an understanding of the
resultant pseudoscience. Who decides the criteria by which to judge intelligence will be discussed in more detail a little later.

The original Binet-Simon test was considered a practical diagnostic instrument; Binet never intended to make a distinction between 'acquired and congenital feeblemindedness'. Certainly Binet called his product an 'intelligence test' but his concept of intelligence differed greatly from that proposed by spokespersons for the then popular eugenics movement based on Social Darwinism (e.g. Galton, 1892; Spencer, 1897). To eugenicists of that time intelligence denoted a relatively fixed, hereditary trait of individuals which was largely responsible for success or failure in most important spheres of life. In contrast, for Binet intelligence was a way of describing behaviour at a particular time in a particular setting. Scores on the test were seen as measures of how children had adapted to schooling - whether they were more advanced or less advanced than peers of their age. It did not explain developmental retardation.

In retrospect, the invention of the intelligence test constituted only a modest advance for the science of educational psychology. The test could predict fairly well how children would do in school but gave little information as to the reasons for the success. At the same time, the advent of the intelligence test constituted a tremendous, revolutionary advance for the development of Galtonian eugenics. The test was adopted as a scientifically legitimate method of measuring general intelligence - the basis of successful achievement in life. Prior to the advent of Binet's scales, Galton's contention that 'genius', 'mediocrity' and 'imbecility' and various categories between, were analogies in their statistical distribution
through the population to certain physical characteristics of that population, could not be scientifically advanced as none of the attributes measured in Galton's laboratory correlated positively with measures of 'genius', 'mediocrity' etc. Binet's test brought together eugenics advocates and key leaders of the testing movement under the name of 'scientific' testing.

The Transformation of Binet's Ideas

Binet's conception of intelligence was almost unrecognisably transformed when imported into the United States. When the intelligence test was introduced into the United States, it was written in English language and designed to predict which persons would be most likely to succeed in American Institutions, primarily schools. New migrants, non-Anglos, and persons from lower classes made significantly lower scores on these tests. Leading spokespersons for 'scientific' testing (e.g. Terman, Goddard and Yerke) interpreted these lower scores to be the results of inferior heredity (Cronbach, 1975) and became active in the political movement to ensure continued 'Anglo' dominance through restrictive immigration laws and sterilisation of those found 'unfit' by their 'mental' tests. (Kamin, 1974, pp.5-72).

The major importers and promoters of Binet's intelligence test were either eugenicists or eugenic sympathisers (Marks, 1973). The three well known early American translations of the Binet-Simon test were authored by Goddard, Kahlmann and Terman. Interestingly, all three had studied with G. Stanley Hall at Clark University where Hall had professed a theory of inherited mental ability similar to Galton's.
Goddard introduced the test first, in 1908, and became a leading hereditarian spokesperson almost immediately. His work at a training school for the retarded gave him expert credentials in the area of feeblemindedness and carried him a place on the Davenport promoted (Davenport, 1911) and Carnegie Institute of Washington-sponsored Eugenics Record Office 'Committee on the Heredity of the Feeble Minded'. He later wrote the infamous 'The Kallikak Family' (Goddard, 1912), a popular book which purported to demonstrate the all powerful influence of 'good and bad' heredity. Goddard's next book 'Feeble-Mindedness - its Causes and Consequence' gave further 'scientific' justification to the then popular notion of the relationship between feeble-mindedness and moral character (Goddard, 1914).

Terman probably had a more significant impact on the course of intelligence testing by authoring the influential Stanford-Binet revision which became the prototype of most subsequent intelligence tests. He wrote that "of the founders of modern psychology, my greatest admiration is for Galton" (quoted in Haller, 1963) and as an index of his admiration Terman published a study in which he estimated Galton's IQ at around 200 (Terman, 1979). The close nexus between the Galton-Spearman eugenics movement and the burgeoning intelligence test movement was voiced by Terman:

"Intelligence is chiefly a matter of native endowment. It depends upon physical and chemical properties of the cerebral cortex which like other physical traits, are subject to the laws of heredity. In fact the mathematical coefficients of family resemblances in mental traits, particularly intelligence has been found to be almost exactly the same as for such physical traits as height, weight, cephalic index, etc. ... The attempt to explain familiar resemblances by any other hypotheses than that of heredity have not been successful. All available facts that science has to offer support the Galtonian theory that mental abilities are chiefly a matter of original endowment." (Terman, 1975, p.201).
Besides his influential academic career, Terman was a leading member of the Human Betterment Foundation which promoted widespread sterilisation in California. Californian sterilisation law was based on race purification as well as criminology. Those who were 'morally and sexually depraved' could be sterilised. Between 1907 and 1928 a total of 6200 Californian residents were sterilised (see Karier, 1972).

Intelligence testing in the United States was given an invaluable boost by the decision to screen army recruits in 1917. Robert Yerkes, the psychologist chiefly responsible for convincing army officials to use the intelligence test was both president of the APA and a member of the Eugenics Records Office - he and Thorndike were members of the Committee on 'Inheritance of Mental Traits'. To construct the army tests Yerkes recruited Terman, Goddard and several other psychometricians.

Two other important early researchers, Edward Thorndike and R.S. Woodworth who pioneered the use of twin studies to estimate the heritability of intelligence were both students of J. McKeen Cattell who in turn had been Galton's research assistant. An indication of psychologists' respect for eugenics is the fact that Hall, Cattell, Yerkes, Terman, Thorndike and Woodworth all became presidents of the APA.

In Britain, a similar story but on a less grand scale can be told. Intelligence tests were quickly adopted by eugenicists - most notably by the social psychologist William McDougall and by his student Cyril Burt and by Karl Pearson. Burt in 1913 became the country's first psychologist appointed to a local authority with
special reference to educational matters. Burt went on to become the leading spokesperson for eugenics, and was instrumental in the introduction of the 11 plus examination. Britain is now just breaking free from an educational system that found its rationale, or justification, precisely in theories about human capacities derived from mental testing. In the nineteen thirties and forties, when the system was being constructed, psychometricians maintained that intellectual capacity was wholly due to genetic endowment and was therefore fixed, unchanging and, in addition, accurately measurable by group intelligence tests. This was the advice tendered to, and accepted by, the Consultative Committee to the Board of Education whose series of reports (Hadow, 1926; Spens, 1938) underlay the structuring of the divided system of secondary education with its concomitant, streaming in the primary school and early selection at 11 plus for secondary education.

However, toward the end of his career Burt lamented the decline of interest in his field:

"With the advent of behaviourism, investigations (in eugenics) virtually ceased. Even to ask such questions is to incur the old jibe that we are 'raising the discredited problem of nature versus nurture'. Nevertheless, let us hope a small band of enthusiasts will still come forward to explore this urgent and fascinating field of research". (Burt, 1968, p.18)

Unbeknown to Burt, the American education psychologist, Arthur Jensen reported that a lecture of Burt's "... was 'impressive indeed' and probably the best lecture (he) ever heard" (Jensen, 1972, p.8). Burt's lament may have been premature as Jensen's subsequent research interests attest. Early in 1969 Arthur Jensen published his now famous article, restating original nativistic views, in the Harvard Educational Review. In the American context this had relevance,
particularly, to theories concerning racial differences in 'Intelligence'. Jensen's views were propagated in Britain more specifically in the social class/educational context, as exemplified, in particular, by the contributions of Burt and others to the series of Black Papers on education from 1970 on. Intelligence testing is, therefore, once more, an important theoretical question in Britain and America and one having serious implications in terms of social policy.

In sum, the British and American early psychometricians, in effect, grafted Galton's theory onto Binet's instruments and by doing so, they transformed intelligence testing into a major form of pseudoscience. They successfully made it appear that the test measured intelligence as defined by Galton and Spencer, when the evidence never justified such a conjecture and Binet himself was not happy with this state of affairs; as late as 1909, he was challenging and criticising the misuse of his test by psychologists. Unfortunately, Binet died in 1911 without stemming the abuse to which his test was being subjected.

Thus the evolving paradigm embraced an array of false notions generated by efforts to construct a scientifically legitimate base for the practise of eugenics. Some of the more salient assumptions of the paradigm were:

(a) that mental capacity or intelligence is something inside the individual which has mediating, causal properties;
(b) that individuals differ in the amount (how much) mental capacity they have;

(c) that mental capacity can be inferred from performance on an intelligence test;

(d) that differences in intelligence are in part genetically determined (and the extent of this influence can be calculated);

(e) that by comparing test results one can determine the different capacities of various races, social class and ethnic groups and lastly,

(f) that intelligence tests measure the mental abilities needed for success in high level occupations.

**Innate Difference in Intelligence**

Terman's claim that differences in intelligence were innate and physiological, rested on an elaborate analogy with physical traits. He gathered and presented data in a way which made intelligence appear to be shaped by the same process as height. Test scores were reported in units of 'mental age' which was a kind of mental status. Like a child's height, his/her mental age grew steadily until s/he was around sixteen, and then growth slowed down and stopped. Thereupon it remained constant through most of adulthood. Like measurement of height, IQ's approximated the bell-shape curve of normal distribution when examined in large numbers. But most important, a child's IQ remained roughly constant as s/he passed through school. This was to be expected since children who
were taller in one year usually would also be taller two or three years hence. The approximate constancy of the IQ gave the impression that it could not be altered by the environment, and that it was therefore shaped almost exclusively by heredity. Finally, there was Pearson's correlation showing that resemblance between siblings and between parents and children was almost the same for IQ as for height. All this taken together, Terman reasoned, was determined in the same way and to the same extent as the physical trait, e.g. height.

In retrospect it appears that this similarity was a kind of mirage. The normal distribution in and of itself means nothing (see e.g. Lippmann, 1976; Simon, 1968). The stoppage of mental growth at around sixteen was an illusion generated by the selection of test items appropriate for school children. The much-touted constancy of IQ scores turned out to be a fallacious argument since again it was largely a product of the manner in which test items were selected. The constancy has been discovered only for children already in school and only appeared after a year or more of formal schooling (Hunt, 1961; Sharp, 1972). Terman's belief that he had substantiated Galton's theory was an illusion which he projected onto his own data.

Although Terman agreed that his test did not measure native intelligence precisely, he insisted that 'nearly all the psychologists believed that native ability counts very heavily' in determining scores (1916, p.311). This was a considerable distortion of Binet's insistence that:

(a) intelligence was not fixed biologically;
(b) mental orthopaedics be provided for children identified as developmentally retarded; and
(c) the resultant test score did not index mental capacity,
and was perpetuated by both American and British psychometricians (Kamin, 1974, p.5). The transformation rendered the assessment instrument vulnerable to abuse.

The abuse of intelligence tests is not primarily a result of misuse by unprofessional practice, as proponents would have us believe (see e.g., Simon, 1953) but is intrinsic to the paradigm which encourages hypostatization, isolation and measurement of acontextual mental traits and ignores the historical-social setting. Nowhere in the history of intelligence testing has intelligence or mental capacity been independently validated, certainly not measured. Intelligence is not something which one has, but is instead an interpersonal judgement about the power of the way one does something (Fischer, 1973, p.16). Intelligence refers to the effectiveness, relative to age peers, of the individual's approaches to situations in which competence is highly regarded by the society (Fischer, 1969, p.665). Note intelligent behaviour always occurs in a concrete situation, must be judged to be effective, and the competence evidenced must be highly regarded. Thus judgements as to what constitutes intelligent behaviour presuppose valued ends - effectiveness in accomplishing what and toward what further goal? In the established 'scientized' literature on intelligence the question of 'who defines?' is seldom, if ever, addressed (i.e. where do the a priori criteria for establishing genius and retardation come from). Who select the criteria by which to judge intelligence? That is, what is the social-political basis of the concept of intelligence?
Who Defines?

Recall that the Binet scales were first validated against teacher judgements - an external criterion. Operating from an implicit Anglo-conformity perspective, the psychometric intelligentsia saw as self-evident that school and vocational success in the 'Angle' core culture were the only socially important criterion measures needed for external validation. But as has been argued earlier:

"The question, what is it really?" "What is its right name?" is a nonsense question ... one that is not capable of being answered ... The individual object or event we are naming, of course, has no name and belongs to no class until we put it in one ... What we call things and where we draw the line between one class of things and another depends upon the interest we have and the purpose of the classification." (Hayakawa, 1957, pp115-116)

The question "who are the persons in our community who are really geniuses/mental retardates etc.?" is a nonsense question. The line we draw between one class of things and another is always socially arbitrated and depends upon our interests and the purpose of classification. Parenthetically, the difficulty in defining science and pseudoscience attests to the validity of this assertion.

Take, for example, Galton's 'Hereditary Genius': Where did Galton draw the line between geniuses and others? What was his purpose? In his famous study (Galton, 1892) a sample of nine hundred and ninety seven eminent men (were there no female geniuses?) were drawn wholly from the established upper and upper-middle class strata. Precluded from the sample were the 'captains of industry and commerce' who had, for the most part, no background in the established class. Thus 'eminence' for Galton was eminence within a certain specific range of activities. Galton, himself the nephew of Charles Darwin, and of
definite upper-middle class background, was faithfully reproducing the judgements and biases of his own stratum as to what constitutes 'true eminence'. A similar comment can be made of the work of Charles Spearman (1904).

Who are the mentally retarded and why do we draw a line between 'them' and us? Do we seek to help or control by classifying. Fernald (1912), a pioneer in the field, was unabashedly explicit in his description of 'them':

"...the feeble-minded are a parasitic, predatory class, never capable of self-support or of managing their own affairs. The great majority ultimately become public charges in some form. They cause unutterable sorrow at home and are a menace and danger to the community. Feeble-minded women are almost invariably immoral. We have only begun to understand the importance of pauperism, crime and other social problems ... Every feeble-minded person, especially the high grade imbecile, is a potential criminal, needing only the proper environment and opportunity for the development and expression of his criminal tendencies. The unrecognised imbecile is a most dangerous element in the community." (pp.90-91)

Intelligence Tests and Social Sanitation

Unfortunately Binet's test was seen as a scientifically validated and therefore unbiased way of recognising the undetected feeble-minded. Certainly Terman was excited by the prospective application of the intelligence test. "... in the near future intelligence tests will bring tens of thousands of these high-grade defectives under the surveillance and protection of society. This will ultimately result in curtailing the reproduction of feeble-mindedness and in the elimination of an enormous amount of crime, pauperism and industrial inefficiency. It is hardly necessary to emphasise that the high-grade cases of the type now so frequently overlooked are precisely the ones whose guardianship it is most important for the state to assume." (Terman, 1916, p.6-7)
There seemed to be little doubt in Terman's mind as to who the defectives were likely to be. He found that a low level of intelligence (IQ's in the 70 to 80 range) "is very, very common among Spanish-Indian and Mexican families of the South West and also among Negroes. Their dullness seems to be racial, or at least inherent in the family stocks from which they come." (1916, pp.91-92) Because of the lack of mental ability Terman argued that the feeble-minded were incapable of moral judgements and, therefore, could only be viewed as potential criminals. He said

"All feeble-minded are at least potential criminals. That every feeble-minded woman is a potential prostitute would hardly be disputed by anyone. Moral judgement, like business judgement, social judgement or any other kind of higher thought process, is a function of intelligence." (1916, p.11)

The same thinking which guided Terman to find a lower morality among those of lesser intelligence had its mirror image in the work of Thorndike, who found a higher morality among those with greater intelligence. Thorndike was convinced that "To him that hath a superior intellect is given also on the average a superior character." (1920, p.233)

Szasz (1976) warned us that we should always look to the relationship between the explainer and explained (In this case tester and testee). It is patently obvious that the explainer/tester in the above questions holds the explaine in contempt and metaphorically confines her by means of degrading imagery. It is easy to be shocked and chilled by the overt racism of the early advocates of scientific testing but since the paradigm that legitimated the hypothesis formation and testing that generated the false beliefs remains largely unchanged, it is feasible to suggest that similar
racist answers are being given to contemporary questions (no doubt in a more sophisticated form). Evidence suggests that this is the case in our 'scientific' response to our 'surplus people'.

(Braginsky and Braginsky, 1971; Braginský and Braginsky and Ring, 1969; Mercer, 1973; Scheff, 1966). The naive, apolitical view of the world and science has been expressed by a well known psychologist as follows: "I don't see why people should be disturbed by unequal representation of different groups in different occupations or educationally, if it should be found that there are real differences."

(Neary, 1970). As Smith (1978) has recently commented, echoing Rychlak's (1968) earlier analysis, although the positive and natural science epistemology on which mainstream psychology rested has been essentially discredited, Neary's remark and even a cursory glance at the content of psychological journals is enough to indicate it remains the prepotent perspective. Psychologists have been slow to explore alternative epistemologies.

Psychologists have been equally slow to learn that 'differences are not deficits' (Dobzhansky, 1973)

"That is, what biology has shown is that all people are genetically diverse, every human being is genetically unique, an unrepeatable individual. It can not be reiterated too often that equality and inequality are sociological, similarity and diversity are biological concepts. A society can grant equality of opportunity to its members, or it can withhold such equality; genetic diversity is biologically given, and could not be stamped out even if this were desirable."

(Dobzhansky, 1968, p.138)

The confusion of differences and deficits can lead to, albeit well intentioned, institutional racism (Baratz and Baratz, 1970; Edwards and Hargraves, 1976).
That is, people are different, not intrinsically superior or inferior; what people are called and where we draw the line between one group of people and another is socially arbitrated and then validated and always serves a purpose. The concept of intelligence has served an ideological purpose since illegitimately grafted on to the eugenics paradigm. The definition of intelligence in the sense of what is to be esteemed has been put forward by the dominant stratum in society to maintain its hegemony over social mobility (Henderson, 1976; Karier, 1972; Rose et al, 1973; Simon, 1971). The decision as to what behaviour will be the datum from which to infer intelligence arises from the power of certain sections of the ruling elite to define as worthy of high esteem those skills that they themselves hold relative to other groups. The dominant group have the power to define certain characteristics, which they already have (i.e. "white skin" or intelligence as they define it) as necessary criteria for entry to their social position. It becomes, from this point of view irrelevant to ask whether black skin or intelligence is genetically acquired or learned. The important and central question lies elsewhere, namely, who defines? Efforts to correct the abuses of the psychometric pseudoscience must aim at its foundation, not its content. For example, blacks will not improve their situation by continuing to demand black norms and testers for standardised intelligence tests. They must reject being judged as deficient by virtue of being different. They must look not to psychologists but politicians to rectify their devalued social situation.
Defining Intelligence

It may seem surprising to conclude that 'Psychology has no established literature on intelligence' but this is what Constance Fischer contends (Fischer, 1973, p.12). She goes on to say: "Our current scientific literature is about IQ, not intelligence. IQ is an artifact of the psychometrics movement" (Fischer, 1973, p.12). Test makers have not shown a great interest in intelligence, the psychometric elegance of their procedures being apparently more worthy of pursuit (Bersoff, 1973). How tests are constructed is interesting and even germane, but how they work is critical. Thus, other questions like "who defines?", "what is the validity of the 'general intelligence' concept?", and "what is the validity of the claim that intelligence tests measure 'intelligence'?" are too infrequently asked.

Unbelievably, psychometricians have no consensual agreed to definition of intelligence (see e.g. Weschler, 1975). They claim nevertheless that intelligence is a worthwhile scientific concept because it can be measured. If, however, the evidence indicates that no existing test can measure anything as broad as general intelligence, then the concept should be abandoned as scientifically meaningless at this time, since it can be neither defined nor measured. Intelligence is defined by Cleary et al (1975) as that behaviour labelled as intellectual by psychologists and exemplified in the Standford-Binet and Weschler tests. It is "the repertoire of acquired skills, knowledge, learning sets and generation tendencies considered intellectual in nature that are available at any one period in time." (p.19) The definition is essentially operational. Such definitions
are in no sense a definition in the scientific meaning of the word (Simon, 1971) and are maximally conclusive to the development of pseudoscience (Blum, 1978) in that they discourage investigators from questioning the validity of the concept.

The unstated assumption of most operational definitions is that intelligence tests measure a person's capacity for intellectual achievement and professional success. However, Jackson (1975) considers erroneous the "assumption that scores on these tests measure intellect when, in fact, what they might be measuring to some extent is the appropriateness of the subject response in relation to the cultural configuration of the white Anglo-Saxon middle class population that dominates American society" (p.89) Thus, the greater sociocultural distance between the individual and the dominant core culture, the lower his or her score will be (Mercer, 1973, 1975; Mercer and Brown, 1973).

The only way to validate the construct scientifically would have been to demonstrate consistent high positive correlations between IQ and reliable ratings of performance in a wide variety of artistic, scientific, and professional activities (the class bias is apparent here—why isn't criminal activity or union organizing included? For example, Williams (1974, p.34) argues that "actual intelligence covers a broad range of human abilities that IQ tests do not even attempt to measure. For example, no test has formally assessed the many verbal and non-verbal skills required to survive in the black community". Such validating data were rarely obtained, whereas a number of negative results have (see e.g., Blum, 1978 pp.69-84).
The Validity of the General Intelligence Concept

In line with the thesis being presented, McClelland concluded a recent review by stating - "it is my contention that the evidence for their validity [intelligence tests] is by no means so overwhelming as most of us, rather unthinkingly, had come to think it was. In point of fact, most of us just believed the result that testers gave us" (1973, p.46). In fact there has never been strong evidence that intelligence tests measure intelligence, or in psychometric parlance external validity of the tests has not been adequately established. External validity is concerned with prediction to real-world performance. It is usually expressed in terms of a correlation between test scores and actual performance in some social role. Claims that external validity had been established rested on two kinds of correlations:

(a) between IQ and performance in schools (i.e. role of student), and

(b) between IQ and occupational status (i.e. role of worker).

(a) Tests and Grades in School

Grades in school are determined primarily by performance on tests and examinations most of which are somewhat similar to intelligence tests. It is scarcely surprising that a positive correlation between IQ and school grades is found since they both depend on similar cognitive skills, stores of knowledge and motivational variables. However, the correlation would support the validity of the construct only if ability (grades)
in school provided a good index of ability in subsequent intellectual endeavours. Such outcomes have been documented carefully by many researchers (cf. Hoyt, 1965) - both in Britain (Hudson, 1960) and in the United States (Berg, 1970). Generally studies show that neither the amount of education nor grades in school are related to vocational success in a wide variety of occupations, including research scientist (Mahoney, 1976).

Using school performance as the external criterion to validate tests was first seriously questioned by Eells and Davis (Eells, Davis, Havinghurst, Herrick and Tyler, 1951). They protested that existing tests underestimated the abilities of lower-class children. Cronbach (1975) accurately perceived the controversy as an early attempt by a politically 'rising' group to assert their definition of the situation as opposed to the definitions of the intelligentsia. In the light of the discussion thus far it is interesting to note Cronbach's evaluation - "The Davis campaign failed for several reasons. He challenged the testers when they were in public favour. He concerned himself with 'persons of low status' and there were no militant voices to take up that cause. And if it were true that his new tests identified potentially able children for whom someone ought to invent better schooling, that advice was too abstract for public debate or action" (1975, p.8). So much for 'let the data talk'.

It was pointed out in Chapter III that there is always a large number of ways of explaining any feature of reality. This is graphically evidenced by comparing two diametrically different
views of schools and intelligence testing. For example, Galton (1907) saw schools as proving grounds where natural ability would inevitably lead to different performance. Henderson (1976) saw schools of the same period as the principal arena 'in which the class contest' was fought over access to position of high status in wider society. For Galton measured intelligence explained success whereas for Henderson the intelligence test constituted a powerful weapon used to justify and perpetuate the existing stratification system.

(b) Test and Occupational Success

Claims that intelligence tests measure capacity for competence in different level occupations rest similarly on tenuous ground. Correlates between IQ and occupational status could be interpreted in several ways. In general, scientific testers embraced the meritocracy interpretation which attributes the correlation to different intellectual demands of different status jobs. This position, stated broadly, maintains that certain people in our society, because they are intelligent, come to find themselves in privileged positions and the positions are privileged or disproportionately remunerated because of the scarcity of such people. It seems equally possible that it is because of the privileged position that some people are 'intelligent'. That is, certain groups within society are able to restrict entry to their positions, and thus to make or keep their ability 'scarce'.

The basic problem with many job proficiency measures for validating intelligence tests is that they depend largely on the credentials the person brings to the job - the habits, values,
accent, interests, etc. - that means s/he is acceptable to the employing agent. Since we know that social-class background is related to getting higher-test scores (Nattall and Fozard, 1970), as well as having the right personal credentials for success, the correlation between intelligence tests scores and job success often may be an artifact, the product of their joint association with class-status (McClelland, 1973).

Recall that the Binet scales were first validated against teacher judgements, an external criterion. Operating from an implicit Anglo-conformity perspective, the psychometric intelligentsia saw as self-evident that school and vocational success in the Anglo core culture were the only socially important criterion measures needed for external validation. For example, in their recent review Cleary et al. state that "intelligence and other ability tests are useful to the extent to which they are correlated with socially relevant and important criteria" (1975, p.23). Nowhere do they explain who decides which performances are 'significant', 'relevant' or 'important'. Such decisions are political, not scientific.

It is crucial to note that once the Binet scales were accepted as 'valid' subsequent tests were validated by correlating them with scores on the Binet scales. Such correlations are measures of internal validity. (Tests of internal validity have little to do with predicting real-life performance rather they focus on rationalising and systematising the knowledge system itself). During the long period in which the psychometric belief system was all but unchallenged, intellectual discourse became increasingly 'scholastic'. Studies of test validity were preoccupied with the internal systematisation of the psychometric belief system. Publications concerning test validity
focused primarily on measures of internal validity, factor analysis of test items, and so forth (see e.g. Buros, 1940-72). Relatively little energy was expended on problems relating to external validity, and what data was available was equivocal.

SCIENTIFIC PARADIGMS AND SOCIAL VALUES

Psychologists, like other scientists, have been, until recently, incredibly naive about the role of power and interest in social outcomes (Blackburn, 1972). Social scientists need to recognise the social substructure of values, beliefs, and ideologies implicit in their research and activity (Buss, 1975; Martin, 1979) and acknowledge the historical-social processes that generate 'knowledge'. In the case of intelligence testing powerful interests have provided the definitions, that is, controlled the criteria against which psychologists have validated their tests. Belonging to the power elite not only helps a young person get into tertiary institutions and get jobs through family contacts, but it also gives access as a child to the credentials that permit access to certain occupations. But over and above this, the power elite are able to select the criteria by which to judge intelligence and they do this 'in their own image' as it were. This does not imply any conscious manipulation or 'plot' on the part of the particular class. Any group in a relation of dominance to another will tend to set up entrance requirements that are in class correspondence with their own standards of acceptability. Notwithstanding this, McClelland (1973) concludes that it is very hard to find 'a good carefully controlled study of the problem (validity) because testers simply have not worked very hard on it. They have believed so much that they were measuring true competence
that they have not bothered to try to prove that they were'. (1973, p.56)

Within or Between Paradigm Debate

What is required is not necessarily more carefully controlled research but critical appraisal and rejection of the defective psychometric paradigm. That is, effort should be directed at exploring and evaluating the conceptual framework rather than its content. The legitimate work of criticism between paradigms consists of offering an alternative viewpoint which is an adequate representation of the phenomena and which is discontinuous in the sense of being incompatible with the other position (Overton, 1973). Kuhn (1970) has suggested the characteristic of this paradigm clash is that each side disputes the legitimacy of the central assumptions of the other and repudiates the meaningfulness of its approach. This has been the aim of this review, and although it would have been possible to marshal much more evidence, the risk of doing so is to unwittingly enter into a within paradigm debate: the risk being that such discourse inadvertently legitimates the contested paradigm by debating specific issues without challenging its assumptive base. Consequently the foregoing should be seen as a conceptual critique rather than a detailed empirical evaluation. As Mannheim pointed out, there is a multitude of realities and 'the problem is not how we might arrive at a non-perspectivistic picture but how, by juxtaposing the various points of view each perspective may be recognised as such as a new level of objectivity attained" (1936, p.226). A more detailed empirical evaluation of the 'attribute model'
of assessment of which intelligence testing is a prestigious example will be presented in Chapter VI.

In conclusion, during the half-century in which the psychometric intelligentsia have been unchallenged, a uniformity of viewpoint has bounded intellectual discussion concerning tests. The traditional paradigm and accompanying assumptions were regarded as self-evident. The resulting psychometric research and theorising represents an example of 'grand' pseudoscience. The major results of this movement have been a highly exaggerated notion of what intelligence tests measure; a widely subscribed-to belief that variations in mental abilities are determined largely by genetic differences, the belief, also widely held, that certain races are demonstrably more intelligent than others; and, a maintenance of the doctrine that social inequality is caused by variations in the biological fitness of individuals. These conjectures while put forward as scientifically validated facts are, on close examination far from that (Chomsky, 1972; Layzer, 1972; Kamin, 1974; Blum, 1978). Although strong supporters of the idea of rigorous science, proponents of psychometrics have departed from the rules of science. When valid objections have been raised against their work, they have sought repeatedly to circumvent and dismiss these objectives which could not be refuted. This pattern of apparently deliberate distortion has appeared in Terman's response to Lippmann (Terman, 1976) as well as more recently in Jensen's evasive use of the terms covariance and interaction, in the attempt to dismiss Kamin's critique, and in the refusal to cite studies with evidence contradicting the meritocracy interpretation (Jensen, 1978). More flagrant deception is found in the fabrications in Sir Cyril Burt's work (see e.g. Lawler, pp.159-173). Like Lysenkoism in Soviet
genetics, the British and American psychometricians had important patrons both from private funds (e.g. Galton and Mrs. E.H. Harrisson) as well as private foundations like the Carnegie Institution and the Commonwealth Fund. The importance of selective funding of scientific research programmes and the resultant development of certain types of paradigms has been highlighted by the Roses and others (Rose and Rose, 1974; Rose and Rose, 1971; Marcuse, 1967). The comfortable congruence between their research program and the status quo was sufficient for the chief proponents to retain important academic posts, attract expensive funding and recruit a sizeable number of disciples and successors to maintain the enterprise. The enterprise has not as yet been widely recognised for what it is - pseudoscience on a grand scale.

On the surface, the dispute over testing appears to be an academic discourse between two opposing factions of intellectuals. In fact, as has been demonstrated, the issue of test validity and test fairness deals with matters of critical importance in a society which is relying increasingly on the results of tests to make important life chance decisions and to decide ultimately who will achieve positions of wealth, power and prestige in Western society (Mercer, 1978-79). In fact, unbelievably, the first sentence in a recent publication reads "IQ is the yardstick by which your children will be judged throughout their lives". (Wilson and Grylls, 1977, p.1) Very little appears to have changed in the last five decades.
In the last chapter it was argued that the scientific enterprise should be intrinsically subversive (Carey, 1972; Russell, 1953). This is not to say that all scientific enquiry necessarily have immediate or important social consequences. The work of astronomers, geologists and physicists only rarely has significance that could conceivably bring them into conflict with the dominant values and ideology of their society. However, the furor and invective which arose when Copernicus banished the earth from its honorific position as the centre of the universe, or when Darwin dared to demonstrate our affinity through evolution, with wart-hogs and monkeys attests to the potentially revolutionary nature of even the natural sciences. It should be noted that as soon as the subversive import of these ideas was recognised the scientists concerned (namely, Copernicus, Galileo, Darwin and Huxley) found themselves in savage embattled confrontation with not only hostile social institutions but also with incredulous fellow scientists. For example, the Aristotelians refused to look through Galileo's telescope to see the four moons of Jupiter and Goss (a Christian biologist) in response to Darwin's theory, postulated that God had deliberately planted fossils so as to separate true believers from doubters. This explicit practice of isolationism was recently documented by Krantz (1971) in the area of behavioural psychology. In fact, in his autobiography, B.F. Skinner sounds almost proud of the fact that he never even bothered to read the criticism against him by Chomsky and Scriven.
Social scientists by contrast, are continuously concerned with questions about human nature and human behaviour, questions which, if answered one way, may challenge the existing system of values and privileges and, if answered another way support and strengthen this system. In the last chapter it was argued that the 'scientific' testing movement arrived at answers comfortably congruent with the power elites' perspectives and thus justified and strengthened the status quo through the practice of grand pseudoscience. (see e.g. Lawler, 1978; Simon, 1971) This accusation could certainly not be made with respect to Freudian pseudoscience.

Both as a thinker and therapist, Freud has had a remarkable, even revolutionary impact on the people of the twentieth century (see particularly Jacoby, 1975). His insights, like those of Copernicus and Darwin have penetrated the very underpinning of our culture and have become a part of us in far more ways than are usually realised. In keeping with the history of truly subversive science, Freud is often portrayed as having courageously proposed his theories in the face of violent public and scientific opposition and as having suffered personal and professional isolation as a consequence (see e.g. Brill, 1948; Jones, 1953). Certainly, Freud in his correspondence with Fliess complained about his lonely struggle against a world hostile to his ideas in the first decade of psychoanalysis (Freud, 1948). However, he also refers to the period as one of 'splendid isolation'.

Today, the accounts of the early opposition and hostility to Freud's ideas are seen to have been greatly exaggerated (see, e.g. Bry and Rifkin, 1962; Ellenberg, 1970); so much so that they completely obscured the fact that Freud was very much a 'man of his times' (Elkin, 1972; Whyte, 1962). Notwithstanding the many signs
of recognition and extraordinary respect accorded Freud, even in his early years, and the convincing demonstration that virtually every single idea of Freud's had been conceived by somebody else before him (see e.g. Ellenberg, 1970; Whyte, 1962) his revolutionary impact can not be underestimated or attenuated. But can the impact of Freud be legitimately compared with that of Copernicus or Darwin? Is the Freudian revolution an example of scientific subversion par excellence, or 'one of the greatest hoaxes of the century' (Sargent, 1964, p.89). Pinckney and Pinckney (1965) subscribe to the hoax conjecture and suggest that 'probably the biggest factor contributing to the psychoanalytic hoax is that analysts strive so deliberately to call their exploitation as a science without fulfilling any of the postulates set down and accepted by scientists the world over which would qualify it as being truly scientific' (p.70).

Critics, doubtful of the validity of psychoanalysis, have created the impression that Freud simply and arrogantly asserted the truth of his statements without bothering to consider the all important questions of validity (see e.g. Cioffi, 1970, 1974; Eysenck, 1963; Jurjevich, 1974). Such a view is not only naive but ill-informed. Nothing could be further from the facts. There is much evidence, both direct and indirect, in Freud's writing of his desire to discover ways in which the validity of psychoanalysis could be established (see particularly Jahoda, 1977, Chapter 7). Further, Freud was adamant that his theories were inferred from 'good, hard' clinical observations yielded by the application of a scientifically legitimate method. According to Freud psychoanalysis was a natural science, "what else could it be?" (Freud, 1925). There can be no doubt about Freud's desire to be scientific but there are grave doubts about the adequacy of his approach.
The aim of this chapter is not to challenge Freud's incontestable status as an innovator, nor to thoroughly evaluate the scientific validity of his ideas, rather it is to argue that his contribution should not have been accorded the 'status' of scientific knowledge; and that his pretense to science has been a major contributing factor to the popularity and pervasive impact of psychoanalysis. This process of false persuasion by scientific pretense has, in earlier chapters, been called pseudoscience. It will be argued that many professionally popular beliefs about personality and psychopathology are largely shaped by such a process.

THE IMPACT AND STATUS OF FREUD'S IDEAS

The impact of Freud's ideas has already been alluded to by comparing their social import with the ideas of Copernicus and Darwin. However, unlike the contributions of the great scientists, Freud's 'correct' place in our intellectual history is far from consensually established. Despite the passage of over three quarters of a century since the publication of 'The Interpretation of Dreams', consensus as to his status is conspicuously absent. Freud has astonishingly been accused of everything from immorality to unscientific dilettantism, yet his ideas remain the essential starting point for much theoretical and clinical innovation in Western psychiatry (Elkin, 1972; Jahoda, 1977).

That there is almost complete disagreement on Freud's scientific status has not apparently ameliorated his social significance. No one could seriously doubt the ubiquity of the influence of his ideas, not just within professional and academic psychiatry and psychology but as a basic part of our cultural substance. Writing in 1932
Franz Alexander summarised the situation well: "All these concepts are today not only generally accepted, but they have already become emotionally assimilated and like the theory of evolution or the cosmological doctrine of the planetary systems are now an integral part of modern thinking"(1932, p.65).

The Freudian omnipresence has all kinds of practical consequences. It not only influences the way in which we 'talk about' and respond to problematic behaviour but at a 'day-to-day' level it regulates the behaviour of parents with respect to their children for example. For years the spectacular mind-expanding nature of Freud's announcements have overshadowed the question of their 'scientific' validity for most people, professional and lay. It was as if a new psychological world had been discovered; people were fascinated with the novelty of its sights, rather than their 'reality'. The perspective's apparent efficacy as evidenced by its ability to explain everything, won many converts. In these terms, Freud's models have already proven themselves to be spectacularly successful. It is impossible to look into any major journal dealing with research in the area of personality or psychological dysfunction without finding important segments of his thought represented. His concepts relating to repression, unconscious motivation, defense mechanisms, and character development are everywhere imprinted. Overall, it is reasonable to say that Freudian concepts have already achieved the important objectives of opening up new phenomena for study and introducing additional ways of interpreting old issues.

Although there can be little dispute over the profundity of Freud's influence, much doubt remains as to our 'final' judgement regarding his intellectual contribution to a science of human
behaviour. That Freud has a prominent place is attested to by his social significance evidenced above but as Jahoda questions: "Where?" Jahoda (1977) points to the seemingly unending flood of publications that provide "contradictory answers: a scientist or a charlatan; the founder of a new psychology or a poet; a philosopher or a philosopher manque; a moralist or a libertarian; an original thinker or a clever propagandist of other people's ideas; an organiser of a movement or a lonely pioneer; a cure or a blessing for science and morality" (p.1). Given that the important, basic 'discoveries' of psychoanalysis were all made and published before 1906, and given the successful social penetration of his ideas, why haven't scholars been able to agree on Freud's rightful status?

Obstacles to the Accurate Evaluation of Freud's Contribution

In short, Freud and his followers must bear most of the responsibility for the manifest lack of clarity as to the scientific worth of their contribution. For unlike the proponents of other subversive belief systems (e.g. Darwinism and Marxism) who took an offensive, confrontational stance vis-a-vis their antagonists, in the case of Freud, the psychoanalytic movement took predominantly a defensive stance. Most Freudians restricted themselves to their professional practice and emerged from their 'sect-like' existence (Fisher and Greenberg, 1977) only occasionally to answer some new accusation. Their truly prolific literature (The Index of Psychoanalytic Writings (Grinstein, 1975) lists over 100,000 books, articles and monographs) is directed primarily to other psychoanalysts (Jahoda, 1977).
A further index of their defensive posture was their recourse, in the early stages, to argue that:

(a) any opposition to their ideas was a form of resistance - a mechanism of defense against the unpalatable truth; and

(b) a true test of psychoanalytic ideas could be made only by those qualified to practice psychoanalysis.

As Medawar (1969) has recently pointed out, such postures are contrary to all the canons of scientific investigation and he consequently dismissed the whole theory as unscientific. He went on to suggest "The opinion is gaining ground that doctrinaire psychoanalytic theory is the most stupendous intellectual confidence trick of the twentieth century and a terminal product as well - something akin to a dinosaur or zeppelin in the history of ideas, a vast structure of radically unsound design and with no prosperity" (Medawar, 1977, p.125). Although these defensive arguments have been thoroughly debunked, at least from the perspective of the natural science paradigm, (see e.g. Eysenck and Wilson, 1973; Fisher and Greenberg, 1977; Glover, 1952) and are rarely seriously advanced by psychoanalysts in response to criticism, their legacy is a lingering doubt about the willingness to 'play the science game'. Obviously there is no imperative on psychoanalysts to play the science game. There are other, equally legitimate ways to approach the understanding of our fellows. Understanding of an 'empathic' kind, called Verstehen may be more appropriate. This procedure which involves careful observation of the individual, combined with a putting oneself in his/her position may make valid the so-called defensive postures described above. This issue will be discussed further towards the end of this chapter.
A second important reason for the uncertain scientific status of the psychoanalytic system is Freud's occasional explicit assertion to the effect that 'psychoanalysis did not need outside validation'. This claim has recently been echoed by Kubie (1960) and Ramzy (1962, 1963) in their contention that each psychoanalytic hour is a miniature controlled experiment in its own right. Initially Freud made the claim cited above in response to a report of an instance of laboratory support for his concept of repression. Freud went on to say: "I can not put much value on these confirmations because the width of reliable observations on which these assertions rest makes them independent of experimental verification" (Freud, 1934 cited in Mackinnon and Dukes, 1962, p.702). Parenthetically it is interesting to note that a comprehensive recent review of the attempts to experimentally validate the Freudian concept of repression concluded that 'surprisingly, it was found that none of the investigations provided support for the predictions' (Holmes, 1974, p.632). This disparity illuminates the different logic and notions of validity that may be adopted in evaluating research.

In a similar vein, other experienced analysts have asserted that their complex concepts cannot be adequately studied within experimental guidelines because they force too much artificial simplification (see e.g. Rapaport, 1960). Thus the subject matter prohibits experimental investigation as usually practised. The scientific validity of these claims will be evaluated subsequently but it suffices to say that the 'defensive ploys', the explicit rejection of experimental evidence, and the rejection of experimental methodology has made it relatively easy for Eysenck to build a case to justify his iconoclastic conclusion that "...psychoanalysis is
a myth; a set of semi-religious beliefs disseminated by a group of people who should be regarded as prophets rather than scientists" (1963, p.32). It is probably timely to point out again that 'differences are not deficits'. The aim of this chapter is not to show that Freud's theories are 'wrong' but merely that they should not be labelled 'scientific' if we subscribe to the narrow, natural science definition of science. This point will be taken up later.

A third and equally important obfuscatory influence is intimated in the quote above, namely, Freud's promotion of psychoanalysis as a proprietary dogma (see also Jurjevich, 1974; Szasz, 1963). Elkin (1972) asserts that whether or not Freud intended to "he created a 'school' in the classical sense and his most devoted followers still believe that in his writings one can find the answers to most pressing contemporary problems" (p.56). Reinforcing this assertion Ellenberger (1970) rather disparagingly described this as Freud's most striking novelty, that is "the founding of a 'school' according to a pattern that had no parallel in modern times but is a revival of the old philosophical schools of Greco-Romano antiquity ... and this is no doubt a noteworthy event in the history of modern culture" (p.550). The noteworthiness of this aspect of Freud's contribution is more than offset by its negative aspects to be discussed below.

THE PSYCHOANALYTIC MOVEMENT

According to Szasz (1963), Freud's adoption of a highly monopolistic attitude towards psychoanalysis and early developed a restrictive and coercive organisation (namely The International Psychoanalytic Association) to restrict its use to those whom he
considered loyal disciples. This when coupled with the later development of a training regime derived from the guild concept acted (in concert) to effectively perpetuate his hegemony. Szasz claimed that Freud failed to evidence the kind of leadership we associate with the progress of science (i.e. the active encouragement of the creation and testing of new ideas) and instead evidenced the kind of leadership typical of big business or of imperial nationalism (1963, p. 146).

In looking over the literature, Johnson (1948) remarks: "one often has the feeling that the Freudian analyst is less concerned with results than with the fact that he is solely in possession of the one system of psychodynamics which purports to be thoroughly scientific" (p.321). Indeed Freud wrote to Ferenczi, in 1913: "We possess the truth; I am as sure of it as fifteen years ago". According to Szasz (1963), in scientific leadership Freud gave an example of what to avoid. By way of example Freud wrote: "Adler's 'Individual Psychology' is now one of the many schools of psychology which are adverse to psychoanalysis and its fuller development is of no concern of ours" (Freud, 1914 cited in Szasz 1963, p.156). According to Szasz "people who advocate this sort of discrediting of the investigation of their colleagues sever their loyalty to science and condemn themselves as propagandists" (1963, p.155).

Unfortunately, despite the pseudodemocratic, pseudoscientific atmosphere created by Freud, the realities of the Psychoanalytic movement were manifestly coercive and restrictive (Ellenberger, 1970; Jurjevich, 1974; Szasz, 1963). The medico-political, rather than scientific aims of the association were identified early by Bleuler.
One of Freud's earliest non-Viennese adherents, Bleuler refused to join the newly constituted International Psychoanalytic Association. One of the episodes which troubled Bleuler was the exclusion of a psychiatrist called Isserlin from attendance at the psychoanalytic meetings because of his persistent criticism. Bleuler was unpersuaded by Freud's claimed open-mindedness and persisted with his conviction that 'the introduction of the closed door policy scared away a great many friends and made some of them emotional opponents' (Bleuler, 1910 cited in Alexander and Selesnick, 1965).

History has vindicated Bleuler as the subsequent smearing, vilification and estrangement of Adler, Jung, Rank and others attests. The later snubbing of Horney and Fromme further evidences the closed door policy. Puner nicely depicts the contradictions expressed by Freud: literalism in verbalisations, ruthless dogmatism in behaviour. He wrote:

"He himself recognised the shortcoming of his theorising and painstakingly and consistently worked on ... (its) improvement. When he spoke this way, with a voice of reason, he spoke accurately and well. But his psychoanalytic children reacted not so much to what he said as to what he did. He spoke of freedom to amend, revise and change his doctrine, but the emotional atmosphere that he generated for his followers was one of rigid, watchful authority under which any attempt to deviate was treated as heresy. So his followers have interpreted every tentative work he wrote as the ultimate crystallisation of God-given truth. A pedantic and strictly defined conceptual system is exactly what has grown up among orthodox Freudians around the body of Freudian literature." (Puner, 1947, p.216)

The end result of Freud's proprietary dogmatism was an organisation referred to by Keen (1972) as the '... psychological establishment - the Freudian Mafia' (p.44). The impacts of the 'Freudian Mafia' in
developments in psychiatry and to a lesser extent psychology are myriad.

The charges range from:

(a) arrogance, evidenced in Jones' (then the president of the British Psychoanalytic Society) response to Glover's suggestion 'that every candidate (psychoanalytic student) should be instructed in the systems propounded by the then important psychoanalytical schismatics, Jung, Adler and Rank'. His reply was "Why waste their time?"

(b) usurpation of power and consequent 'brainwashing' of trainees (Brody, 1971; Eysenck, 1963; Glover, 1952; Marmor, 1968; Sargent, 1964; Shakow, 1967); and

(c) clandestine censorship of professional journals and books (Jurjevick, 1974; Kaplan, 1964; Pinckney and Pinckney, 1965; Samra, 1969).

On this last point it is interesting to note that Thomas Szasz' most influential paper 'The Myth of Mental Illness' was submitted to every major psychiatric/psychoanalytic journal between 1957-60 only to be repeatedly rejected. Ultimately it was published in probably the most hostile journal to psychiatry, the American Psychologist. Szasz sardonically described the course of events in the following way: "They all said it was no good, they rejected it. After that there came the second line of defense. There are the psychiatrists, the high priests, the bishops, then comes the psychologists, the local clergy. So I had to send it to the American Psychologist". (Szasz, 1974).
Religious symbolism vis-a-vis psychiatry and psychology is not new. White, for example, in response to early anti-German feelings proposed that 'the time had come to free American psychiatry from the domination of the Pope at Vienna' (White cited in Oberndorf, 1964, p.135-36). Similarly Kazin remarked that 'psychoanalytical literature has replaced the Bible as the place to which people turn for an explanation of their suffering and a source of consolation' (1958, p.15).

The point being made here should not be lost in the symbolism. Recall it was argued earlier in Chapter III (pp.64-66) that science has to a large degree assumed the authority previously given to religious teachings, and further, that psychiatrists and psychologists have recently assumed some of the roles previously ascribed to the elders in the church. In the Braginsky's words, psychiatrists and psychologists have become the "high priests of the middle class" (1973, p.15). Thus, in his conclusion to the chapter entitled "The Nuisance of the Freudian Establishment" Jurjevich asserts that in American psychiatry and clinical psychology it is clear that the 'high priests' who determine scientific 'truth' at this time are the members of the Freudian establishment. They do so without warrant (1974, p.67).

The tension between scientific responsibility and the pseudo-religious aspects of psychoanalysis has been recognised for many years. Friedlander voiced most of the contemporary concerns and objection to Freud's theories when he spoke before the international medical congress held at Budapest in 1909. Friedlander argued:
"First, instead of the quiet demonstrations usual with scientists in their discussions, psychoanalysts make dogmatic affirmations punctuated by emotional outbursts; psychoanalysts are unique in equating Freud with such men as Kepler, Newton, and Semmelweiss, and for the vigor of their attacks on their adversaries. Second, instead of proving their assertions in a scientific manner, psychoanalysts content themselves with unverifiable statements. They say: "We know from psychoanalytic experience that ..." and lay the burden of proof on others. Third, psychoanalysts do not accept any criticism nor even the expressing of the most justified of doubts, terming these 'neurotic resistance'. Friedlander quoted from Sadger: "The prudery of physicians in their discussions of sexual matters is due less to principle than to psychological background ... Rather than accept themselves as hysterics, they prefer to be neurasthenics. Even if they are neither, they prefer to declare the whole theory invalid and condemn it a priori". Friedlander agreed with Aschaffenburg that such argumentation was unacceptable among scientists. Fourth, psychoanalysts ignore what has been done before them, or by others, thereby claiming to be innovators. It is as if, before Freud, no hysterical patient was ever cured, and no psychotherapy ever practiced. Fifth, sexual theories of psychoanalysis, are presented as scientific fact, though unproven, as when Wulffen says: "All ethical powers in the interior of man, his sense of shame, his morality, his worship of God, his esthetics his social feelings, originate from repressed sexuality". Wulffen reminds one of Weininger when he said: "Woman is a born sexual criminal; her strong sexuality when successfully repressed easily leads her to illness and hysteria, and when it is insufficiently repressed, to criminality; often it will lead her to both". Sixth, Friedlander objected to the practice of psychoanalysts of addressing themselves directly to a wide lay public, as if their theories had already been scientifically proven; by so doing, they make those who do not accept the theories appear ignorant and backward." (Friedlander cited in Ellenberger, 1970, p.802-3)

In an attempt to clarify the nature of the opposition to Freud, and hopefully remove some of the barriers to accurate evaluation, Alexander and Selesnick (1965, chapter 13) proposed the merits of distinguishing between 'psychoanalytic thought' and the 'psychoanalytic movement', their point being that much of the opposition to psychoanalysis, as the previous pages evidence, is due to the dubious organisational features of the psychoanalytic associations
rather than wholesale hostility to psychoanalysis itself. Freud's assertion that psychoanalysis was his personal discovery and his attempts to make something special and sacred has obscured and all but prevented dispassionate evaluation of his ideas. Szasz contends that insofar as Freud, Freudians and post-Freudians have shown a special loyalty to psychoanalysis itself, as something other than a part of the study of persons, "they betrayed their faith in science and in humanity. Such loyalty can be purchased only at the cost of disloyalty to the ideal of the open society" (1963, p.156).

FREUD'S PSEUDOSCIENCE

Because most people see 'science' as good and 'pseudoscience' as bad they imagine that the two are diametrically opposed. They assume that there is a natural hostility between the two. That this is not the case was indicated in Chapters III and IV. The inability of scholars to satisfactorily arbitrate on the scientific status of psychoanalytic theory further attests to this difficulty. In the previous section some of the obstacles to evaluation erected by the psychoanalytic movement were discussed. Although important, these hindrances are peripheral to the central issue of the scientific status of Freud's theories. In this section support for Cioffi's assertion that psychoanalysis is most appropriately labelled pseudoscience will be presented.

Cioffi (1970, 1974) cogently argues that psychoanalysis is an example of pseudoscience; this he defines as being constituted 'not merely by formally defective theses but by methodologically defective procedures ... This notion of a pseudoscience is a pragmatic and not
a syntactic one (1970, p.474). He goes on to argue that underlying all of the defects of psychoanalytic theory is 'the same impulse: the need to avoid refutation'. The relevance of Popper's writings in judging the scientific merits of such a posture should be patent. With respect to Freud's and Adler's theories Popper stated: "The two psychoanalytic theories were in a different class. They were simply non-testable, irrefutable. There were no conceivable human behaviour which could contradict them." (1963, p.37).

With respect to Cioffi's accusation there can be no doubt that Freud was very sensitive to, but resistive of, demands from others that he revise his formulations. Note Bleuler's comments in a letter he sent to Freud concerning the intolerance of psychoanalysis towards dissent or criticism:

"Scientifically I still do not understand why for you it is so important that the whole edifice [of psychoanalysis] should be accepted. But I remember I told you once that no matter how great your scientific accomplishments are, psychologically you impress me as an artist. From this point of view it is understandable that you do not want your art product to be destroyed. In art we have a unit which cannot be torn apart. In science you make a great discovery which has to stay. How much of what is loosely connected with it will survive is not important."

(Bleuler cited in Alexander and Selesnick, 1965, p.6)

In the discussion to follow, psychoanalysis will be evaluated against several criteria currently used to differentiate 'genuine' from 'fictitious' science. But first it will be shown that Freud not only aspired to have his theories recognised as science but was adamant that they were scientific. This is a key issue in labelling Freud's theories examples of pseudoscience.
Freud's Assumption to Scientific Status

In our technological age, one of the most successful promotion tricks is to get under the scientific penumbra. Earlier (Chapter III), the ready acceptance of many preposterous ideas offered in the name of 'science' was documented (see MacDougal, 1968; Gardner, 1957). The status of science is evidenced by: "Science of Hotel Management", the Institute of Sartorial Science; Christian Science and Spiritual Science, to name a few. The Communists call their doctrines 'Scientific Socialism' supposedly to distinguish their teachings from the bourgeois tainted socialism which has departed from 'science'. La Pierre (1959), a sociologist, draws a parallel between the status of Marxist 'science' and that claimed by Freud.

"The Marxian doctrine of social evolution was not, as Marx and all his disciples since have believed, derived from the empirical study of social history. Like Freud's doctrine of man, it was imposed upon facts, not deduced from them. Marxist is therefore, unscientific in the same sense that Freudianism is unscientific." (p.60)

However, like the scientific socialists, Freud was at pains to emphasise in his lectures and articles that what he stood for was 'science'. Writing in 1925 about the scientific development of psychoanalysis, Freud unhesitatingly compared his troubles to the persecution Galileo suffered for the sake of promoting scientific 'truths' (Freud, 1959). On another occasion he proposed that his work was more worthy of admiration than that of Einstein since "He had the support of a long series of predecessors from Newton onward. While I had to hack every step of the way through a tangled jungle alone" (Freud cited in Hartmann, 1959, p.17). That this was patently not the case has been graphically documented (see particularly Ellenberger, 1970) and will be taken up in a later section.
Freud was, of course, trained as a natural scientist and was disturbed, as he indicated in his response to Bleuler's criticism cited earlier, that his case studies did seem to lack 'the serious stamp of science'; he grudgingly admitted that they sounded more like fiction (Freud, 1911 cited in Alexander and Selesnick, 1965, p.7). Nevertheless he was adamant that psychoanalysis was science like physics and chemistry and further he vehemently disclaimed the purported similarity between psychoanalysis and philosophy. He wrote:

"Psychoanalysis is not like a philosophical system which starts from a few strictly defined fundamental principles, uses them to embrace the totality of the world and, once perfected, has no room for new discoveries or improvement. On the contrary, it remains linked with the facts which are produced in its field of activity, it tries to solve the immediate problems of observation, tentatively continues its experience, it is always incomplete, always ready to correct or modify its theories. Like physics and chemistry it allows that its fundamental concepts are vague and its assumptions provisional. It does not expect a more rigorous definition than future work ..." (Freud cited in Marthe, 1966, p.173)

Freud was not alone in his claim to scientific status, Bailey (1964) has shown that Freud's followers continued to emphasise the scientific nature of their beliefs. He contests that 'psychoanalysis is certainly not a 'natural science' (Naturwissenschaft). Its followers, however, try desperately to make it appear so, because of the prestige our society attaches to science (1964, p.63). In contrast, an influential follower of Freud, Fenichel proposed that 'science is its [psychoanalysis] strong suit ... Its virtue of virtues of which all analysts are proud and which belongs to psychoanalysis alone, is that it is built upon scientific insight' (1945, p.14). In a recent lecture Hartmann
asserted that psychoanalysts' 'objectivity is scientific objectivity, his [her] truth is scientific truth' (1960, p.10). As a concession to the critics, he added cautiously 'This attribute has frequently met with misunderstanding'. As a last example, Ernest Jones' belief in the scientific significance of Freud is revealed in the following:

"The future world may well speak of a Pre-Freudian and Post-Freudian era of thought. Man's conquest of nature has been proceeding for many thousands of years, and fumbling attempts have often made in his more difficult task of self-conquest, but Freud's life work represents the first serious endeavour to apply to it the methods of science. After all, Freud has been educated, not as a psychologist or mythologist, but in the tenets of orthodox neurology. Undeterred by this bias, however, Freud determined to examine the facts themselves and let nothing but their evidence influence his conclusion." (1956, p.121, 124).

If one believed Jones it would appear that what Copernicus and Darwin started, Freud has finished!

La Pierre, in his book 'The Freudian Ethic' (1959), suggests that the persistent claim to science would no doubt have been brushed aside if it had not been sponsored by '...men [women] of medicine ... Since they were Doctors of Medicine before they became Freidians, and since they did not abandon their status of physicians in assuming that of Freudian psychoanalysts, they gave to the practice and advocacy of Freudianism the traditional authority of medicine' (p.42). Such faith (unconditional) in professionally qualified people was poorly founded as the empirical research on suggestion and diagnostic judgements indicates (Masling, 1960; Temerlin and Trousdale, 1969).
Masserman elegantly demonstrated the gullibility of physicians when he told a learned audience about a new psychosomatic syndrome of 'onychoneurosis' (the stubbed toenail neurosis). To his surprise the audience had taken his playfulness seriously and were convinced that they had heard a new revelation (Masserman, 1953). Recent research, to be introduced in Chapter VIII indicates that professional scientists and statisticians generally think as unscientifically and non-probabilistically as the lay population (see e.g. Johnson-Laird and Wason, 1977; Mahoney, 1976). There is no reason to think that physicians would be any different. Indeed many argue that psychoanalytic training in fact promotes a decline in the scientific ability of trainees (Brody, 1971; Marmor, 1968; Shakon, 1967).

Psychoanalysis Ought to Become Scientific

After the insistent and sometimes passionate claims to scientific status by Freud and his followers, and after the passage of seventy-five years, a recent friendly reviewer concluded that psychoanalysis 'can be and ought to become a science' (Rapaport, 1968, p.32). Similarly Fisher and Greenberg in the conclusion to their mammoth survey of experimental literature pertinent to Freud's theories express their optimism that 'it is possible to approach Freud's work in a scientific spirit' (1977, p.396) and that the time 'has come to face up squarely to the sparseness of what has generally been offered as support for Freudian formulations' (p.6). They recognised that many of the Freudian principles have been and are currently perpetuated by faith and authority.
Glover (1952) also highlights this. He suggests that most disputes within psychoanalytic circles are settled by invocation of authority rather than examination of the pool of scientific evidence. Invariably looking back at Freud's original statements and case illustrations replaces data induced conflict resolution and decision making. Glover describes the common way in which 'new facts' are established in psychoanalytic circles: "An analyst, let us say, of established prestige and seniority produced a paper advancing some new point of view or alleged discovery in the theoretical or clinical area. Given sufficient enthusiasm or persuasiveness, or even just plain dogmatism on the part of the author, the chances are that without check, this view or alleged discovery will gain currency, will be quoted and requoted until it obtains the status of an accepted conclusion" (1952, p.403). Thus what change manages to occur reflects the power, status, persuasiveness or plain persistence of the innovating proponent. The early response of the psychoanalytic 'establishment' to Eric Fromm reflects this process graphically. Fisher and Greenberg (1977) suggest that the discarding of ideas is also informal and equally unscientific. They use as an example some of Freud's ideas (e.g. inherited dispositions, death instinct) that have, for all practical purposes, been rejected by the psychoanalytic establishment, yet this occurred without explicit testing of their validity and subsequent rejection or without a direct statement that they would no longer be given serious weight.

Speaking before the American Psychiatric Association in 1965 Engel asserted that psychoanalysis had not even begun to assume the status of a science; that it was still at a level of pre-science.
Incidentally this was the conclusion arrived at by Popper in the early 1930's. Engel asked: "What are the reasons that psychoanalysis as a science has failed in seventy-five years to progress significantly beyond the stage of an observation, data collection and theory building ... Theoreticians and would-be theoreticians we have in abundance but no more than a handful of scientists skilled in and dedicated to the critical examination of natural phenomena" (1968, pp.214-215). In a response to Engel's address Walter Stein admitted that he had touched 'a raw nerve of the body psychoanalytic, the state of crisis in psychoanalytic research ... our collective failure to line up to the enthusiastic overselling of our therapeutic promise that ushered in our sweeping acceptance by educated America ... Psychoanalytic research has never really gotten off the ground' (Engel, 1965, p.216).

Notwithstanding Silverman's (1976) recent presentation of several sound research programs, the general absence of an ongoing research program could be held responsible for what Ford and Urban (1976) called the stagnancy in psychoanalytic literature. In their review of Freudian therapy for the 'Annual Review of Psychology' they conclude that 'there is little substantive novelty in these writings ... the innovative stream is gone out of the psychoanalytic movement. Major theoretical and technical advances will probably come from other orientations' (1976, p.33). Psychoanalysis it appears has failed to avail itself of the subversive nature of scientific examination to rid itself of that which is defective and redundant. Jurjevich suggests that future historians in their summing up as to the status of Freudian science will conclude that 'they talked big, they promised a lot, they wrote copiously, but it is hard to find scientific grains on the piles of chaff' (1974, p.263).
PHILOSOPHICAL CONSIDERATIONS

Participants in the debate about the scientific standing of psychoanalysis rightly refuse to accept the *ex cathedra* assertions of Freud and his followers. But as we have seen before, the issues do not lend themselves to easy, black or white decisions. Whether one accepts or rejects Freud's claim is as was mentioned earlier inextricably enmeshed with one's definition of the concept 'science'. Science may be conceived of broadly or narrowly. The broader concept accepted by most European thinkers sees it as any productive, factual investigation which has theoretical significance. The narrower concept, expressed by most Americans, identifies science with an experimental or quasi-experimental procedure in which general theories produce specific hypotheses to be tested against predictions made with reference to a set of quantitatively tabulated data. Most critics of Freud use the narrower conception and, besides Eysenck, tend to be primarily American.

Freud's Theories - Progressive and Degenerating Program

Surprisingly, many of the leading philosophers of science are unanimous in their judgement that psychoanalysis is not in fact a scientific theory. We have already seen that Karl Popper (1963), who is regarded by some as the world greatest living philosopher (e.g. Magee, 1974), considers that because Freud's theories can explain everything, can predict nothing and are therefore judged to be prescientific since falsification is impossible. "What makes a theory, or a statement, scientific," according to Popper, "is its power to rule out, or exclude, the occurrence of some possible events - to prohibit, or forbid, the occurrence of these events".
Thus, the more a theory forbids, the more it tells us (Popper, 1976, p.92). To briefly recapitulate on Popper's position, he asserts that all theories are ultimately false and can not be proved right but will ultimately be proved wrong. Thus he suggests instead of dividing scientific theories from pseudoscientific ones we should rather divide scientific method from pseudoscientific method: scientific method being characterised by the criteria of falsifiability. Thus, if a theory specifies experimental conditions which could lead to disproof of the theory (as Einstein so elegantly and courageously did) then, and only then, that theory is scientific, if not, not. Consequently, a proposition may petrify into pseudoscientific dogma or become genuine scientific knowledge, depending on whether we are prepared to state observable conditions which would refute it. Using this reasoning and criterion Popper decided that both Marxism and Freudism are pseudoscientific theories (Popper, 1963).

There is much support for Popper's judgement. For example, Ernest Nagel (1959) also concluded that psychoanalysis is not a science since its concepts and formulations are so vague that they can not be tested. Yet, contradictorily he went on to reject psychoanalysis on a second account, namely, that some experimental evidence and anthropological data failed to support it. (This point will be developed subsequently.) Farrel (1964) is another to conclude that psychoanalytic theory is not scientific. He agreed with Nagel's comment that the Freudian theories are not sufficiently determinant, logical or empirical but in addition suggested that the data yielded to the psychoanalytic method was of doubtful scientific validity. And lastly, Thomas Kuhn whose international standing is
comparable to that of Popper, concurs with him with respect to the scientific status of psychoanalytic theory. He stated that in "Examining the vexing case, for example, [of] psychoanalysis or Marxist historiography, for which Sir Karl tells us his criterion was initially designed, I concur that they cannot now be properly labelled 'science'" (1970, p.

Unfortunately, but as one would expect and as was alluded to in Chapter III, there are serious difficulties connected with Popper's view. At a conference called to discuss his and Kuhn's conjectures, it became very apparent that falsification theory itself had significant weaknesses (Lakatos and Musgrove, 1970). The problems were made apparent by studying the success of Newton's theory of gravitation for although shown to have been literally riddled by anomalies and contradictions, its proponents did not give up the theory: paradoxically they behaved very much like convinced Marxists and Freudians. Lakatos speculated that 'had Popper ever asked a Newtonian scientist under what experimental conditions would he [she] abandon Newtonian theory, some Newtonian scientists would have been exactly as non-plussed as some Marxists and Freudians' (Lakatos cited in Eysenck, 1976, p.343). Incidentally, Kuhn's views have also been found to be wanting as the change from one paradigm to another as described by Kuhn, is not supported by historical evidence (see e.g. Chalmer, 1976; Weimer, 1979). The historical evidence refutes both Popper and Kuhn: 'on close inspection both Popperian crucial experiments and Kuhnian revolutions turn out to be myths: what normally happens is that a progressive research program replaces degenerating ones' (Lakatos op cit, p.345).
Imre Lakatos' answer to the demarcation issue is now perhaps the most widely accepted view among contemporary philosophers of science. To a considerable extent Lakatos' view combines and transcends those of Popper and Kuhn, and importantly bears directly on the matter under discussion. That is, is psychoanalysis an example of a progressive or degenerating research program?

Lakatos (1970) introduced the notion of research programs that may take decades before they get off the ground and become empirically progressive. Each program has a 'hard core' (i.e. a set of assumptions, laws, conjectures or in Kuhn's terminology a paradigm) that is tenaciously protected from refutation by a vast 'protective belt' of auxiliary hypotheses. But even more importantly, research programs have a 'heuristic', that is, a powerful problem-solving machinery which, with the help of sophisticated techniques, digests anomalies and even turns them into positive evidence.

According to Lakatos, Newton's theory of gravitation, Einstein's relativity theory, quantum mechanics, Marxism and Freudism, are all research programs, each with a characteristic hard core that is stubbornly defended, each with its more flexible protective belt and each with its elaborate problem-solving machinery. Each of them, at every stage of its development has unsolved problems and indigested anomalies. He points out that all theories, in this sense, 'are born refuted and die refuted'. But all theories are not necessarily equally 'good'. Some theories are scientific or progressive programs whilst others are pseudoscientific or degenerating. But how does one tell the difference?
Lakatos answers that a progressive research program theory leads to the discovery of hitherto unknown novel facts. In contrast degenerating program theories are fabricated only in order to accommodate known facts. To use Newtonian physics as an example, advocates were able to predict the existence and exact motion of small planets which had never been observed before. Equally Einstein's program predicted novel facts that are only now being corroborated. With respect to Marxism, Lakatos (a confessed Marxist), concedes that it never successfully predicted a stunning novel fact. On the contrary; it had some monumental failures to predict (i.e. the absolute impoverishment of the working class, the first socialist societies would be free from revolution).

"To sum up: the hallmark of empirical progress is not trivial verification. Popper is right that there are millions of them. It is no success for Newtonian theory that stones, when dropped, fall towards the earth, no matter how often this is reported. But so-called 'refutations' are not the hallmark of empirical failures, as Popper has preached, since all programs grow in a permanent ocean of anomalies. What really counts, are dramatic, unexpected, stunning predictions: a few of them are enough to tilt the balance, where theory lags behind the facts, we are dealing with miserable degenerating research programs." (Lakatos, ibid, p.345)

What then of Freud's theories - progressive or degenerating program? Eysenck sarcastically suggests that the claims made by Freud and his followers are certainly stunning, but he adds, there is a complete lack of empirical support for the predictions (Eysenck, 1976). Before evaluating Eysenck's claim, the predictive strength of psychoanalysis will be discussed.
Prediction and Freud's Theories

It is important to note that Freud himself claimed that the psychological determining factors can be established retrospectively only, and once discovered cannot be used predictively in other cases. His explanation for what he called "this disturbing state of affairs" was rather simple: we cannot yet measure exactly the relative strength of psychological factors with the result that in a circular fashion "we only say at the end that those which succeeded must have been the stronger" (Freud 1901, cited in Jahoda, 1977, p.15). Analysis can therefore recognise causality retrospectively with certainty, whereas prediction is impossible. Empirical evidence tends to support this conjecture. To restate this important point, "Freud 'felt that psychoanalysis could reconstruct the patients' pattern of defenses from events that had already occurred but that it was not able to plot them in advance" (Fisher and Greenberg, 1977, p.8).

Notwithstanding Kaplan's (1964) defence of reconstructive or postddictive models as an appropriate research methodology for clinical research, postdiction and ad hoc explanations are generally eschewed by empirical psychologists. Although after-the-facts accounts of an observed phenomenon may illustrate some of the assets of a particular inference or construct and may be suggestive, they also invite conceptual 'cop-outs'. For example, theorists can isolate themselves from the corrective feedback of empirical evidence by equipping their theory with an ever-ready supply of ad hoc explanations. The Freudian defence mechanism of 'reaction formation' like the 'divine paradox' effectively produces a closed philosophical system. That is, Freud postulated reaction formation in order to
account for the fact that observation occasionally revealed exactly
the opposite sort of effect to that derived from the theory. Such
a construct provides an effective way of 'avoiding refutation'.

Returning to the question of Freud's theories vis-a-vis
Lakatos' definition of science in terms of a progressive program and
pseudoscience in terms of a degenerating program, there is little
doubt that Eysenck is correct in concluding that Freud must join
Marx in the latter category. However, philosophical considerations
make up but one of the many aspects of the problem: equally
important is the empirical literature. It was noted earlier that
many critics of psychoanalysis dismissed its claim to scientific
status because they judged the Freudian system insufficiently
determinant to permit testable hypotheses to be deduced. Such a
stance has been empirically shown to be a myth. In fact Fisher
and Greenberg claim the quantity of research data pertinent to
Freudian theories grossly exceeds that available for most other
personality and development theories. They conclude: "We have
actually not been able to find a single systematic psychological
typey that has been as frequently evaluated scientifically as has
Freud's" (1977, p.396).

EMPIRICAL VALIDATION

The quantity of research literature certainly attests to the
amenability of Freud's theories to experimental scrutiny. However,
quantity is no guarantee of quality. To attempt to arrive at a
position with respect to the quality of the vast literature is
beyond the scope of this essay - in the limited space available only
the briefest overview is possible.
There have been only a few large-scale attempts to judge whether Freud's theories are empirically corroborated. An early example was the study by Sears (1943). He concluded that "other social and psychological sciences must gain as many hypotheses and intuitions as possible from psychoanalysis but that the further analysis of psychoanalytic concepts by non-psychiatric techniques may be relatively fruitless as long as those concepts rest in the theoretical framework of psychoanalysis". Obviously Sears' work is quite out of date and since his advice was not taken - research has continued to burgeon as is evidenced by the recent publication of Fisher and Greenberg's (1977). They concluded, after consulting approximately 1500 research reports, that a considerable scientific literature supports Freud's fundamental view that motives, feelings, and fantasies may exist in an individual without her/his awareness of them and that they may significantly influence his/her behaviour. However, their conclusion must be tempered by a consideration of the criteria applied in the selection of articles to be included. Although they excluded case histories they decided not "to rule out studies that had defects in their experimental design or that were based on over simplistic notions concerning Freud's model" (1977, p.15). Their rationale was to propose that it "seemed more sensible to make a sweep of the total empirical data, flawed or otherwise, and to draw conclusions from overall trends" (1977, p.15). The wisdom of this decision seems questionable. Research to be discussed in Chapter VII will evidence the difficulty of the task set by Fisher and Greenberg in evaluating such a vast literature without employing explicit a priori decision making criteria.
A third large scale study described by Fisher and Greenberg as a 'particularly dedicated effort', by Jahoda as 'meticulous', and by Eysenck as 'conscientious, scholarly and well-documented' was published by Paul Kline in 1972. After reviewing some 500 research reports, Kline concluded that "far too much that is distinctively Freudian has been reported for the rejection of the whole psychoanalytic theory to be possible"(1972, p.350). However, after a 'searching examination' of the evidence presented by Kline, Eysenck (1972) retorted "that the marshalled evidence for Freudian theories leaves the reader little option but to conclude that if this is the best that can be offered by way of support, then the only conclusion can be that there is no evidence at all for psychoanalytic theory" (1972, p.321). According to Eysenck and Wilson (1973) it did not seem that "he (Kline) was entirely impartial in looking at the studies being surveyed, but tended to overlook reasonable alternative hypotheses, accepted tests of poor reliability and no proven validity, paid little attention to problems of sampling, and was rather naive in considering the adequacy of the statistical treatment given to the results" (1973, p.386).

The last study to be discussed was instituted in response to the publication of Kline's 'Fact and Fantasy in Freudian Theory' (1972). Eysenck and Wilson (1973) undertook to evaluate a limited number of articles which were widely believed to be the most convincing, the best designed and the most conclusive amongst those which confirm Freudian theories. They adopted and applied the same evaluative criteria as Kline had and claimed that they did not apply these more stringently than is normal practice in experimental psychology in other fields: 'if anything, we have been inclined to bend over backward to temper the wind to the shorn lamb' (1974, p.386).
Their 'dispassionate' conclusion was that the studies scrutinised gave little, if any support to Freudian concepts and theories. In fact they postulate:

"There is not one study which one could point to with confidence and say: "Here is definitive support which is not susceptible of alternative interpretation, which has been replicated, which is based on a proper experimental design, which has been submitted to proper statistical treatment, and which can be confidently generalised, being based on an appropriate sample of the population." After three quarters of a century this is a serious indictment of psychoanalysis, whether it can be interpreted as an overall disproof is of course another matter" (Eysenck and Wilson, 1973, p.392)

It would be difficult to attribute impartiality to Eysenck when it comes to evaluating Freudian concepts - nevertheless his conclusion should be responded to not with emotion and polemic but with data. This has not been forthcoming, although as mentioned earlier, Silverman's (1976) report on some contemporary research programs looks promising.

By way of a tentative conclusion one must agree with Eysenck; but to emphasise that refutation alone never disproves a theory. As has been seen, whether a theory is subscribed to and survives depends not on the empirical data alone but on the balance between positive and negative instances, on the ability of the problem-solving machinery to digest anomalies, and on the ability of the theory to predict novel facts. Against these criteria Freudian theory does not hold up well. Its survival seems due more to the vagaries of the Zeitgeist and the perceived absence of a viable alternative paradigm.
CHARGES OF DELIBERATE SCIENTIFIC PRETENCE

One of the less attractive features of the debate about the scientific status of psychoanalysis has been the occasional vitriolic personal attack on Freud. To this point the argument has not suggested that psychoanalysis was the idiosyncratic product of Freud's wilful, deliberately deceptive charlatanism. Such claims have, however, been made. Despite Freud's convincing display of sincerity, cogent arguments have been advanced that his pseudoscience contains important elements of wilful pretence to science. These propositions will be examined below. By way of a palliative for those who find ad hominem argument obnoxious, it is only fair to say that they were also used by psychoanalysts and were a frequent ploy of Freud (as evidenced by his discrediting of Adler, Jung, Stekel, Rank and others who refused to abide by the original 'revealed' truths).

The Sources of Freud's Ideas—Observation or Literature

Much controversy surrounds Freud's claim that he learned all he knew from his patients. Twenty years ago Bakan (1958) suggested that Freud's inferences, rather than stemming from clinical observation, were surreptitiously imposed on them by Freud, the source of inspiration being the mystical fantasies of medieval and ancient Kabbalists, and particularly the esoteric, mystical and sexualised fantasies of Zohar. To substantiate his claim Bakan produced some amazing parallels between Kabbalistic thought and psychoanalytic ideas, particularly with regard to the role of sexuality. Rabbi Grollman (Grollman, 1965) agreed with Bakan that
Freud unmistakably utilised many of the quasi-insights of mystical medieval Jewish dreamers to form some of the central propositions of his theories and practice (i.e. alleged sexual character of psychic energy, the exegetic method of dream interpretation).

Bakan leaves it open as to whether Freud deliberately disguised Talmudic teaching as clinical observation or whether Freud perpetrated a 'sort of reversed Piltdown hoax in psychology' (p.12). To use Jahoda's (1977) words, Freud never admitted publicly to his ideological indebtedness to Jewish mysticism, but in his private correspondence with Fliess, Abraham and Pastor Pfister he did speak of the Jewishness of his doctrine. In fact, publicly Freud denies in print any knowledge of Yiddish or Hebrew, although Jones (1953) informed readers that Professor Hamenschlag taught Freud the 'Scriptures and Hebrew'. Bakan considers such a dismutation strange. Equally strange is the reported absence of Zohar from Freud's library in New York even though a Jewish visitor had seen it among Freud's books in Vienna (Bakan, 1958).

Suspicion about Freud's claim for the data-dictated nature of his inferences has more recently been suggested by Cioffi (1970, 1974). Cioffi undertook the most meticulous textual analysis of Freud's various statements of his account of how he arrived first at and later rejected the theory that hysteria was the result of a childhood sexual seduction. The results of this scholarship was the explicit accusation of pseudoscience, more specifically, that Freud engaged in the 'habitual and wilful employment of methodologically defective procedures' (Cioffi, 1970, p.474). As was mentioned earlier, he went further to argue that underlying all the defects of psychoanalytic theories is the same impulse - the need to avoid
refutation.

By way of example Cioffi examines the hypothesised link between adult neurosis and infantile sex life. He suggests that such a link may be justified and corroborated by the accuracy of those portions of the reconstructions which are held to characterise childhood in general, thus being capable of 'confirmation' by the contemporary observations of children. Freud seemed to concur when he wrote "I can point with satisfaction to the fact that direct observation has fully confirmed the conclusion drawn from psycho-analysis" (Freud, 1958, p.594). Indeed he claimed that the 'facts' of infantile sexuality were so obvious to the naked eye, once they had been stated boldly that it was a sheer miracle that they had not received full recognition before. Freud regarded the various manifestations of infantile sexuality as 'facts' but 'facts' in science alas are not as straightforward as he would have us believe. An event (fact) does not have a unique interpretation, as Geertz (1973) has emphasised. You cannot literally see the difference between a wink and an involuntary twitch. As Jahoda (1977) highlighted, facts and inferences are often difficult to disentangle in Freud's work, but testing the inferences as a way of approaching the factualness of infantile sexuality is, of course, a legitimate procedure.

Several empirical psychologists have collected relevant research together on the issues that were the focus of Cioffi's analysis and generally found little support for Freud's conjectures. For example Orlansky (1949) surveyed the research on infantile sexuality and concluded that there was no supportive evidence. Similarly, in the
Sear's (1947) survey little support for the Freudian construction of infantile sexuality was forthcoming. A further negation of the primacy of sexual factors in early childhood came from Valentine (1942). He reported no evidence in support of the supposed Oedipus complex and went on to conclude:

"Whether the ideas of infantile sexuality reported by patients are indeed (a) suggested by psychoanalysts - as Freud at one time himself suspected - or (b) are entirely or partly the patient's own interpretation of and exaggeration of relatively slight sensations and impulses, or (c) whether they are largely true but only in a few abnormal cases, this is not the place to discuss. But the fact that the reports of patients, which Freud himself took at first to be facts, proved to be mere fantasies, is very significant (1942, p.351)

Cioffi does not rest his case there, he goes on to point out the way in which Freud retreated to the esoterically observable (i.e. analytic observations) when he was faced with disconfirmatory evidence. Obviously such a retreat renders the direct observations of children futile for the purpose of validating psychoanalytic method and renders the theory impervious to refutation. It is not being suggested that analytic observations are intrinsically deficit or invalid, but to evoke them as a ways of avoiding disconfirmation in the natural science paradigm is not legitimate - given Freud's insistence on the natural science nature of psychoanalysis.

Cioffi also takes up a point made by Farrell (1964) and others (Escalona, 1952; Flowerman, 1954; Ludwig, 1947; Stuart, 1970) that Freud wrote in such a manner as to make disproof difficult if not impossible. His writings they claim are often contradictory, curiously indecisive and generally abstruse. Flowerman (1954)
has offered some striking examples of how Freud could at times formulate an idea in such a temporizing way that it was truly impossible to test its validity. Similarly Escalona (1952) describes the potential complexity that might be involved in corroborating the existence of Oedipal conflicts:

"Thus, if the child gives daddy a good night hug and insists that he, rather than mummy, tuck him in, his behaviour may also confirm our original hypotheses. His desire to have father put him to bed rather than the mother could be the result of a fearful state, i.e., as long as the father is with him the little boy can be sure the father is not doing anything to harm him. On the other hand, or also simultaneously, it may be an act of aggression towards the father in that it separates him from mother for the time being. Or yet again, it may be because the little boy fears that if mother puts him to bed her seductive powers will prove too much for him ..." (p.16).

That Freud wrote in a manner that effectively erected barriers to refutation can not be denied. Whether it was deliberate is inconclusive.

Free Association - The Methodological Key

Freud's claim that psychoanalysis is a natural science rests to a considerable extent on the psychoanalytic method of observing the unconscious. He believed that most important source of data in psychoanalysis was made available through free association which was 'the methodological key to its results' (Freud, 1960, p.403). The analyst was seen to be an impartial scientist who took utmost care not to interpose on the flow of data from the patient. Freud claimed to have taken up the position behind the couch to create an immaculate experimental situation: "Since while I listen, I resign myself to the control of my unconscious thoughts, I do not
wish my expression to give the patient indication which he may interpret or which may influence him in his communications" (Freud cited in Campbell, 1957, p.146). How well Freud achieve this non-directive posture and how honest was his image is evidenced by his comment to Fliess on how gullible some of his patients were in their acceptance of sexual etiological speculation. Meanwhile things have grown livelier in the usually struggling medical practice. The sexual business attracts people, they go away impressed and convinced, after explaining: No one ever asked me that before". (Freud cited in Bonaparte et al, 1954, p.354).

Unfortunately for Freud and his 'methodological key', it has been shown that psychoanalysts (like all psychotherapists) are not detached observers; they influence what the client says and do so in all sorts of uncontrolled and unintentional ways. The subtlety of this influence is graphically evidenced by recent work on non-verbal communication (see e.g. Henley, 1977; Scheflen, 1974). Haley (1963) goes as far as to suggest that the insistence that clients lie on the couch places the client in an inferior, manifestly sick role which may serve to heighten the psychoanalyst's indirect, suggestive power. Be that as it may there are grounds for considerable doubt that the analytic hour approximates in any way to a controlled experiment (as claimed by Kubie, 1960, Ramzy, 1962, 1963). Data suggests that psychoanalysts who witness the events of an analytic session have evidenced difficulty agreeing in their interpretation of these events (see e.g. Marmor, 1955; Seitz, 1966)
Interpersonal influence in psychotherapy is ubiquitous as Frank (1968) has pointed out and this applies to all forms of therapy. Frank argues that "Evocative therapies may influence patients as much as directive ones. The powerful influencing effect of psychoanalysis and other 'permissive' therapeutic methods have long been rated ... To recapitulate, the content of the patient's utterances on non-directive therapy follow the therapists' unwitting leads, and patient's values shift toward those of the therapist in the course of successful treatment" (p.22).

Contemporary psychoanalysts would benefit from reading the vast literature on experimenter induced biases and demand characteristics (Orne, 1962; Rosenthal, 1963). The importance, subtlety and robustness of these influences are convincingly documented in a recent publication entitled 'Self-fulfilling Prophecies' by Jones (1977). Barber's (1976) recent elucidation of the ways (ten in all) in which the behaviour of the experimenter may threaten the internal validity of an experiment seems particularly relevant too. For as Pinckney and Pinckney contend: "Although it is vigorously denied today, Freud admitted that the process of psychoanalysis puts thoughts as well as words in the patient's mind and mouth. You cannot apply the scientific method to test a hypothesis where you create your own evidence out of words, discard whatever evidence displeases you, and then say that your concept is correct simply because you say so" (1965, p.76).
In short, evidence suggests that this 'methodological key' is so permeated by the analysts' comments and influences that it destroys the scientific validity of psychoanalysis. 'It makes Freud's findings of dubious scientific value; the contaminating variables of the analysts' overt and covert suggestion render Freudian 'finding' scientifically worthless' (Jurjevich, 1974, p.283-284).

Along with free association, a frequent bone of contention was the apparent arbitrariness of Freud's interpretation and his employment of symbolism in particular. Writing in 1912 Wells referred to symbolism as the phase of psychoanalysis to which most legitimate objections are raised (Wells cited in Cioffi, 1973, p.18). The accusation is that Freud and his followers make symbols mean what they want. Thus it is claimed that although Freud was aware of the influence of suggestion and instituted rules designed to minimise its occurrence, his insistence that the latent dream thought was outside the suggestive power of the analyst is untenable. The skepticism about the interpretation of this content is captured by Aldous Huxley:

"It was the machinery of symbolism, by which the analyst transforms the manifest into the latent dream content, that shook any faith I might possibly have paid in the system. It seemed to me, as I read those lists of symbols and other obscure allegorical interpretation of simple dreams that I had seen this sort of thing before. I remember, for example, that old-fashioned interpretation of the Song of Solomon ... I had never, even in infancy whole-heartedly believed that the amorous damsel in the Song of Solomon was, prophetically the Christ, and her lover the Saviour. There are no better reasons for believing that walking upstairs or flying are dream equivalents of fornication than for believing that the girl in the Song of Solomon is the Church of Christ." (1928, pp.316-17)
Once again, the evidence seems to point overwhelmingly to the conclusion that the foundation of Freud's claim to science, namely, the psychoanalytic method of observing and interpretation, is seriously flawed. This is not to say that Freud consciously perpetrated a bias by ignoring the serious weaknesses in his methodological key but it is suggested that this weakness was a major constituent in the Freudian pseudoscience, irrespective of Freud's motives.

ALTERNATIVE EPISODEMOCYIES

There are other issues and controversies, serious and trivial, which will not be discussed here as the case seems well established. But before leaving the issue it should be restated that the ascription of the label pseudoscience to Freud's theories is not meant to demean his contribution to our understanding of ourselves. It should be restated that science is only one of several ways of gathering knowledge and is not intrinsically better or worse than other methods. Judgements depend on the purpose and criteria used. And as has been evidenced, science is an ambiguous and value laden term and cannot be clearly defined. In the foregoing discussion a narrow or, in Maruyama's (1977) terminology, a hierarchical and non-reciprocal causal definition of science was used where as many supporters of Freud's ideas suggest that this conceptualisation of science cannot be appropriately applied to psychoanalysis (Habermas, 1971; Jahoda, 1977; Ricoeur, 1970).

Even though Freud claimed that 'psychoanalysis is a natural science - "what else could it be?" (i.e. Naturwissenschaft) others propose that psychoanalysis is better seen as hermeneutics (see
particularly Habermas who takes psychoanalysis as an outstanding example of a hermeneutic science). Psychoanalysis as hermeneutics belongs to the Geisteswissenschaft, for which the term humanistic science is the only available term. It is argued that like history or literacy criticism, psychoanalysis should not be expected to establish its validity in the manner of the natural sciences by prediction from theory. Rather psychoanalysis should above all evidence coherence of part and whole interpretation, using emerging evidence from 'factual' events (Habermas, 1971).

There has been a much criticised (see e.g. Giorgi, 1970; Koch, 1961; Rychlak, 1977) tendency in psychology, and the social sciences in general, to see the natural science epistemology as being the only 'scientific' approach to studying phenomenon. However, this is not the case. In a recent paper Maruyama (1977) delineated five different epistemologies available and currently employed in scientific research. More recently, Sampson (1978) has called for psychology to embrace an alternative to the natural science paradigm, in particular, a paradigm that recognises the impact of sociopolitical and historical factors on knowledge generation and validation. In light of the obvious relativity of the concept of science Freud may have been mistaken to identify with the Naturwissenschaft rather than the Geisteswissenschaft, for it appears that in dealing with the person, rather than with part psychological process, the former approach has been fairly unsuccessful. O'Neil (1972) has argued that "Those of us who have to deal with persons in our day-to-day work seem to be much more helped by a humanistic approach with its intuitive judgements of the concrete situations and its empathic judgements of what follow from there" (pp.88-89). Had Freud explicitly identified
with the Geisteswissenshaft this legitimate approach to understanding persons may have been further developed at this stage in history, notwithstanding, the powerful arguments of Habermas (1971) and others (Jahoda, 1977; Ricoeur, 1976). Hebb (1974), has suggested that attempts to make science simultaneously scientific and humanistic is misguided, if not impossible. In accord with what has been said above, he claims that what occurs is the confusion of two different ways of knowing human beings; one is science, the other is art. Science imposes limits on itself and makes progress by attacking only those problems it is fitted to attack by existing knowledge and methods. The other way of knowing about humans is the intuitive insight of the poet, novelist, historian and biographer. This alternative way of understanding people is a valid and deeply penetrating source of insight into ourselves and others but it is not 'scientific' and would be pretentious if it claimed to be.

In a recent paper, Farson (1978) reviewed the progress of 'humanistic psychology' and illuminated a marked disparity between the early claims and current practice within this orientation. Farson highlighted the dearth of scholarly scientific work conducted by humanistic psychologists and suggests that it will be unlikely to make a serious contribution to the discipline. However, the ideas of Rychlak (1968, 1977), particularly as expressed in his recent book 'The Psychology of Rigorous Humanisms' are an obvious exception to Farson's disparagement.
It is beyond the scope of this essay to discuss and evaluate the dilemma which besets psychoanalysts vis-a-vis their choice of self-definition (see Salter, 1972, for a neat delineation of the alternatives). The aim of the chapter was simply to substantiate the claim that Freud's theories cannot be legitimately called 'scientific': if science is defined as a conventionally accepted narrow, natural sciences sense. The inappropriate labelling of Freud's theories as scientific has been costly for the development of an efficacious approach to psychodiagnosis. It has lead to the widespread acceptance of the assertion that behaviour is relatively independent of the situation in which the person interacts. Consequently, personality was perceived as a set of needs, drives, psychodynamic structures or trans-situational traits, that initiate and guide behaviour. Thus the function of psychodiagnosis is to accurately measure these dispositions and to translate them into classificatory descriptions. The adequacy of these notions will be discussed in the next chapter.
Historically, it has been demonstrated that clinical psychology, as a viable profession in the mental health field, emerged at the end of World War II and began to bloom thereafter. However, as we have seen, because of the restrictive policies of the 'psychiatric establishment' of the time, psychologists were not permitted to engage in the prestigious practice of psychotherapy but were relegated to the auxiliary assessing function. Their subsequent development of exclusive expertise in the administration of esoteric measuring devices facilitated the demarcation of an area of influence within the mental health profession. It was suggested in Chapter III that because of the professional exigencies psychologists uncritically adopted and routinely applied measurement devices without questioning the purpose of the exercise. Using indirect projective tests and objective personality inventories containing disguised properties (i.e. MMPI, California Personality Inventory) clinical psychologists ended up performing a role analogous to an X-ray technician. Whether the fruits of their labour had any impact on the ensuing decision making and treatment plans sadly seemed immaterial. Their role was that of assessor.

In embracing this role early, clinical psychologists took over the objective assessment devices that have been shown in Chapter IV to be the product of research and development conducted within a seriously defective paradigm - at least against the criteria adopted. It was argued that the resulting psychometric theory and research represented an example of 'grand' pseudoscience. Further, they also took over the Rorschach ink blocks, Murray's TAT and other projective and associative techniques that were designed to provide inferential
evidence of the state of the examinees' unconscious motives and conflicts. These techniques were a direct outgrowth of Freud's own assessment technique of free association (Goldberg, 1975). Many clinical psychologists became quite proficient at describing clients in psychoanalytic jargon, using a maximum of inference and a minimum of data and were highly valued by their psychiatric confreres for this reason (see Yates 1970). Thus, although they did not presume to take over the psychoanalyst's couch, they did become pseudo-psychiatrists, "preaching and practising Freudians without encroaching upon the psychoanalysts' preserves" (La Pierre, 1959, p.46). The scientific status of Freud's theories and practice was evaluated in Chapter V and found to be wanting - again against the criteria adopted.

In retrospect, Bersoff's (1973) claim that psychoanalytic and psychometric theory provided a faulty foundation for the emergence of clinical psychology, more particularly psychodiagnostic practice, seems justified. The practice of psychodiagnosis, like any other area of technology, rests not only upon a varied assortment of learned skills, experiential expertise, and special procedures devised empirically from practice, but also upon various assumptions about the nature and function of the field. This latter domain of understanding can be thought of as the philosophical basis of assessment. Such assumptions are of tremendous influence in determining aspects of practice; yet they are rarely examined in detail.

The objective of this chapter is to undertake an analysis of the underlying assumptions and conceptual model of the traditional dispositional approach to assessment and to evaluate the empirical support for some of its central assumptions. This will lead to an
examination of the model of causality that is implicit to the approach. It will be argued that the quasi-medical model of abnormal behaviour represents the logical extension of the dispositional model to the explanation of grossly deviant behaviour. Lastly the way in which this approach to deviance leads to the medicalization of social problems will be discussed, with special attention given to the politics of clinical judgements.

THE ASSumptIVE BASE OF THE ATTRIBUTE MODEL

The theoretical structure underlying an approach to psychodiagnosis can be delineated in the form of a conceptual model that reflects the data to be obtained and the uses to which they are to be put. Traditionally, psychodiagnosis has been conceptualised in terms of what McReynolds (1970) refers to as the 'attribute model'. Both psychoanalytic and psychometric theory are encompassed under this label. According to McReynolds (1970) assessment theory typically has been developed and presented as if this were the only possible model; as if it necessarily reflects the fundamental nature of assessment. This, however, as Cronbach and Gleser (1965) have reminded us is not the case: indeed there are various possible assessment models that could be put forth, depending on the function that assessment is concerned to serve. Arthur (1966) was able to conceptualise five distinctly different models. There is little doubt, however, that the attribute model is the classical and conventional approach to assessment. The model is founded on "Darwin's emphasis on differences among individuals, and all theoretical work behind test scores has attempted to conceptualise differences in abilities and traits" (Cronbach, 1971, p.440). The current
dominance of the attribute model is illustrated by Kelly's (1967) introduction to his text, in which assessment is defined as "any procedure for making meaningful evaluations or differentiations among human beings with respect to any characteristic or attribute" (p.1).

The length of its lineage does not necessarily imply that the field has not been dynamic and innovatory. Goldfield (1977) suggests that one of the clearest barometers of the field's progress over the years is the ever-changing title of the journal specifically devoted to psychological assessment. The journal, launched in 1936 was entitled the "Rorschach Research Exchange"; eleven years later, primarily as the result of the expanded scope of assessment devices the journal title changed to "The Rorschach Research Exchange and Journal of Projective Techniques". Research and development on objective assessment devices led to the next name change and in 1963 the title became the "Journal of Projective Techniques and Personality Assessment". And finally, in the light of disappointing research findings on the reliability and validity of projective techniques the journal changed again, this time to its present title the "Journal of Personality Assessment". Goldfield speculates that the next name change may reflect the growing interest in behavioural assessment (1977, p.3).

Central Assumptions

Despite the apparent innovations in assessment evidenced above, the central assumptions remained protected and unchanged. Its assumptive base was probably first articulated nearly a hundred years ago by Galton (1883). Essentially the aim of the approach was to measure given attributes of a given individual with the purpose of
either

(a) assigning the individual to a given category; or

(b) placing the individual at some point in the continuum represented by the attribute. This enterprise is given meaning through subscription to the shared common assumption that what is measured are certain relatively stable and interrelated motives, characteristic traits and dynamics that are causally responsible for the individual's overt actions. Thus, in order to fully understand why an individual behaves in a particular way, one needs to obtain a comprehensive understanding or picture of the underlying states and traits.

From this vantage point, to simply observe and describe overt behaviour in various life situations is inadequate in that the causal essence of the person is deeper and more inferential than that which may be directly observed. Assessment must therefore focus on the structural or dynamic components assumed to make up the individual's personality structure. This may be done by means of pen-and-paper questionnaires or by projective tests that presumably enable skilled personnel to infer the latent content of the manifest behaviour. The implicit assumption in the administration of such tests is that the salient variable is the 'personality' of the examinee. Thus any differences in the Rorschach protocols, for example, produced by two different people are presumed to reflect differences in their personalities, preoccupations and concerns.

The model is predicated as an additional three assumptions that can be stated with respect to a given attribute of given individuals. Firstly, it is assumed that all individuals in some sense 'have' or can be characterised by the attribute, disposition or trait in
question. Secondly, it is assumed that individuals differ among themselves with respect to the attribute. Thirdly, it is assumed that for each individual there is a true placement (score) on the attribute continuum, or correct categorical label, which is only approximated by assessment techniques. In a general sense, the assessment judgement is composed of the 'true' characteristic of the individual plus an error component or deviation of the observed score from the true score. This represents a Platonic perspective on the philosophy of measurement, that is - to take the view that scale scores are imperfect representations of psychological properties that exist in a pure form elsewhere (in this case inside people).

Although Hogan, DeSoto and Solano (1977) suggest that few contemporary personality researchers would endorse this Platonic view, in practice the view, although inarticulated, seems pervasive. In fact in 1959 Kogen represented the relationship with the equation \( X = T + E \) in which \( X \) is the observed score or rating, \( T \) is the individual's 'true' characteristic and \( E \) is the difference between the observed and 'true' score. In line with this, the attribute model conceptualises the assessment device (tests, interview, etc) as a measurement instrument and stresses the importance of accuracy of measurement, emphasising particularly the concept of reliability and validity.

The attribute model does not assume that the traits and states being assessed are ephemeral or situationally specific. Quite to the contrary, the inferred dispositional attributes are seen to determine behaviour in a consistent, at least at a covert level, way over time and across different situations or settings (see e.g. Cattell, 1950, Sanford, 1963). Whether one uses the language of factors, or of habits, or of basic attributes, or of dynamics and
character stimuli, this fundamental assumption is pervasive. Our persistent belief in personality traits, and stubborn adherence to the view that there are pervasive cross situational consistencies in an individual, are ancient convictions (see eg Allport, 1937). And despite Hogan's et al. assertion that "no one believes that speedometers 'really' measure speed or that the Stanford-Binet Intelligence Scale measures intelligence, nor do we think that personality scales measure traits", (Hogan, De Sotto and Solano, 1977, ps.256-257) in practice this important distinction is lost.

Test Responses as Signs or Samples

The attribute model is characterised by what Mischel (1972) refers to as a 'sign' approach to assessment. This applies to the assessment of states and traits. Goodenough (1949) is accredited with first distinguishing between the 'sign' and the 'sample' approaches to test responses. She suggested that when test responses are viewed as signs, an inference is made about the performance as an indirect or symbolic manifestation of some unobserved characteristic. In the trait approach, test responses are viewed as direct signs of the personality trait, with the strength of the trait determined by the number of signs. Importantly, the trait label is considered not merely a descriptor, but rather is seen as an underlying determinant of the observed behaviour, i.e. the signs (Allport 1937, 1966). Of course, not all trait theorists ascribe a causal role to traits, some see traits as terms to denote stylistic consistencies in behaviour (see e.g. Hogan, De Soto and Solano 1977). But unfortunately again this important point is often lost in practice - resulting in reification.
In the more psychodynamic approaches to assessment, responses to the assessment stimulus are seen as indirect signs (or symbols) of an underlying personality organisation and dynamic forces. Any aspect of behaviour, for example expressed fears or obsessive thoughts, may be a sign that reveals important underlying conflicts but are in themselves of only trivial significance. Not only is the overt behaviour secondary, but it may be misleading, distorted and disguised by protective defenses.

Thus the primary objective of psychodiagnosis is the discovery and accurate measurement of the individual's underlying dispositional structure or pervasive motivational system, i.e. their essence (Wallace, 1966). Since these attributes are not directly observed but are inferred from the assessment performance either directly or indirectly, the language used to describe the individual is predominantly 'genotypic' (Stuart, 1970).

The Language of Assessment

Genotypic language is essentially concerned with explaining behaviour in terms of sub-surface factors where as the alternative language system, referred to as 'phenotypic', focuses more on describing the surface diversity of behaviour. That is, phenotypic language stresses what a person does in quantitative terms whereas genotypic language stresses what a person 'is' or 'has'. The question of what is meant by statements to the effect that an individual 'has' certain attributes (e.g. introversion, inferiority complex, high need for achievement) is a thorny one.

Mac Corquodale and Meehl (1948) in a useful discussion of
psychological inferences distinguished between two classes of inferred variables:

(a) those which are presumably potentially observable but are to date unobserved they referred to as 'hypothetical constructs', (i.e. traits in Allport's 1937 and Cattell's 1946 systems), and

(b) that other type of inferred variable they called an 'inter­vening variable' and is not seen as being even potentially observable - it serves purely as a conceptual convenience (i.e. traits in Gough's (1969) Holland's (1973) and Welsh's (1972) systems).

If these distinctions are not kept clearly in mind the ever present danger of reification may be realised.

In an attempt to ameliorate this danger Tyler (1965) distinguished between traits and 'dimensions' - with the former supplying differences in amount and the latter implying difference in distance on a scale. The danger of reification of concepts is a major disadvantage of the attribute model and is clearly enhanced by the use of genotypic language and the assumption of 'true' scores.

Qualitative Differences Between People

Genotypic assessment relies upon postulated qualitative difference between individuals. When applied to psychodiagnosis it is assumed that individuals who are 'mentally healthy' are qualitatively different from the 'mentally ill' and that the 'schizophrenic' is qualitatively different from the 'neurotic'. These qualitative differences are purportedly codified in the current Kraepelinean typology used by the mental health profession to classify clients. A major function
of psychodiagnosis is the accurate arrival at a categorical label for the examinee. It is imperative in this process to note that the dispositions are derived by inference from the behavioural data and they are not isomorphic with this data. That is, dispositions are logical rather than empirical entities.

Dispositional diagnosis draws both on 'signs' which are objective indicators of 'abnormality' and 'symptoms' which are indirect descriptions of subjectively perceived abnormalities (Holmes, 1946) in classifying clients. The traditional model of psychiatric classification relies almost exclusively on a genotypic approach which describes the inferential dynamics of the individual whose behaviour is problematic. The resultant categorical label presumably conveys information indicating the conditions associated with its origin (i.e. etiology), the planning of intervention or corrective strategies, and the generalised predictors of therapy and post-therapy behaviour. These assertions may be readily analysed as there is a plethora of empirical research concerning the validity of genotypic diagnosis. This will be undertaken a little later.

Reliability and Validity

Because of the assumptive base of the attribute model, the concepts of reliability and validity are part and parcel of the associated assessment theory. Reliability is related to the correlation between the obtained and true scores. In a sense, reliability is conceptually a synonym for consistency. Psychometric reliability means that the measurement device can be expected to consistently discriminate between individuals from one occasion to
the next. The variance of the test scores has a great deal to do with this type of consistency - the less the variance the more likely the discrimination will change across occasions. Psychometric variance is properly estimated in terms of error variance, product moment reliability coefficients, and standard errors of measurement. All of these statistics are dependent on variance, if there is no variance then by definition there can be no psychometric reliability.

Validity, on the other hand, is related to the correlation between the obtained score and some other estimate whether obtained or inferred, of the attribute. To evaluate empirically the psychometric validity of a device, individual differences in the device may be compared with individual differences on another variable that is assumed to be highly related to the attribute being measured. If the device discriminates between individuals in approximately the same manner as the criterion variable, then this is positive evidence of the validity of the device. These interpretations of reliability and validity are not wholly applicable to other models (see for example the debate between Nelson et al (1977) and Cone (1977) vis a vis their application to behaviour assessment). The highly sophisticated development of these concepts has done much to enhance the dominance of the attribute approach (McReynolds, 1971) but as Bersoff (1973) has indicated quite often the psychometricians appear to have spent too much time focusing on statistical niceties at the expense of evaluating the external validity of many devices.

DISPOSITIONAL EXPLANATIONS OF BEHAVIOUR

Implicit in the attribute model, at least as usually carried out, is a theory of human behaviour, namely, that behaviour is
essentially determined by intraorganismic factors. The attendant assumption is that if we understood the individual well enough, we would be able to predict her/his behaviour without detailed knowledge of the contextual milieu. Hence Mischel's criticism of the tendency in psychodiagnosis '... to use a few behavioural signs to categorise people enduringly into fixed slots on the assessor's favourite nomothetic trait dimension and to assume that these slot positions were sufficiently informative to predict specific behaviour and to make extensive decisions about a person's whole life' (1979 p.146). Although the conceptual and empirical adequacy of this assumption has been occasionally questioned (see for example Lehman and Witty, 1934; Newcomb, 1929; Thorndike, 1906; Vernon, 1964) the major figure in the current round of the debate has been Walter Mischel (1968, 1969, 1972, 1973, 1977, 1979).

A little over a decade ago Mischel published 'Personality and Assessment' in which, after reviewing both past and current research in a very considered way, he argued that personality attributes are constructs of the observer which may have little or nothing to do with generalised, trans-situational behavioural patterns of the observed. The stubborness of our adherence to the concept of personality traits reveals the pervasiveness of 'magical' thinking and its power in influencing our perception of reality. Everyday personality traits (e.g. dependent, aggressive, friendly) are not correlational patterns to be found in conduct: they are clusters of meaning evoked by conduct (Shweder, 1975).

Personality traits are symbols or interpretative categories that link together items of behaviour and are linked to each other by conceptual relationships that have little to do with frequency or
correlation (Shweder, 1977 p.452). Mischel (1968) went further to assert that the predictive utility of an attribute based approach to personality remained undemonstrated and that situational specificity of behaviour appeared to be the rule rather than the exception. Although evidence was available to signify temporal stability there was much evidence for the discriminativeness and ideographic organisation of behaviour patterns. Summarising his review, Mischel concluded 'Although behaviour patterns often may be stable, they usually are not highly generalised "across situations"' (1968, p.282).

Mischel's Critique of the Attribute Model

Although other contemporary authors have drawn similar conclusions (see e.g. Fiske 1974, Peterson 1968, Shweder 1973, 1975) it was Mischel who provoked the most controversy by arguing that the commonly observed +.30 cross situational correlation coefficient probably reflected 'true' behavioural variability rather than an artifact of imperfect methodology, as has often been claimed. It remains to be seen to what degree the erratic and uneven relationship, typically found when cross-situational consistency is studied, reflects methodological problems (as Block, 1977 recently suggested) or the actual discriminativeness of social behaviour across psychologically non-equivalent situations (Mischel, 1977). 'In my view, better measures will surely provide better support for the existence of meaningful organised behaviour patterns. It should be noted that discriminative behaviour and ideographic organisation implies neither chaos nor unpredictability'(Mischel, 1977, p.742).

In his early review Mischel (1968) clearly announced that the data was sorely lacking that would corroborate the notion that personality dispositions were important components of an individual's
behaviour which are consistent across situations. Since personality
tests like the MMPI, the CPI and the TAT are commonly assumed to
measure such traits, the relevance of these tests for the development
of psychological theory was questioned. Evidence marshalled by Mischel
was clearly an embarrassment to trait theorists (see below) but, as
we have seen, research programmes are relatively indifferent to
anomalies and contradictory evidence. The 'protective belt' has
staved off the lethal blows. Exactly when the accumulating refut­
ations will promote disenchantment appears to be an individual matter.
Several notable researchers have expressed the disillusionment, for
example, early on Peterson expressed his dissatisfaction in the
following way:

the generality of these (personality) measures over method
and situation was still not high enough to justify perpetuating
the traditional concepts of personality. The findings required
abandonment of a line of research to what I had devoted ten
years of my life as a psychologist. The results also required
a change in beliefs about the nature of personality. This
research, per se, did not say which way the conceptual shift
should go, but it suggested very strongly that traditional
conceptions of personality as internal behaviour dispositions
were inadequate and insufficient (1968, p.23).

As one would have expected, very few of Peterson's confreres
imbued in the traditional paradigm allowed the data to induce
similar renunciations. Rather the perceived attack on personality
gave rise to a lively 'between paradigm' debate (as e.g. Alker, 1972;
Bem, 1972; Block, 1977; Wachtel, 1973) and an acceleration of relevant
research (see e.g. Bem, and Allen, 1974; Epstein, 1979; Olweus, 1977).

It is well beyond the scope of this chapter to review the contemp­
orary debate, or even to attempt to adequately summarise it. The sub­
sequent discussion will instead restate Mischel's central argument
and briefly discuss the major sources of data. A more detailed
review of the reliability and utility dispositional diagnosis will be undertaken in the next chapter.

Mischel's original book (Personality and Assessment, 1968) was often mistakenly perceived to be a situational manifesto aimed at undoing the role of dispositions. This was far from his intent. His self-professed aim was to 'defend individuality and the uniqueness of each person against what I saw as the then prevalent form of clinical hostility' (Mischel, 1979, p.742). To do this it was necessary to cogently document the flimsiness of evidence available to support the vitality of ascribing global dispositional labels. The central finding was that, with the exception of cognitive variables, the utility of inferring personality disposition from behavioural signs is underscored by the data. To recapitulate:

Response patterns even in highly similar situations often failed to be strongly related. Individuals show far less cross-situational consistency in their behaviour than has been assumed by trait-state theories. The more dissimilar the evoking situation, the less likely they are to produce similar and consistent responses from the same individual. Even seemingly trivial situational differences may reduce correlations to zero (Mischel, 1968, p.177) ... with the possible exception of intelligence, highly generalized behavioural consistencies have not been demonstrated, and the concept of personality traits as broad dispositions is thus untenable (Mischel 1968, p.146).

Empirical Evidence for the Consistency Position

The attack on the assumed cross-situational consistency of state/trait theories rested primarily on three major sources of empirical evidence. Most important was the large number of studies that were unable to produce correlation co-efficients greater than +.30 when behaviour in one situation was correlated with behaviour
in another situation. As Mischel (1969) noted '... a correlation of +.30 leaves us understanding less than 10% of the relevant variance. And even correlations of that magnitude are not very common and have come to be considered good in research on the consistency of any non-cognitive dimension of personality' (p.1012). Certainly to predict individuals with reasonable accuracy, correlations of the vicinity of .80 or .90 are required.

Although a few studies have reported correlations of such magnitude when behaviour was averaged over many days (see e.g. Epstein 1979) the vast majority of studies have yielded meagre correlations. So consistent are the correlations less than .30 that Bem (1972) contends that the issue of stability in personality cannot be resolved by dispute but only by data: specifically, correlation coefficients in excess of .30. He suggested that when personologists are able 'to predict behaviour across situations better than +.30, Mischel will fold up his tent and steal away' (Bem, 1972, p.18). Bem went further to suggest that:

There was nothing silly about the initial assumption of personologists that everything was glued together until proved otherwise. But since it has now proved otherwise it seems only fair to give a sporting chance to the counter-assumption that nothing is glued together until proven otherwise. Instead of assuming cross-situational to be +1.00, let us begin by supposing them to be 0.00 until we can explicitly construct them to be otherwise (Bem, 1972, p.28).

A second source of evidence consists of findings from the apportionment of variance in analysis of variance designs. This procedure was independently introduced into the assessment of stability of personality research by Raush (1965) and by Endler and Hunt (1968, 1969). The research strategy purports to yield the relative separate quantitative contributions of the person and
situation, as well as the person and situation interaction. Early results of this line of investigation suggested that the variance attributed to individual differences was usually much smaller than the variance attributable to the situation and to the interaction of the situation and person (Endler and Hunt, 1968; Moos, 1968, 1969). Later, however, in an influential position paper Bowers (1973) challenged this conclusion but was forced to concur that too little of the total variance was due to the person to justify a thorough-going trait position.

It may be pointless to analyse the research further as it is now evident that the procedures for partitioning variance are so severely flawed that they render interpretation meaningless (see e.g. Golding, 1975; Olweus, 1977). In addition to the misuse of analysis of variance in the studies, the very nature of the question is questionable and more likely to create a pseudocontroversy that pits the person against the situation, than clarify the issue. The assumption is implicating the question in direct contrast with Mischel's (1968) aim of calling attention to the specific reciprocal interactions between persons and contexts. It also obscures the fact that the answer must always depend on the particular situation and person sampled. Presumably studies could be designed to demonstrate almost any answer - this was Moos' (1972) contention. He recognised that a study could be designed so that:

any result is possible. I think that all one can say is that given relatively real life situations (e.g. patients in wards, or in outpatient psychotherapy, or your (Mischel's) delay of gratification studies) that the major proportion of variance simply does not appear to be accounted for by individual difference variables. One could certainly, however, easily design studies in which the major portion of the variance would be accounted for by individual differences variables.
Frankly, this is why I have stopped doing studies of this sort. It seems to me that the point has now been amply demonstrated, and it is time to get on with other matters (Moos cited in Mischel, 1973, p.171).

The last source of evidence to be cited is the tendency to attribute more stability to individuals across situations than is objectively warranted (see e.g. Jones 1979; Shwedler 1975). Additionally, earlier reference was made to as the 'fundamental attribute error: a tendency to underestimate the importance of situational determinants and overestimate the degree to which actions and outcomes reflect the actor's disposition' (Ross, 1977, p.193-194). There are many persuasive reasons postulated by person perception researchers to explain the apparent stubborn adherence to untenable beliefs. Rather than briefly introduce the relevant research here, it will be discussed in Chapter VIII. It should be noted, however, that the presence of overestimations of stability does not establish that there is no stability in behaviour apart from such bias.

The Categorization of People

It is probably timely to recall that in marshalling negative evidence Mischel's aim was not to demonstrate an absence of stability but rather that specificity is the case; that specificity more accurately reflects our impressive discriminative faculty. Mischel (1968) argued not against the concept of traits but that, when the relationship between the observed behaviour and the attributed trait are relatively direct, traits serve essentially as summary terms for the behaviour that has been integrated by the observer. As such they may serve a useful purpose. Global characterisations of salient personal qualities and broad, highly abstracted categories may be useful with minimal moderations or specific situational
qualifiers. But Mischel was adamant that: for the purpose of more specific communication; for explaining the phenomena of personality; for making statements about individual behaviour; and for predicting specific behaviour in relation to specific conditions, careful discriminative limits must be included.

Eleven years after the publication of 'Personality and Assessment' Mischel contends that 'although temporal stability in the patterning of individual lives, in self-perceptions, and in how others view us is not in dispute, there is serious disagreement about the nature, degree, and meaning of cross-situational breadth of behaviour assessed by objective measures of the behaviour as they unfold (1979, p.742). The challenge remains essentially unchanged, namely, that global dispositional labels are inadequate causal explanations and lack clinical utility. Obviously "when the consistency issue is viewed in terms of the utility of inferring broad response tendencies from behavioural signs and not in terms of more metaphysical questions of the existence or validity of personality dispositions, the answer must depend upon the particular objective or purpose for which the inference is made" Mischel's original treatise, is oriented primarily on current psychodiagnostic assessment practices. It was in this context that Mischel cautioned that even though the statistically significant relationships found in personality research are:

"sufficient to justify personality research in individual and group differences ... their value for making statements about an individual are severely limited. Even when statistically significant behavioural consistencies are found, and even when they replicate reliability, the relationships usually are not large enough to warrant individual assessment and treatment decisions except for certain screening and selection purposes" (1968, p.38).
Even though there are occasional discussions of the difference between clinical and statistical differences in the published literature (Lick, 1973) the importance of this topic is not apparent in much discussion of research. Many investigators appear content to rest on their statistical laurels and not to worry overly about the actual practical value of their results. This difference is at the heart of Mischel's critique, and the more general assault on the utility of dispositional approaches to personality that has come to be known as the 'situationalists' position. Both these critiques have been much misunderstood.

The critiques generally do not imply that people show no consistency, that individual differences are insignificant, or that situations are the most important determinant of behaviour. The real issue is the clinical utility of decisions and predictions based on global trait/state inferences. The critique questions the traditional personality paradigm that promotes traits/states as the intrapsychic causes of behavioural consistency (and occasional phenotypic inconsistencies). It is argued that such a perspective obscures the impressive discriminative faculty and sensitivity of individuals to the subtle nuances of their contexts.

THE QUASIMEDICAL MODEL

It has been stated several times that Mischel's critique of the attribute model of assessment was provoked by the aversive consequences of its application (see e.g. Stuart Chap 5, 1970) to the identification and categorical labelling of individuals who threaten the normative penumbra. In this section the implications
of construing deviance from the perspective of the attribute model will be delineated. Or, alternatively, the attribute model's answers to the question "who is normal"? will be discussed. The threatening antiquity of mental disorders (who is mad? Who is sane?) leads us to take our system of perceiving 'mental illness' for granted when it is just that system which should be the object of study since it defines our experience of 'mental illness'.

The Clinical Definition of Normality

Because the attribute model represents a union of psycho-dynamic and psychometric beliefs in the one system, not surprisingly it provides two definitions of normality. The first answer to the question of who is mad comes from the disease theory perspective and was primarily the result of Freud's influence. The second answer, was contributed by the psychometricians and is based on the statistical probabilities of differences occurring in specific characteristics. Unfortunately, the two approaches are quite often used interchangeably and clinicians may think in one system while operating in the other with unfortunate consequences for the client.

(a) The Pathological Model of Normality

The disease theory or pathological model of normality was developed in medicine as a conceptual device for comprehending and controlling disease processes and organic malfunction. Medical concern has traditionally been aroused when conditions occur which interfere with biological functioning and homeostatic processes of the organism. Consequently, the focus of the pathological model is
the presence of pathology and its removal. Bordua (1967) has suggested that the model appears to have six principal elements:

1. There is an underlying pathological condition in the sick person which can usually be diagnosed by a syndrome of physical symptoms;

2. An effective or ameliorative treatment is known for many types of cases;

3. Treatment involves doing something to the sick individual;

4. The disease condition may get worse, leave permanent damage, or take an inordinately long time to return to a healthy condition without treatment;

5. In most cases the treatment has little or no harmful side effects; and

6. A principle of medical diagnosis asserts that it is generally preferable to treat a well person erroneously diagnosed as sick than to leave a sick person untreated.

The implications of these assumptions will become apparent as discussion proceeds.

A clinical diagnosis based on the pathological model begins with an abstract concept of the nature of symptoms that constitute a particular disease syndrome, and then detection of whether the
pathological symptoms are present in the individual case. If they are, the person is sick, if not, then the person is well. The pathological model is thus conceptually a bipolar construct. At one pole is normal, which is equated with health and absence of pathology. At the other pole is abnormal, defined by the presence of pathological signs and equated with disease. To be abnormal is unhealthy, or in the light of the principles listed above, is 'bad' and should be prevented or alleviated. To be normal, that is healthy, is 'good' and should be sought after and maintained. It can be seen from this that the pathological model is essentially evaluative.

(b) The Statistical Model of Normality

In contrast a non-evaluative definition of normal is provided by the statistical model. Unlike the pathological model which defines the symptoms of pathology by functional analysis, the statistical model defines abnormal according to the extent to which an individual attribute varies from the mean of a particular population of measurement. The definition rests on the assumption that an individual's attributes can be gauged by his/her relative position in a frequency distribution of other persons. Consequently normality is defined as the zone range of one standard deviation either side of the arithmetic mean - this comprises approximately 68% of the population.

Again, unlike the bipolar pathological model, the statistical model defines two types of abnormal - those who have unusually high and low measures of the characteristic in question. Thus there are gradations of abnormality depending upon the frequency
of occurrence or distance from the mean. The model is non-evaluative or neutral in that the desirability or otherwise of being abnormally high or low depends on the characteristic being measured and is socially arbitrated. The statistical model is not restricted to establishing norms for biological characteristics but is also used to establish behavioural norms.

Initially the pathological model was directly applied to the understanding of deviant behaviour, as it had been earlier by the ancient Greeks and Romans (Rosen, 1968), but by the end of the nineteenth century it was becoming apparent that psychopathology could not be adequately accounted for in terms of the prevailing model of underlying organ disease. To satisfactorily account for many of the troublesome behaviours for which no organic cause could be found, Freud reformulated the medical model. Retaining the central notions of an underlying disease to which the symptoms can be attributed, Freud suggested that the disturbed behavioural symptoms were the result of psychological disorders. The supposed causal agents underlying 'abnormal' or deviant behaviour in this 'medical analogue' theory (Ullman and Krasner, 1975) is a set of diseased personality attributes (Buss, 1966).

We have already recorded that this extrapolation of the 'genuine' medical model, which utilizes the model but not its content has been called the intrapsychic or quasimedical model. This extrapolation and reasoning by analogy, in which the model is taken for empirical truth, without considering how well it applies almost inevitably results in the generation of ideology (Leifer, 1969).

The history of the transplantation of the medical model into
the field of problematic behaviour has been described by Ullman and Krasner (1965, pp.2-45) and others (Foucault 1965, Rosen 1968, Rothman, 1971) and need not be repeated here. That the model worked well in the treatment of physical ailments is probably the chief reason for its ready acceptance by psychiatrists and clinical psychologists. However, the influence of psychoanalysis and the early dominance of the clinical arena by medically trained psychiatrists should not be under emphasised. Irrespective of the reasons postulated, the end result was the "transplantation of concepts and methods of proven usefulness in physical illness into the fields of 'mental illness'. Thus, disorders of behaviour have been regarded as diseases for which an etiology must be found, which will ultimately lead to a specific treatment. Hence the stress on the importance of diagnosis and the belief that specific causes would be found for specific mental illnesses" (Yates, 1970, p.4).

In order for the model to be legitimately applied to clinical psychological practice, it must be shown to be appropriate and useful. "There is", according to Stuart (1970, p.7) "good reason to believe that it is neither". Certainly, a number of the assumed parallels between physical and mental illness simply do not materialise (see e.g. Cowen, 1973; Schofield, 1964). And although the treatment of 'mental illness' by psychotherapists and psychoanalysts has been nominally isomorphic with this medical model, in practice, some elements of the model are inappropriate for dealing with social behaviour (Turner and Cuming, 1967).
The Transformation of Behavioural Differences into Signs of Pathology

The problems elucidated above are compounded when the genuine or quasimedical model is used in conjunction with the statistical model and the two definitions of abnormality are coupled, as in, for example, the official nomenclature for defining mental retardation (Braginsky and Braginsky, 1971; Mercer, 1973). When a diagnostic evaluation is conducted using both models as is the case in most psychodiagnostic situations - there is a tendency to think in terms of one model while assessing within the other. That is, behavioural differences (particularly relatively infrequent behaviour) may be illicitly translated into a pathological sign. An example of the implicit logic that underlies this transformation is spelt out by Mercer (1973, p.6) as follows: Low IQ is undesirable and 'bad' in Western cultures; Bad equals the presence of pathology in the medical model; thus a low IQ indicates pathology. Through this implicit process IQ which is not a biological manifestation but a behavioural score becomes conceptually transformed into a sign of pathology carrying all the implications of the disease theory outlined above. Psychodiagnosis is replete with examples of this illicit and faulty reasoning that results in the transformation of differences into signs and symptoms. It is almost inevitable given the two definitions of normal outlined above which are intrinsic to the attribute mode.

An important result of the existence of these two definitions is the ever present opportunity for psychodiagnosticians to serve politically conservative ends by taking the behavioural norms of the dominant culture (WASP) and construing gross deviation from
these as signs of pathology. This is only possible if psychologists fail to acknowledge that the statistical definition of normal is not trans-society (as the medical model is). When the statistical model is used to define normal performance, the emergent norms can not be legitimately generalised beyond the population sampled. Equally important is the requirement that the characteristic being measured is distributed normally. Not infrequently psychologists should remember Boring's (1920) assertion that there is nothing normal about Normal Distributions - that they don't inhere in nature, or as Simon (1968) conjectures "its occurrence [Normal Distribution] is entirely caused by the researcher and its appearance 'means' that the researcher may consider that his(her) research work is complete" (p.436). And lastly, when the statistical model is used, if a minority group is included in the sampled population they will be abnormal - it is intrinsic to the model. Whether this abnormality will be valued positively or negatively, and whether it will be taken as a sign of pathology will depend on socio-political factors.

FROM SINNER TO SICK DEVIANT

The attribute model by combining together two definitions of normal - one evaluative and trans-societal and the other neutral and population specific - readily lends itself to the medicalization of deviant behaviour (Conrad, 1975; Divoky and Schrag, 1975; Szasz, 1971). Attendant with this medicalization of social problems is the rise of a rationale of 'treatment' in reaction to deviant behaviour; this phenomenon has been referred to as the therapeutic state (Kittrie 1971, Szasz 1970). Examples of the
pervasiveness of this ideology can be found in all areas of social deviance. Kurpas for example states that "The criminal is a sick person, crime is a disease; a symptom of mental aberration" (1954, p.218) and similarly Lindner asserted that 'crime is not an act against the law, nor against 'nature', nor necessarily anti-social but "crime is a symptom, actually expressed, of internal maladjustment and conflict" (1946, p.38). Designating undesirable conduct as illness rather than crime has been a major hallmark of this century's rise of the therapeutic state. The measure of the transition from a penal to a therapeutic model is dramatically illustrated in a report which concluded that "racism is the number one public health problem facing America today" (cited in Kittrie, 1971 p. XVI).

The Social Impact of a New Label

The question of whether a person who, say, transgresses a civil law should be called criminal or mentally ill is not merely an academic question. Kittrie evidences this by stating: "The clients of the therapeutic state increasingly recognise it as a two-edged sword. They fear that the label of treatment engenders in the public as much suspicion and hostility towards the one who is being treated as does the criminal label, they fear that in the name of therapy society seeks to impose controls over people and behaviour that should be free of societal intervention; and they fear, finally that the therapeutic state possesses tools of human control that far exceed in their threat to individual liberty the sanctions possessed by the criminal model (1971, p.XVII).

On this score Sarbin offers an interesting observation, which fits in nicely with Scheff's (1966) notion about the reactions of
others helping to stabilize mental illness. Sarbin (1967) notes that "... because of the inherent vagueness in the concept of mind, its assumed independence from the body, and its purported timelessness (derived from the immortal soul), there is a readiness to regard this special kind of sickness as permanent" (p.451). It does appear that people react to those labelled mentally ill as if they are permanently disabled. This is attested by the recent Eagleton affair. Eagleton was proposed as vice-president of the United States by McGovern only to have his nomination withdrawn after evidence of three treated occurrences of depression was reported. The disabling aspects of Eagleton's history was not so much that he had been depressed, but that he had seen a doctor who 'treated the mind'.

Along a similar line, Szasz (1970) argues that in criminal cases, when a successful insanity plea results in commitment to a mental institution, the deviant would have been better off not to make that plea - if being better off is spending less time incarcerated. Reinforcing this assertion Chu and Trotter (1974) revealed that in 1965, 94% of all prisoners in federal prisons of the United States had been incarcerated for five years or less. In contrast, over half of the patients of a large mental hospital surveyed had been inpatients for 10 years or more, and 37% had been inpatients for 20 years or more. Further, the evidence indicates that 'mentally ill' persons are still regarded quite negatively by lay persons, therapists and other psychiatric patients (see, for example, Rabkin, 1977; Wills, 1978).

Notwithstanding the view expressed above, therapeutic correction has been advocated by professionals not only for formal
law breakers but for drug abusers, school truants, homosexuals, political activists, lonely homebased parents, prostitutes ... the list seems infinite.

The 'Insane' as Patient or Victim

Today there is an active contest between two fundamentally different and conflicting approaches to explaining grossly deviant behaviour. On the one hand there are those who subscribe to the attribute model with its emphasis upon the individual nature of mental illness. From this perspective the 'mentally ill' are construed to be the carriers of a 'disease' that leads them to become patients. On the other hand there are those that subscribe to a psychosocial perspective (variously called 'labelling theory' the 'societal reaction model', 'interactionist theory') with its emphasis upon the communal nature of mental illness. From this perspective the mentally ill are seen as victims of external contingencies that lead them to be ascribed a new social role as an institutional inmate or as psychiatric patients. (See e.g. Goffman 1961; Perrucci, 1974; Scheff 1966).

Superficial reading of historical accounts of social reactions to mental illness suggest genuine progress from prescientific periods of ignorance and inhumane treatment justified by recourse to supernatural and demonological theories, through the enlightenment to the modern scientific era of informed and humane treatment. Most historical accounts emphasise and dwell upon the different ways in which the mentally ill have been treated in different periods. However, most histories of abnormal psychology are written from the viewpoint of the medical model, they give detailed descriptions of
the long struggle of medicine to gain its 'rightful' dominance in the field of 'sick' behaviour (see for example, Spanos 1978; Zilboorg and Henry, 1941). Although our beliefs about mental illness have changed dramatically, the change is not related solely, or even predominantly, to the history of medicine but rather to the history of culture, societies, theologies, philosophies, jurisprudence, politics, economics and the sciences (see Foucault, 1965; Szasz, 1961, 1971; Ullman & Krasner, 1975).

When one looks beyond the assertions of change and progress there appear to be relatively consistent and stable features of 'madness' in a social context. The very obvious difference between the ancient practice of expelling unwanted mad people from the community and the modern practice of committing unwanted mad people to mental hospitals should not blind us to the similarity of the ultimate objective - to cast the mentally ill out of the community and to separate them from their society.

The Social Construction of Mental Illness

A close examination of the historical accounts of our response to abnormal behaviour justifies the assertion that the concept of abnormal behaviour is a human creation and that models, labels and deductions are human acts that are maintained because of social contingencies. A model or paradigm will, as we have seen, continue to be used until it breaks down under the burden of inconsistent data or until a new paradigm is demonstrated to be more effective at solving problems. Sociological analysis has highlighted the ways in which modern psychiatry 'reflects' the values of a larger society
(Bastide, 1972; Opler, 1967). Others have more explicitly depicted how psychiatry and the mental health establishment fulfil a social control function in contemporary society (Goffman, 1961; Myers & Bear 1968; Scheff, 1966). The various ways of thinking about psychiatry pursued by these authors suggest that the study of mental illness should not exclusively focus on cataloguing differences between the healthy and ill or between different kinds of mental illness, but should also focus on the field of psychiatry itself which is responsible for the definitions and treatment of mental illness.

In part, the intuitive appeal of the historical accounts of the rise of scientific psychiatry and its inadequacy seem to arise for the same reason: they are based on a frame of reference for conceptualising behaviour that is taken for granted and is therefore unexamined. In psychiatry formulations, the everyday social reality of the literate, industrialised, modern person is usually unexamined. Many social rules are so embedded in our every activity that they 'go without saying'. Behaviour is more or less automatically deemed to be symptomatic of some mental aberration to the extent that it is not readily understandable in terms of this socially constructed reality (Coulter 1973). That is, there are some rules that seem as much part of our conceptions of reality that violations of these seem strange and uncanny. Partly because such rules are central to our conceptions of what reality ought to be, our vocabularies for labelling violations of these rules are relatively underdeveloped. Scheff (1966) refers to the breaking of these rules for which we have no clearcut labels as 'residual rule-breaking' and argues that we tend to lump various sorts of residual rule-breakers into the category 'mentally ill'. 
The Latent Function of Psychiatric Help

In our enlightened, humane society it would be totally unacceptable to just expel the unwanted mentally ill - instead we help them. But the salient question is, is the ultimate objective of both strategies the same? On this point Thomas Szasz (1973, p.97-98) asserts that:

"Mental illness is a myth whose function is to disguise and thus render more palatable the bitter pill of moral conflict in human relations. In asserting that there is no such thing as mental illness I do not deny that people have problems coping with life and each other."

The moral dilemma is resolved by converting the deviant into a sick person and the agent of social control into a helper. Parsons (1951) refers to illness as not merely a 'condition' but also a social role and as Stainbrook (1959) noted our society has transformed the locus of the sick role from the family to the centrally located hospital. He asserts that today "the modern hospital has become increasingly the social system in which the sick role as a meaningful mode of participating in society is lived out". (Stainbrook, 1959, p.152). Our contemporary image of the mentally ill person and the helping role of the healer are superimposed on an existing role structure which had been developed to deal with physical rather than behavioural problems.

Within the illness model behaviour may be discounted, and neither the patient nor the labeller need be responsible. Secondly, if the person is sick, his/her incarceration may be called his/her own good and the 'good' of society. The use of sick role concepts permits social control over deviance without guilt concerning deprivation of civil liberties (Szasz, 1963, 1965; Leifer, 1969). Lastly, the latent socio-political function of the model is highlighted by Leifer
(1971) who states that:

"To abandon the medical model and view mentally ill persons as social deviators instead of diseased would expose the structure and quality of our social life to criticism. To say that so-called mental illness is one of the most important public health problems today ... means that there are many more deviating persons in our society than we have publicly acknowledged and it means that there are many more unhappy people than we have admitted."

In a thought provoking lecture Thomas Szasz (1976) addressed the question: "Medicine - Cure or Control?" He asserted that ultimately what justifies treatment in physical illness is not disease but consent. Therefore, if the mentally ill were 'really' ill - in that 'real' pathogenes were detectable - then a similar posture would apply to them. That is, the right to refuse treatment. As things stand in contemporary society, the hypothesised presence of mental illness may justify involuntary diagnosis, hospitalisation and treatment, whereas someone suffering from a physical illness, even a lethal one, is free to disregard medical advice, refuse treatment, and risk his own life in the process. Yet Birley (1973) in contrast contests that:

"Compulsory psychiatric care is not a threat but a right. Every citizen should have the right to be admitted against his(her) will to the care of a first-class psychiatric service".

It was pointed out earlier that Szasz alerts us to look at the relationship between the explainer (psychodiagnostician) and explained (client) and asserts that "explaining a person's behaviour against their will, especially when one holds him(her) in contempt, is, albeit ostensibly an explanation, actually a metaphorical confinement ... Confining by means of a contemptuous and degrading imagery" (Szasz, 1974 audio tape).
In his earlier work Szasz (1961) highlighted the abuse of ordinary language has suffered as a result of the medicalisation of deviance. He argues that as a result the language of the psychiatrist is no longer serviceable for the proper description of abnormal behaviour and our reaction to it. The language is being debased by systematic fraudulence, by the overwhelming effort on the part of the mental health establishment to impose its own image of the world on others, and by justifying any means used to achieve this end (for an excellent example of professional dominance of a definition of normality, see Winkler, 1977).

By way of example of the way the language of the medical model is used to differentiate ostensibly identical processes Szasz asks: Who defines grand-mal seizure a disease (epilepsy) and electro-convulsive therapy (artificially induced grand-mal seizure) a treatment? Who defines heroin taking as a disease and methadone (a synthetic opiate) taking a treatment? And who defines a surgeon operating on a perfectly healthy brain a psychosurgeon? Szasz argues that the fundamental problems of psychiatry centre around a struggle for definition. For example, the 'patient' says he is Jesus; the psychiatrist says he is not Jesus, but a schizophrenic. Although the contest sometimes looks like a debate, it is actually a bitter fight, and like all such struggles, it is decided not by logic, but by power. Terms like mad, mentally ill, etc refer to a particular form of ongoing relationship. Laing and Esterson (1964) make this point in their presentation of the cases of supposed schizophrenics - in each of their cases the actual behaviours of the individual diagnosed as schizophrenic make eminent sense when viewed in relation to the covert and overt behaviours of the individual's entire
family. With reference to the interview material collected from the family of one of their patients, Laing and Esterson (1964) comment that, once the attribution of illness is made, it comes "... to be taken as a fact, and ... she is treated accordingly. Such is the spell cast by the make-believe of everyone treating her as if she were ill, that one has constantly to pinch oneself to remind oneself that there is not evidence to substantiate this assumption, except the actions of the others" (p.203).

It appears that the behaviour of an individual engaged in social interaction cannot be interpreted except as part of a system, a system which is in part composed of the behaviour of the other person or persons with whom the individual is interacting. Once this is clearly understood - that a behaviour can only be studied and interpreted within the interaction or context in which it occurs - then, as Watzlawick et al. (1967, p.46) point out, terms such as sanity and insanity or mental illness and deviancy become meaningless as attributes of individuals. Or, to put it another way: "Behavioural events do not speak for themselves; they evoke meanings that are not to be found in the behaviour itself. The meaning may seem to be inextricably part of the event once an interpretation is made, but this phenomenological illusion of inherent meaning should not mislead us into thinking that meanings are discovered. The temptation to confuse one's interpretive categories with the events they describe is at the basis of magical thinking ..." (Shweder, 1977, p.455).
CHAPTER VII: STEREOTYPING AND PSYCHODIAGNOSIS

As we have seen, and as Claire (1976) has clearly indicated:

"The concept of mental illness appears to permit a bewildering number of interpretations. Is it a label for socially unacceptable behaviour behind which the deviant is permitted to take refuge and thereby evade the consequences of his antisocial activities? Is it an arbitrary concept which only serves to mislead people, by virtue of its medical connotations, into believing in 'mental sickness' when, more often than not, what it describes consists of disordered, interpersonal relationships wherein one person is scapegoated to carry the responsibility for the disturbances of the group? Is it merely a political expedient which enables those who hold power within society to devalue and degrade the dissenter and, by defining him as mentally ill, to violate his freedom and destroy his dignity? Or is it a concept, analogous to physical illness, which is applied to a patient who manifests not physical pathology but 'psychopathology', who experiences psychic rather than physical suffering, exhibits disturbances in his psychological rather than his physical functioning, and who, in some instances at least, suffers a serious impairment of his judgemental capabilities and his personal responsibility?" (p.1).

In recent years, several authors have analysed the relationship between psychiatry, psychology and mainstream values (Haley, 1967; Leifer, 1969; Sarbin, 1967) exploring especially the 'correctional stance' towards deviance. Laing (1967) and Szasz (1960) have contended that the essential function performed by the mental health establishment is a political one; that of preserving the status quo by discrediting persons whose behaviour does not conform to social expectations. Still others (Goffman, 1967; Halleck 1972) have stressed the moral political outcomes of the determinations rendered by psychologically oriented administrations and practitioners. Similarly, Mowrer (1960) has suggested that the latest function of therapy is moral re-socialization.
Over the past decade, the value ladder nature of the mental health enterprise has been starkly illuminated by the invective of the 'anti-psychiatry' movement. The central belief of the 'movement' is that mental illness is a reductive smear that obscures and defiles the despairing cries of the downtrodden and exploited against an alienating and dehumanized society. Psychiatric intervention is portrayed as a violent assault perpetrated under the guise of treatment, and the psychiatrist is deemed to be an agent of the dominant political order, and an agent of repression and of power. Anti-psychiatrists demand the abolition of existing psychiatric institutions and insist that psychiatrists either acknowledge their true role as society's thought police or become agents of personal and social change.

Given the fractious and acrimonious debate which has raged over the interpretation of the concept of mental illness it is not surprising that it has been difficult to reach a reasoned, consensually agreed to view on the function of psychodiagnostic assessment and classification. However, today, probably as a result of the debate, few mental health professionals would argue that psychodiagnosis and treatment can be value free but few subscribe to the 'Szaszian' concept of the latent socio-political function of the enterprise and the clinician's covert social mandate. Certainly few would agree with the Braginsky's (1974) assertion that psychodiagnosis is nothing more than a translation of social stereotypes and prejudices into the scientific sounding labels of mental illness.

However, we have seen earlier that perceptions of clients develop rapidly and stabilize somewhere between the first and
fourth interview (Meehl, 1960; Parker, 1968). And that there is much anecdotal and experimental evidence indicating the contempt and derogatory attitudes held by many of the 'helping profession' towards their clients (Rabkin, 1972; Rosenhan, 1973; Wills, 1979). It will be argued that a picture emerges which supports the Braginskys' assertion that psychodiagnosis represents a transformation of stereotyped thinking and social prejudices into clinical descriptions that then serve as the basis for 'diagnosis'. In this way the 'attribute model' converts undesired 'differences' into 'deficits' and 'symptoms' and facilitates the medicalization of everyday deviance.

By translating a society's important stereotypes, under the guise of science, into categories of psychopathology the mental health profession not only legitimates arbitrary stereotypes as scientific knowledge but also assists society maintain its normative structures by identifying and labelling those who threaten the boundaries. As the Braginskys point out:

"The diagnostic enterprise does not end with the assignment of labels. Indeed, the label is but the first step in helping to keep societies' house in order. The identification and classification of deviants is not an academic exercise, but, instead, starts the rather elaborate process of social sanitation of removing the deviant from mainstream society." (1974, p.132).

But rather than debate the issue further, in the discussion to follow empirical evidence of the value bias in the drawing of clinical inferences will be presented. This aspect of clinical judgement is typically unacknowledged. Clinicians are seen to base the decisions solely on impartial and 'objective' evaluation of relevant client behaviour. It will be argued that assessment and categorisation is an interpersonal process in which many non-
relevant aspects of the client inevitably influence clinical judgement.

THE POLITICS OF PSYCHODIAGNOSTIC LABELLING

To put the point bluntly - psychodiagnosis, even in extreme cases, is susceptible to distortion as a function of extraneous personal attributes such as sex, social class and political ideology. Several writers (c.f. Szasz, 1970, 1971; Scheff, 1966) have extended this charge to the commitment process, suggesting that it is equally influenced by all sorts of irrelevant interpersonal and intrapersonal variables, resulting in, as Chu and Trotter (1974) phrase it:

"In general, admissions to State mental hospitals come from a large 'residual population' made up of the poor, the aged, the abandoned, the members of minority groups, and others who are brought for psychiatric treatment not because they have been diagnosed according to any medical or psychological criteria but because they have disturbed, bothered or shocked the sensibility of someone or some group". (p.43).

The Influence of Task Irrelevant Client Characteristics in Psychodiagnosis

Demographic studies have clearly shown that 'lower class' people are far more likely to receive severely pathological diagnosis for conditions which are diagnosed as less severe in middle class patients (Dohrenwend & Dohrenwend, 1969; Holling & Redlich, 1958). Analogue research also clearly attests to this assessor bias.

'Lower class' individuals are clinically devalued more constantly than are other minority persons, women or social non-conformists (Abramowitz, Curtiz and Dokecki, 1977; Efren, 1970; Lee, 1968). Although there is evidence of diminution in the size of this effect (Trachtman,
A recent study still evidenced a sizeable bias (Di Nardo, 1975).

A second set of client characteristics that may be influential in determining the diagnostic conclusions drawn about the client are his/her attitudes towards being examined or tested (Wilcox & Krasnoff, 1967). The Braginskys (1974) cite some of their unpublished data which indicated a strong reversal from negative to positive impressions of clients following the utterance of statements to the effect of the helpfulness, kindness, competence, etc of the examiner. And in reverse, a spectacular change in the assessor's perception of the client occurred when the client directed derogatory remarks to the mental health assessor. Individual differences in the tendency to say what is socially desirable (Edwards & Walsh, 1964) and to seek the approval of others (Crowne & Marlowe, 1964) no doubt produce bias in assessment. More subtle influences have also been found to be operating. For example, Couch and Kenniston, (1960) showed that the general tendency to say 'Yes' and 'No' may influence test results.

Other extraneous variables thought to influence clinical judgements are the sex, and race of the client. Although caucasian clinicians have for the most part shown immunity to negative racial expectations, both analogue and correlational data offer evidence of a negative halo (Satter, 1970; Siegel, 1974). In a recent study Abramowitz and his colleagues (Abramowitz, Abramowitz, Jackson and Gomes (1973)) found left oriented politically active women were unfairly stigmatised in clinical evaluations. This lends experimental credence to the claim of women liberationists regarding the existence of such bias (e.g. Chesler, 1971) and casts in sharper
relief the assertion that 'the sex orientation of this society is not only shared, but also promoted by its clinical personnel' (Neuberyer, 1968, p.554).

Wenger and Fletcher observed and took notes on admission hearings of 81 persons against whom petitions had been filed for incarceration in a State mental hospital in the United States. The purpose of each of these hearings was to determine if the individual was 'sane'. The median time for the hearing was 5.03 minutes. But more importantly for our present purpose, of those individuals who did not retain a legal counsel, over 92% were admitted, whereas of those who did retain a lawyer, only 27% were admitted. It is, of course, possible that the 'saner' individuals were more likely to hire a lawyer, as Gove (1970) has noted, but Wenger and Fletcher had independent observers categorise each patient as to whether or not the legal criteria for insanity were met. "The patients classified as borderline or as criteria-not-met, and who had legal counsel, were more likely to be released than similar patients not represented by a lawyer" (1969, p.7).

Parenthetically if Szasz and Sarbin are even partly correct in their assertion of the permanence of the disability associated with being labelled mentally ill, the time devoted to hearing this grave charge in Wenger and Fletcher's study must cast serious doubts on the fairness of this procedure - to put it mildly.

It is clear, although only a small sample of the research has been examined (see the reviews of Abramowitz & Abramowitz 1977; Dohrenwend & Dohrenwend, 1974; Wills, 1978), that psychodiagnosis is extremely susceptible to bias and distortion as a function of irrelevant personal attributes such as social class, political
attitudes, skin colour and sex. These findings lend support to the assertion that diagnostic labels do not necessarily reflect salient characteristics of the person that differentiates him/her from other persons but inevitably reflect the ideology and value premises of the observer. "As such, diagnosis is nothing more than a translation of stereotypes and prejudgements about other people". Braginsky & Braginsky, 1974, p.129).

IMPRESSION FORMATION AND PSYCHODIAGNOSIS

The Braginskys (1974) were quite blatant in their charge that psychodiagnostic categories are little more than scientific sounding social stereotypes (see also Stuart, 1970 Ch.5). Their critique specifically focussed on the content of the psychiatric categories and the characteristics of the people labelled by psychodiagnosticians. It is possible, however, to suggest that psychodiagnosis is a form of stereotyping in a different sense; namely, that the psychological processes involved in both social stereotyping and psychodiagnosis are similar and amenable to similar situational and information processing sources of bias and distortion.

Stereotypes and Implicit Personality Theories

Adinolfi (1971) draws an explicit parallel with his suggestion that it is just as easy and tempting to apply the same cognitive strategies for stabilizing the clinicians' world as it is to stabilize the lay person perceiver's world and that it is done with the same results. That is, the parsimony and establishment of stability and consistency brought about by the extension of physiognomic apprehension, for instance, to personality trait attribution is
enormous and is not limited to untutored or underdeveloped perceptions of others. The accusation is that the inscrutable Orientals, sensual Mediterraneans, and coldly efficient Nordics find their clinical inferential counterparts in the analogical interpretation of the Rorschach, the Walter Mitty extension to active voice and the creative implications of 'proper' associative networks (Adinolfi, 1971, p.173).

Notwithstanding the evidence of extraneous client variables affecting clinical judgements it may seem inappropriate, and even provocative, to discuss social stereotyping in the same breath as psychodiagnosis - after all psychodiagnosis is conducted by trained experts, professionally committed to their clients' welfare. Certainly, neither the commitment nor the integrity of most clinicians is under question here, rather the focus is on the conceptual and empirical adequacy of their assessment theory and behaviour. However, it will be argued that if a stereotype is defined as "...a region of one's implicit personality theory to which access is gained by a small number of cues, within which the correlations among components approach unity and which has relatively few connections with other regions" (Jones, 1977, p.54; see also Cauthen, Robinson & Krauss, 1971, p.118) then it is possible to show that the psychological process used by lay or intuitive psychologists in arriving at social stereotypes is similar to those used by many clinical psychologists who operate within the attribute model.

Social stereotypes are a special case of interpersonal knowledge and are usually simple, overgeneralised, and widely accepted. However, social stereotypes are no different in principle, from
other types of interpersonal expectations and include those interpersonal expectations arrived through psychodiagnostic assessment and codified in the currently used psychodiagnostic labels. In this section it will be demonstrated that the process of psychological assessment and categorization carried out by professional clinicians is influenced by the same variables that influence our day-to-day perceptions and expectations about others. Subsequently, in Chapter VIII, it will be shown that psychodiagnosis - being essentially an interpersonal encounter - is an inevitably fallible endeavour: the human element in psychodiagnosis make this irrefutably the case. Both the fallibility of clinical judgements, and paradoxically, the illusion of validity are understandable given the nature of the judgemental task, the ways in which most of us code and process information and the heuristic principles used in judgements under uncertainty.

Professional Training and Clinical Judgement

It was mentioned a little earlier that the commitment and integrity of the professional clinician was not under scrutiny but claims to professional training and the use of arcane language do not amount to evidence in favour of special skills or exemption from the human tendency to fall prey to pervasive subjective biases and distortions that characterise everyday personal perception (see, for example, Ross, 1977). In fact, it is apparent that the extent of professional training and experience does not necessarily increase predictive accuracy.

For example, an early publication by Luft (1950) reported two studies designed to determine if professionally trained clinical
psychologists and psychiatrists had an expertise about the
behaviour of others that is not shared by equally 'intelligent'
lay persons. In both studies the experts proved to be no more
accurate in predicting from known behaviour (part of a case
history) to behaviour not known (another part of the same case
history) than the non-clinicians. Similar results have been recorded
by others (see, for example, Breland, 1959; Grigg, 1958; Soskin,
1959). In a review of the literature in the area Sarbin, Taft and
Bailey (1960) found that of the fourteen studies included that
compared the accuracy of clinical psychologists and psychiatrists
with various groups of non-professionals, six showed no difference,
five favoured the trained professionals and three favoured the
non-professional. Luft conjectured on the basis of his study that
"Prediction from such material may therefore call for general
intelligence rather than special clinical understandings" (1950,
p.757).

Alternatively, it could be argued that the judgements and
predictions based on the cues contained in 'such material' tap
into shared aspects of the implicit personality theories of the
judges (Jones, 1977) and the utilization of similar heuristics
governing intuitive prediction and judgement (Tversky and Kahnemann,
1973). These conjectures form the central thesis of this section
and will be discussed in detail subsequently.

Recourse to professional training cannot be accepted as a
substitute for empirical evidence in support of the reliability
of diagnosis based on the attribute model. In the discussion to
follow data will be presented to refute the assumption that the
salient, if not the only, variable that determines psychodiagnostic
judgments is the personality of the client. Rather it will be argued that the assessment situation is more accurately construed as an interpersonal encounter in which situational, clinician and client variables interact in an intricate reciprocal influence processes. If the attribute model is used to construe this interaction human error is augmented and bias and distortion are almost inevitable. Some of the error will be attributable to characteristics of the clinician, some to the characteristic of the situation (including the assessment devices), and others, as we have seen in the previous chapter, to task irrelevant characteristics of the client. The data to be presented below indicates that mistakes and differences in psychodiagnosis are common and that the process of diagnosis is afflicted by many of the same biases, distortions, preconceptions and expectations as our everyday person perception that often results in stereotyping.

RELIABILITY OF PSYCHIATRIC DIAGNOSIS

The traditional methods of assessment, as we have seen, include the administration of psychological tests, taking social histories, and the 'mental-status' examinations; all of which are primarily designed to match the individual's symptomatology with a categorical descriptor or syndrome. A central question in evaluating psychodiagnostic practice is to ask how consistently do different diagnosticians place the same people in the same category, or alternatively - what is the rater reliability of psychodiagnostic labelling? While occasional criticisms and reanalysis of data obtained by studies assessing reliability do appear (e.g. Levy, 1972) and challenge the consensus, the overwhelming evidence suggests that when assessments
of clients are made by independent clinicians the resulting inter-rater agreement is of a low magnitude (see Frank, 1969, 1975 for comprehensive reviews).

As an example Beck and his colleagues (Beck, Ward, Mendelson, Mock & Erbaugh, 1962) selected and especially prepared four experienced broad level psychiatrists (i.e. accredited specialists) who were paired up to interview successively (about five minutes apart) 153 outpatients. Prior to interviewing, the psychiatrists conferred with each other and reached agreement on diagnostic criteria. The results present the 'most gratifying' study in the literature in terms of degree of agreement for subtype of disorders, but considerable disagreement still existed. When both psychiatrists were 'certain', they agreed in 81 per cent of the cases; when both were 'uncertain', they agreed in 25 per cent of the cases; all other combinations of certainty yielded around 48 per cent agreement. Finally, there was an agreement rate of 70 per cent for the three broad categories psychotic, neurotic and character disorder.

Beck and his colleagues (Ward, Beck, Mendelson, Mock & Erbaugh (1962) followed up the work of Beck at al (1962) by using a portion of the 153 assessment reports and analysing them in an attempt to determine why the psychiatrists had disagreed. The authors report three sources of error:

(1) in 5 per cent of cases the primary reason for disagreement was inconsistency on the part of the patient who gave different information of different interviewers;

(2) inconsistency on the part of the interviewer accounted for an additional 32.5 per cent of the disagreement; and
(3) intrinsic inadequacies of the diagnostic systems categories accounted for the remaining 62.5 per cent.

The general conclusion to be drawn from this and many other studies (see, e.g. Rubin, 1948; Thorne, 1961) vis a vis the logic of the categorical system can best be summarised by the following statement:

'A certain degree of relationship has been discovered between symptoms manifestation and diagnosis. However, the most striking finding of the present study is that the magnitude of these relationships is generally so small that membership in a particular diagnostic group conveys only minimal information about the symptomatology of the patient. One is faced with the perplexing finding that the occurrence of a wide variety of symptoms may be related to more than one diagnostic category' (Zigler & Phillips, 1961, p.73).

It appears that the current psychiatric nosological system does not provide a cogent and explicitly logical framework for clinical judgements: in fact, it is a major contributory factor to inaccuracy and lack of agreement. As a result many studies have documented the clinician's influence on categorical decision making. For example, based on four psychiatrists' diagnostic judgement of the same 100 case files Goldfarb (1959) typically concluded by saying: 'The fundamental question raised by this study was whether or not the diagnosis given to groups of patients depends to a significant degree upon which clinician made the evaluation. The results suggest that the answer to this question is substantially "Yes"' (p.396). Grosz and Grossman (1968) report a similar result based on a study of diagnostic decisions of five psychiatrists who evaluated comparable samples of approximately 30 patients. The psychiatrists appeared to differ systematically in their tendencies to rate consistently high or consistently low levels of "abnormalities" in their patients.
Several recent studies have highlighted another aspect of the reliability shortcomings of diagnostic systems; namely, the change of diagnosis over time. This was graphically evidenced by Blum (1968). Blum reasoned that if the psychiatric nosology provided a logical and explicit framework for clinical judgements one would expect symptoms to predict diagnosis at different points in time. However, over a twenty year period, it was found that the changes in patients' primary symptoms could account for only half of the variability in ascribed diagnostic labels. Diagnostic patterns seem to change over time in ways not fully explained by the symptoms the patient presents (Blum, 1978; Kuriansky, Demig & Garland, 1974; Morrison, 1974).

In a similar vein Sandifer and his coworkers (Sandifer, Hordern, Timerry & Green, 1968) convincingly demonstrated the trans-situational variability in diagnostic practice in their study of diagnosticians in London and Glasgow. The conclusion was that each area appeared to build up its own norms that differ somewhat from other areas. Kadushin (1969) indicated how this might occur:

'Because psychiatric theory and concepts are not well agreed upon the rules by which diagnosticians move from the manifest response of patients to their underlying psychiatric diagnoses are incoherently or poorly specified. Therefore, the social situation of clinics became more important in determining diagnosis than the characteristics of applicants. The divergent theories of the various clinics also account for much of the differences among them in diagnosis. Each clinic's position is fixed in the process of routinizing the complex scientific vocabulary of diagnosis through social interaction, each clinic develops its own set of norms for the application of diagnostic terminology, for each seems to have a favourite diagnostic category. Finally, administrative reasons force many clinicians who would not otherwise do so to make diagnoses.
Consequently, the diagnoses themselves seem to follow a clinic's administrative exigencies." (p.111)

The Influence of Examiner Characteristics

Since the psychiatric nosology does not appear to provide guidelines for making clinical judgements and additionally since assessment is essentially an interpersonal encounter, it is to be expected that the clinician may unwittingly influence the outcome of assessment in many subtle ways. On this score, there is a substantial body of research to the effect that "... the personality of the examiner is significantly related to the type of Rorschach protocol which he[she] obtains when his[her] technique has been standardized and his[her] subjects have been selected at random" (Sanders & Cleveland 1953, p.47). Similarly Gibby (1952) found strong and significant differences in the Rorschach protocols as a function of the examiner. In a study designed to delineate some of these differences Filer (1952) concluded that "the three most frequently mentioned defence mechanisms in (the) reports are more characteristics of the clinicians than their patients" (p.336). In fact, Kessel and Shepard (1962) suggest that clinical judgments may reveal more about the clinician than it reveals about the client whom is being assessed. George Kelly (1955) would no doubt agree. According to Kelly, "when the examiner tries to guess what the subject is thinking, we call it a projective device" (1955, p.332).

However, as Masling (1959) noted, the ambiguity inherent in projective devices may render them particularly susceptible to interpretation effects discussed above. Consequently, Masling set out to see if similar effects occurred with more objective devices;
that is, "we call it an objective test ... when the subject is asked to guess what the examiner is thinking" (Kelly 1955, p.332). Masling (1959) was able to demonstrate differential responding by the clinical administrator of an intelligence test as a function of the warmth exuded by the examinee. There are good reasons to speculate that in the naturalistic setting such interactions could influence the final IQ ascribed to examinees. A study by Hersh (1971) replicated Masling's (1959) study on intelligence testing and pointed out that psychological testing is seldom purely a dyadic situation - usually there is a referring third party. The impact of the expectations of this third party on the examiner and the resulting judgements was documented by Towbin (1960). In Towbin's study children were referred for assessment under two conditions - positive referral that stressed the child's social resources and negative referral that stressed the child's relatively poor standing academically. The examiner administered the Stanford-Binet Intelligence Test and the results indicated a 5.5 point difference in favour of the positive referral group.

Despite the widespread use of tests such as the Rorschach, the MMPI and various intelligence scales, most clinical judgements are based on interviews with clients, not on specific tests. In fact, Wade and Baker (1977) found that only a relatively small percentage of respondents in their survey reported that they would accept the 'opinions' of their tests if they conflicted with personal hypotheses developed through other avenues. And secondly, the usual concern of psychodiagnosis is with relatively global categorisation of the client and not with scores on specific tests. Thus the research cited above may not be particularly germane for actual practice.
However, the research of Temerlin (1968) appears particularly
germane since it examined the effects of suggestion on the making
of a diagnosis in an interview situation.

**Suggestion and Psychodiagnosis**

In the much cited study Temerlin had a professional actor enact
the role of a well adjusted, super-normal man, he was happy and
effective in his work, he established a warm, gracious and satisfy­ing
relationship with the interviewer, he was self-confident and
secure without being arrogant, competitive, or grandiose. He was
identified with the parent of the same sex, was happily married and
in love with his wife, and consistently enjoyed sexual intercourse.
He also had a good sense of humour, no hallucinations, delusions, or
psychosomatic symptoms, a happy childhood, and reasonable worries
like concern over Vietnam.

The experiment had three conditions. The first two groups
represented the first condition; that of no suggestion as to mental
health; either no suggestion at all or the designation of an
employment interview. In this condition roughly one third of
the raters indicated the man was neurotic, none indicated psychosis.
In the second condition a prestigious figure gave an 'unintended'
suggestion of mental health. The result was that 100 per cent of
raters indicated the man was mentally healthy.

In the third condition the tape was played for several large
groups of subjects including clinical psychology graduate students,
clinical psychologists (Ph.Ds) and psychiatrists (M.D.s) with the
suggestion that the man was, albeit unapparently, mentally ill.
Operationally, just before the tape was played, a distinguished mental health professional remarked to the man next to him that the person on the tape was "a very interesting man because he looks neurotic, but actually is quite psychotic". For the mental health condition, the suggestion was given in the same way, notably that it was a tape of a truly rare man, one who was normal.

The rankings were dramatically affected in condition three, and within that condition subgroups displayed differential response to the negative suggestion. For example, psychiatrists were most influenced resulting in 60 per cent rated the man psychotic, 40 per cent neurotic and none thought him healthy. The clinical psychologists were next influenced; 28 per cent rated the man psychotic, 60 per cent neurotic and 12 per cent healthy. The undergraduates were least affected with only 11% rating the man psychotic but the vast majority rated him neurotic (78 per cent) and only 11 per cent rated him healthy.

The early Temerlin (1968) study was replicated and extended by Temerlin and Trousdale (1969) with very similar results. Diagnosis of some form of 'mental illness' ranged from 84% to 100% in the experimental group which received a suggestion of mental illness. Such diagnoses were made by 0 per cent to 43 per cent of the subjects in the control groups depending on their composition. Further, not a single experimental subject (out of a total of 300) wrote a descriptive rather than inferential report of the interviewee when requested to do so. Not a single experimental subject described the behavioural basis for his/her diagnosis. Temerlin and Trousdale (1967) concluded that "inference, unchecked by systematic and repeated observations of behaviour, may produce gross errors of interpersonal perception" (p.28).
Diagnostic Decision Making

When it comes time to write a 'case' report, the recourse to inference and the security of labelling the client and therefore imply understanding (to name is to know) is probably what lies behind this astounding finding. In fact, Gauron and Dickensen (1966) pointed out that clinicians are at times so personal in their criteria for the selection of certain diagnostic inferences that they are unable to offer formal explanation of their choices. In their study, the information (cues) sought out by psychiatrists in arriving at their diagnostic decision were correlated with an index of actual cues used by psychiatrists. They found that the two were not related. Gauron and Dickinson (1966) use this to argue that: "... In actual fact the psychiatrist derived the clues underlying his diagnosis from sources other than he[she] thought he[she] did" (p.203). Further, Rommetveit (1960) has rather convincingly shown that the dimensions which perceivers admit to using in their differentiation of others are unrelated to the dimensions on which they actually make their differentiations.

These findings are not surprising in the light of recent research on our access to higher order mental processes. Nisbett and Wilson (1977) cogently demonstrated that people often cannot report accurately on the effects of particular stimuli on higher order, inference-based responses. "Indeed, sometimes they cannot report on the existence of critical stimuli, sometimes cannot report on the existence of their responses, and sometimes cannot even report that an inferential process of any kind had occurred." (Nisbett & Wilson, 1977, p.233).
The situation in clinical diagnosis appears to be exacerbated by the evidence that most clinicians accept their personal hypotheses over test results if the two were in conflict (Wade & Baker, 1977, p.879). Furthermore, the great majority of clinicians reported using personalised evaluation procedures for projective tests while over one third reported using such procedures with objective tests. Wade and Baker (1977) went on to show that respondents claimed they would evaluate test results according to their personal or clinical hypotheses and opinions. They indicated that most clinicians use a clinical data combination system (see, for example, Sawyer, 1966). Sawyer has demonstrated that behavioural prediction will be decreased when both kinds of data (i.e. clinical and mechanical) are combined, rather than by using mechanical (test result) data alone. Thus the respondents in Wade and Bakers study clearly indicated that in their everyday practice clinicians behave in ways that are likely to lessen the judgement and prediction accuracy.

Before leaving Temerlin studies it is important to note that only one stimulus person was employed and he represented the values of the middle class. The suggestion appeared to produce dramatic differences between the experimental and control groups, but rather than being purely a function of the suggestion, the difference may be due to an interaction between the suggestion and the particular stimulus person employed (Jones, 1977, p.69). The person perception literature does seem to have demonstrated the strong relationship between being actually similar to or highly familiar with the perceived person and accurately perceiving him/her. The sequelae of this relationship are visible in the clinical psychology area, as we have seen in Chapter VI.
In a recent discussion of the literature on behaviour classification Phillips and Draguns (1971, p.467) point out that research on the clinician's diagnostic activity seems to converge in suggesting that social distance is a mediating variable. As we have seen, the middle-class diagnostician is likely to employ different categories for lower-class 'patients' and is likely to note fewer symptoms in the process of assigning lower-class 'patients' to those categories. The point here is that the kind of 'suggestion' employed by Temerlin is relatively artificial and may not occur in most naturalistic settings. This, however, does not deny the possibility of large suggestion effects in practice.

The Diagnostic Utility of Different Kinds of Data

If we conceptualise clinical diagnosis as a process of categorisation, a process in which certain cues or combinations of cues are used to refine the placement of "... the patient from a supraordinate stereotype (i.e. class of events) to more subordinate ones as increments of information are accumulated" (Kostlan, 1955, p.486) then it becomes important to know whether or not the clinical judge employs a given combination of cues consistently. That is, can a consistent judgement strategy be identified and separated from the judge's unreliability?

Research has been undertaken to determine if there are certain commonly employed sources of information about clients that lead to more valid inferences about clients. Using a 'functional omission' design Kostlan (1954) had twenty clinical psychologists, each with at least two years of psychodiagnostic experience, make diagnostic judgements on the basis of only three out of a possible
four pieces of available data (i.e. the Rorschach protocol, the MMPI results, the Stein Sentence Completion Test (SCT) protocol, and social case history). Each judge viewed the material from one of the 'patients' under one of the five conditions (i.e. missing the Rorschach, MMPI, SCT, case history or everything except "Barnum" data).

The interesting finding was that when deprived on the case history, clinicians could make inferences of accuracy no greater than they could on the basis of the 'Barnum' profile (i.e. age, marital status, occupation, education, and the source of referral). Meehl (1959) summarised these results in the following way:

"... clinicians knowing only the age, marital status, occupation and sources of referral of a patient (that is, relying essentially upon the Barnum Effect for the ability to make correct statements) yield an average of about 63% correct statements about the patient. If they have the Rorschach, Multiphasic and Sentence Completion Test but are deprived on the social history, this combined psychometric battery results in almost exactly the same percentage of correct judgements" (pp.115-116).

In a more recent review of the role of personal history data in clinical judgement Potkey (1973) concluded that such data, as a source of clinically descriptive or predictive information, is at least as effective as information derived from psychological test sources. Paradoxically, however, Oskamp (1965) found that self-confidence in judgements made by clinicians was found to increase as a function of the amount of information available to them but without any corresponding increase in judgemental accuracy. Thus the answer to the questions posed above appears to be that some
of the commonly employed devices used by clinicians in the assessment process, namely, projective and objective tests are not likely to increase the validity of clinical judgement beyond that gross level of categorization that can be attained with face sheet data.

A line of research addressing the question from another direction indicated that an increase in the amount of information brought to bear upon particular decisions may not enhance the accuracy of these decisions (Cline, 1955; Dailey, 1952 and as Oskamp 1965 indicated above). This research, when coupled with that of Bieri (1962) who studied the effects of increases in the amount of undirectional information revealed that clinicians reach an early appraisal of their patients and do not modify these impressions as a function of increased information or even as treatment progresses (Meehl, 1960; Parker, 1958). This paints a gloomy picture of psychodiagnosis.

**Barnum Statements in Psychodiagnosis**

Similarly Sine (1959) in a study designed to establish the diagnostic utility of different kinds of data found that omitting the Rorschach was associated with greater validity than including it. He also found that absolute improvement from the initial biological data sheet to final clinical judgement based on additional information was minimal. Conversely several studies have shown an inverse relationship between the amount of information and the accuracy of judgement (Cline & Richards, 1960; Sawyer, 1966; Schwartz, 1967).

Additional evidence could be cited (see e.g. Frank 1969, 1975; Stuart 1970, Chap.4) but presumably the point has been made. The
clinician has finite and perhaps quite limited capacity for processing information. In spite of what is claimed for the usefulness of projective data, decisions do not improve their accuracy of prediction or reliability of categorization by employing such data. Apparently the important cues or, at least, the cues utilized are contained in other aspects of the clinical situation. And as Meehl (1959) has forcefully suggested, that for one third of the clients "... the application of a stereotype personality description based upon actuarial experience in this particular clinic provided a more accurate description of the patient than the clinician's judgement based upon any, or all, of the available tests, history, and interview data!" (p.116) Dawes (1970) has noted that after 20 years of research demonstrating the superiority of statistical over clinical prediction, clinicians continue to ignore the former and use the latter.

Of interest to us here is Cline's (1964) review of his work in person perception in which he suggested that the superiority of the statistical or actuarial type prediction over clinical prediction rests primarily on the stereotyped component of actuarial prediction. It would seem, then, that the traditional clinician has been criticised for relying too often on Barnum statements (stereotyped accuracy) rather than on differential accuracy. The clinician is then compared in his/her predictive accuracy with methods that patently rely on actuarial data (stereotyped accuracy). Neither approach appears sufficiently sensitive to the individual client. In summary, it can be said that psychodiagnostic assessment and clinical judgement which is designed to identify dispositional characteristics in their clients have not demonstrated a high level of reliability.
Evidence reveals that clinicians see different cues in the same client, and use the same cues to infer different dispositions (Cole & Lavery, 1969). Mistakes and differences in psychodiagnosis are common and the psychodiagnostic process seems to be susceptible to many of the same influences that shape biases, distortions and preconceptions in intuitive psychologists' person perception.

Although much of the lack of reliability can be attributed to weaknesses in the conception of the diagnostic typology, a significant part is intrinsic to the situation in which judgements and predictions are made and to the clinician's cognitive information coding and processing of outcome data. Evidence supporting the conjecture that trained professional clinical psychologists are generally as fallible in their inferential judgements as intuitive psychologists will be presented in the next chapter. Additionally, it will be shown that trained professionals are equally vulnerable to the temptations of magical thinking. As we have seen stereotyping and psychodiagnosis are, in principle, no different and therefore we would expect magical thinking to be as much a characteristic of the diagnosticians as the modern lay person; as indeed the so called 'savage' mind. That is, most of us have a 'savage' mentality much of the time and this includes the professional psychodiagnostician.

The language of psychodiagnosis should not blind us to its nature. As has been pointed out many times, it is a mistake to make judgements about thought processes on the basis of their content (see, e.g., Levi-Strauss, 1966). Extraordinary beliefs follow quite logically from extraordinary premises. However, our saturation in
our culture quite often prevents us from identifying our extra-
ordinary premises - particularly when accorded the hallowed status
of being 'scientific'. In the discussion that follows it will be
demonstrated that magical thinking is as much a characteristic
of our technological society as other 'primitive' societies. It
is a very rare individual who is bereft of all such thoughts. But
more importantly, the attribute model of assessment and the
assessment situation do little to help clinicians guard against
some of the important pitfalls of information processing in
psychodiagnostic assessment.
CHAPTER VIII: THE FALLIBLE PSYCHODIAGNOSTICIAN

In the decade since 'Personality and Assessment' (Mischel, 1968) was published the study and catalogue of judgemental fallibility of clinical psychologists has continued but, fortunately, research in cognitive and social psychology, particularly investigations of how people categorise, simplify and process information has flourished. This has greatly illuminated the cognitive basis for many of the catalogued distortions. The same research also helps to explain the blatant contradiction between the considerable evidence of fallibility of clinical judgement and the seemingly unshakeable confidence clinicians show in their judgement ability. It appears that the illusions of validity persist very largely because clinicians not only use tests of dubious validity but because they fall prey to many pitfalls of information processing.

So far, the discussion of the mistakes and differences in psychodiagnosis has been mainly descriptive. This chapter will attempt to promote an understanding of some plausible processes that might account for such biases and lack of reliability, as well as the persistent illusions of validity. For example, Wade and Baker (1977) revealed that clinicians approach testing in a way that maximises the possibility of falling prey to phenomena, such as illusory correlates (e.g. Chapman and Chapman, 1971) and the Barnum effect (e.g. Meehl, 1956). In addition, the structures of judgemental tasks in the natural environment prevent clinicians from using important information for drawing valid inferences. This shortcoming, when coupled with information processing shortcomings such as the disinclination to think probabilistically and the
inclination to eschew disconfirmatory evidence, predisposes the clinician to form and maintain the illusion of validity. In the discussion to follow, recent research in cognitive and social psychology that bears on the clinical judgement process will be reviewed with the aim of isolating factors and processes that may explain the fallibility of the psychodiagnostic enterprise.

It is interesting to note that much of the relevant research has been conducted in the area of attribution theory and intuitive psychology. Attribution theory in its broadest sense is concerned with the attempts of ordinary people to understand the causes and implications of the events they witness. According to Ross (1977) this perspective has led to the elevated image of psychological 'man' from passive reactor to a status equal to that of the scientist who investigates him/her. That is, people, in the perspective of attribution theory, are intuitive psychologists who seek to explain behaviour and to draw inferences about actions and their environments.

Such a perspective, although not new (see again Kelly, 1955), does suggest that both the professional and academic psychologist and the lay person, are struggling with the same task, namely to discern, describe and apply systematic covariation between events. Consequently, the onus is on psychologists to demonstrate the superiority of their investigating methodology if we are to trust their knowledge and explicit personality theories in preference to the intuitive, implicit personality theories.
PITFALLS OF INFORMATION PROCESSING IN PSYCHOLOGICAL ASSESSMENT

The research to be discussed below suggests that the professional clinical psychologist and the lay psychologist are faced with similar problems in their attempts to uncover causal relationships. Both appear equally vulnerable to fall prey to pitfalls of information processing, not only because of their shared human qualities but because of important similarities in the context in which judgements are made. Consequently, one would expect many of the same biases, distortions, preconceptions and expectations in psychodiagnosis as one finds in everyday perceptions of lay persons that result in stereotyping.

Magical Thinking and the Disinclination to Think Probabilistically

One of the keys to understanding psychological diagnosis is an appreciation of the probabilistic relationship between covarying events. For example, between behavioural signs and the dispositions or entities one infers from these signs. Since there is seldom an absolute, one-to-one relationship between a given sign and an inferred characteristic, it becomes incumbent upon the diagnostician to determine exactly which cues can be relied upon. For given the probabilistic nature of the relationship between clinical signs and inferred attributes, the well developed propensity for positing causal relations may lead to magical thinking, (i.e. ascription of a causal relation in a case when there are no objective grounds for it (Johnson-Laird and Wason, 1977).
Common observation and experimental investigation (Michotte, 1963; Jenkins and Ward, 1965) evidence the proclivity of observers who in seeking to discover a pattern in events will (too) readily construe them as causally related. It is all too easy to move from observation of concomitant events to interpretation that the first event is the cause of the second event. Although systematic covariation is a prerequisite for accurate prediction, it is hardly sufficient for the inference of a cause-effect relationship. This, of course, relates to the oft-quoted difference between correlation and causation, or what David Hume regarded as the distinction between sequence and consequence. Hume established that this step — from covariation to causation — was logically unwarranted on the basis of observation alone.

Blissfully ignorant of Hume's judgements, causal inferences from observation are the hallmark of intuitive and professional clinical psychologists. As Hume first pointed out, there is nothing other than 'habit' or 'custom' to justify induction. It has no logical warrant.

Inferring Causality in Person Perception

It is plausible to suppose that a causal interpretation of the world is advantageous because it confers some power of prediction to its adherents: fallacious theories are usually more useful than no theory at all (cf. Jones, 1979). With few exceptions, most students of person perception have endorsed the view that individuals construct images of other people in ways that serve to stabilise, make predictable, and make more acceptable their view of the social world (e.g. Heider, 1958; Kelley, 1971). The ascription of
causality is accordingly an interpretation of events, even in the case of the simple but powerful causal illusions of the laboratory (Michotte, 1963). "In the interpretation of the events of daily life, it is almost invariable that in order to set them within a broader causal or intentional framework, it is necessary to go beyond what is explicitly given" (Johnson-Laird and Wason, 1977, p.438).

Although covariation, no matter how systematic can never demonstrate causation, this is also true of the best conceived experiment: causality is inferred and is thus a conceptual rather than empirical entity. In ordinary life, in or outside the clinic, a correlation combined with 'common sense' is usually sufficient to lead to causal claims. The inherent dangers in this process will be demonstrated below.

The problem, which on the surface appears to be one of simply learning which events are associated with which cues, presents difficulties to most humans. Or more accurately, the problem is not so much learning to associate an event which follows a cue but learning what happens to the association when the events appear or do not appear in the absence of the cue. Apparently, these important bits of evidence pertinent to the association are generally ignored. This is readily revealed experimentally (Jenkins and Ward, 1965; Smedslund, 1963; Ward and Jenkins, 1965).

Correlation as a 'Non-Intuitive' Concept

Ward and Jenkins (1965) for example, constructed a series of problems having to do with the relationship between rain and cloud seeding. Subjects received information on a trial-by-trial basis -
a series of statements such as: "On the first day the clouds were seeded and it rained, on the second the clouds were again seeded but it did not rain, on the third ... etc." The results indicated that most subjects judge the strength of relationship between cloud seeding and rain by the frequency of positive instances, while generally ignoring information with respect to negative instances as well as false positives and false negatives. They were generally very inaccurate in deciding whether or not seeding was related to the probability of rain. In fact, only 17 per cent of the subjects followed a logically defensible rule for making their judgement.

As Ward and Jenkins (1965) note: "In general, our results lend support to the conclusion that statistically naive subjects lack an abstract concept of contingency that is isomorphic with the statistical concept. Those who received information on a trial-by-trial basis, as it usually occurs in the real world, generally failed to assess adequately the degree of relationship present" (p.240). As we will see a little later, correlation is a 'non-intuitive' concept, like many statistical concepts, and is generally absent from the thinking of most normal adults including social scientists (see e.g. Smedslund, 1963; Shweder, 1977).

Causality and Correlation and Clinical Diagnosis

What is the relevance of this research to clinical diagnosis? One could argue that experimenters like Ward and Jenkins were unrealistic in expecting subjects to take in, store, retrieve, and appropriately compare such a large amount of information.
Certainly the only condition in which a majority of students displayed correlational thinking was when they received summary information in the form of a 2 x 2 contingency table without prior trial-by-trial experience with the data. Further, one might argue that naive subjects are not, after all, trained diagnosticians most of whom have formal training in statistics and probability.

On the other hand, as Jones (1977) argues, the task facing the clinician who has a hunch about a particular cue being related to a particular problem is much more complex and difficult than the typical laboratory task. Firstly the clinician will probably not be exposed to all the necessary information since those who do not have the problem are not likely to see the clinician. Hence, the clinician will never know whether or not they have the symptom. That is, half of the 2 x 2 contingency table is usually not available. Consequently in the naturalistic setting judges are prevented from using information usually available in laboratory studies (cf. Bjorkman, 1966). Secondly, even the information the clinician does obtain relevant to the relationship of interest is likely to be gathered over the course of months or years rather than in a single experimental session. The problems introduced by these two factors will be elaborated at length later in the chapter.

In sum then, the analogue studies evidence a strong disinclination to think correlationally and an inclination to infer the strength of a relationship according to the number of positive instance of co-occurrence. Support for this view has been made in a somewhat different manner by Smedslund (1963) with reference
to Piaget and Inhelder's (1961) theory of the development of probabilistic notions in children. Smedslund notes that the stage of concrete reasoning, which precedes the ability to apply "correlational reasoning", functions only on the basis of observable events. He suggests, although adults are capable of using disconfirming information for inferring relationships, they frequently fail to do so and operate at lower, cognitive level, for example, concrete reasoning. In other words, judgemental habits are based on experience with lower levels of cognitive functioning. As Shweder (1977) contested, because correlation is not an intuitively available concept, normal adults are not likely to master it without explicit instructions to do so.

The research cited above indicates that the frequency of positive instances of co-occurrence is the major determinant of perceived correlational relationships. On the basis of this Einhorn and Hogarth speculate that "If outcomes are coded in memory as frequencies rather than probabilities, this has major implications for explaining the persistence of the illusion of validity" (1978, p.400). Recent research on probability learning is relevant to this issue and corroborates the conjecture that outcomes tend to be coded in memory as frequencies rather than probabilities.

**Probability Learning**

In an extensive series of experiments Estes (1976a, 1976b) investigated the coding of outcomes in memory and how subjective probability and predictive behaviour are based on such coding.
Estes' (1976a) results indicate that subjects had a strong tendency to predict the more frequently occurring variable (in his study, the winning candidate) even when that variable had a lower probability of occurring (winning). In the contingency tasks set by Estes the experimental evidence is consistent with the notion that frequency is more salient in memory than probability.

Again the studies by Estes evidence the inability of people to deal adequately with non-occurrence of events resulting in a favouring of coding outcomes a frequency of concrete events rather than probabilities. Interestingly, it should be noted that the notion of probability itself has developed relatively late (sixteenth and seventeenth centuries) and David (1962) points out that this late development is remarkable when one considers that notions of gambling and games of chance existed in antiquity. This supports the non-intuitive nature of correlation and probability proposed by Shwedler (1977).

Shweder (1977) suggests that failure of intelligent but statistically ignorant people to grasp the concept of correlation is a potential mechanism underlying magical thinking. He asserts that "as soon as events can be meaningfully linked to one another, magical thinking makes its appearance; Normal adults substitute the readily available intuitive concept of resemblance for the unavailable non-intuitive concept of correlation" (1977, p.451). That people readily substitute a mere similarity between two states of affairs for an actual correlation between them has been corroborated by other researchers (e.g. Tversky and Kahneman, 1973). Thus magical thinking seems to occur when adults assess the degree of
empirical relationship among events that also conceptually affiliate or exclude one another in their minds. Thus what D'Andrade (1965) has called a 'hazard of science' is an appropriate definition for magical thinking. That is, a confusion of 'propositions about the world with propositions about language'. Everyday empirical claims about 'what goes with what' in experience typically according to Shweder (1977), turn out to be claims about non-correlational relationships among interpretive categories.

Resemblance in Reasoning

What is striking about Shweder's claim, however, is not that people untutored in statistics have difficulty with correlation - we have already seen this - but the nature of the concept with which they replace it: resemblance between things. In fact, Shweder claims that resemblance, not correlation, is a fundamental conceptual tool of the everyday mind. If this claim is corroborated then homeopathic magic (i.e. in which one event is supposed to cause another event of a similar sort) is almost unavoidable. Most people regardless of their culture, Shweder argues, will be prone to argue, if A resembles B, then A tends to be associated with B. When this is coupled with the principle of causal inference discussed earlier, one obtains: A tends to be associated with B, and if A proceeds B then A causes B. By these two inferential steps, people may come to believe that the application of fowls' excrement cures ringworms (Tambiah, 1973) or that penicillin cures a sore throat (Malleson, 1973), or even perhaps that experiences of early childhood play a crucial role in determining adult personality. Such propositions may, of course, be true - that is a matter of scientific testing.
The point to be emphasised is that 'magical' thinking is simply a useful heuristic for generating ideas. The danger arises when these 'ideas' pose as scientific knowledge as may be the case in clinical inferential learning where 'resemblance' may be used as an index of contingent or correlational relationships in experience.

CONFIRMATORY BIAS

One thing does seem clear, however, and that is that in practice a clinician who would like to check out a conjecture about a relationship between a behavioural cue and a personality disposition by observing how these covary in the next, say, twenty clients who display the cue is inevitably going to end up making an inaccurate assessment. For as was noted above, and as Wason (1968) points out, "... scientific inferences are based on the principle of eliminating hypotheses, while provisionally accepting only those which remain. Methodologically, such eliminative induction implies adequate controls so that both positive and negative experimental results give information about the possible determinants of phenomena" (p.219). At the very least, the clinician would have to determine for all clients whether or not they have the particular disposition, regardless of whether or not they display the cues. The propensity to ignore information other than positive has been briefly alluded to above.

It will be shown below that thinking in terms of positive instances is an example of the difficulty people have in making use of 'disconfirming information', by which is meant the information that can be gained by the non-occurrence of an action or prediction.
Furthermore, a principal cause of this difficulty is the structure of judgemental tasks in the naturalistic environment. The propensity to ignore disconfirmatory information is so persistent that Snyder and Swann, in a study designed to induce accurate hypothesis testing asserted "that we have yet to develop any procedure that will induce individuals to eschew such strategies (confirmatory) in favour of either disconfirmatory or equal opportunity strategies" (1978, p.1210).

Disconfirmation and the Logic of Lay Persons

There is now available an exciting research literature on the ability to use disconfirming information for making inferences. Central to this literature is the series of papers by Wason (1960, 1966, 1968) in which he explored this issue in detail. In his first published study (Wason, 1960) he presented subjects with a task that gave direct assessment of logical abilities. The task employed was as follows. Subjects were given a triad of numbers (2, 4, 6) and told that these numbers conformed to a simple mathematical rule. Their task was to discover this rule by generating (testing) other triads of numbers and then deducing the rule. Each time they announced an experimental triad, they were told whether it conformed to the rule. Subjects were told to assume their hypothesis about the rule only when they were very confident of its correctness. The correct solution to this task should involve a search for disconfirmatory evidence rather than the accumulation of confirming evidence.

Most individuals seemed to prefer confirmatory rather than disconfirmatory reasoning. As we have already seen, this pattern
is a logically dangerous one in science since "It is easy to obtain confirmations and verifications for nearly every theory - if we look for confirmation" (Popper, 1963). However, confirmation here is always a disguised form of affirming the consequent. The fact that only 6 out of 27 subjects were correct in making their first statement that they had found the rule the first time they thought they did illustrates the dangers of induction by sample enumeration. As Wason (1960) pointed out, the solution to this task must involve "... a willingness to attempt to falsify hypotheses, and thus to test these intuitive ideas which so often carry the feeling of certitude" (p.139).

Logical Reasoning by Scientists

For our present purpose the parallel studies of Mahoney and his associates into the logical fallibilities of individual scientists may be more readily generalised to the clinical population under discussion. In a survey study (Mahoney and Kimper, 1977) physicists, biologists, sociologists and psychologists were asked to rate the validity of four forms of implication. Although most of the scientists recognised the validity of modus ponens, over half of them failed to appreciate the validity of modus tollens (disconfirmation). Twenty-eight per cent of the social scientists thought that denying the antecedent was logically valid and ten per cent thought likewise of affirming the consequent. To further test their reasoning abilities an analogue task patterned after the one employed by Wason (1968) was included (the task will be described in detail later). Fewer than eight per cent of the scientists were able to identify the logically irrelevant experiments and fewer than ten per cent were able to specify the only two experiments which had the critical potential for falsification.
In a second study by De Monbreun and Mahoney (1976), the critical reasoning skills of psychologists and physical scientists were compared to those of relatively uneducated Protestant ministers. The task was that used by Wason (1960) described above. Most of the subjects did not solve the problem and of those who did, only two were errorless - both were ministers. Like Wason's undergraduate students, neither the scientists nor the non-scientists were very inclined to use disconfirmation. Over eighty-five per cent of the self-generated hypotheses were confirmatory. Interestingly, although the scientists and non-scientists were not differentiated on the basis of logic - both were equally poor - differences did emerge on several other measures of reasoning and conservativeness. "Contrary to the story book model", Mahoney notes, "scientists showed a tendency to be more speculative than non-scientists - generating more hypotheses, more quickly and with fewer experiments per hypothesis! Likewise, they were apparently more tenacious as reflected by their differential tendency to return to already falsified hypotheses" (Mahoney, 1976, p.156).

Thus, as Greenwald (1975) has shown 'even' researchers in the behavioural sciences tend to design empirical investigations that seek to confirm, rather than disconfirm their hypotheses. This conclusion converges with those of Wason and Mahoney and suggest that our confident presumption of superior reasoning ability in the scientist may be ill-founded. Professional training may do little to foster critical reasoning skills and logic.
The Clinical Judgemental Task and Disconfirmatory Evidence

It is important to emphasise that in the analogue research cited above, a search for disconfirmatory evidence was possible. However, where the actions of the investigator are based on judgements (as in the case in clinical settings), learning based on disconfirmatory evidence becomes more difficult to achieve. For example, consider how a clinician might erroneously learn the rule 'my judgement is highly predictable'. Suppose the clinician is required to assess the suitability of clients to be admitted into a specific treatment program. The crucial factor here is that action (admittance or not) is contingent on judgement. Therefore, at a subsequent date, the clinician can only examine admitted candidates to see how many have benefited. If there are many improvements (and there are important reasons to think there may be (see Tversky and Kahneman, 1973)) these instances all confirm the rule but, as has been pointed out above, there are several important sources of data not examined or even available here. Thus the tendency not to test hypotheses by disconfirming instances may be a direct consequence of the task structure in which actions are taken on the basis of judgements. Furthermore, as Wason (1960) pointed out, "In real life there is no authority to pronounce judgements as inferences: the inferences can only be checked against the evidence" (p.139). Therefore, large amounts of positive feedback can lead to reinforcement of a non-valid rule and hence to the illusion of validity.
Additional Analogue Demonstrations of our Logical Fallibility

In an equally illuminating study Wason (1968) gave subjects conditional rules of the form *if* $p$ *then* $q$, where $p$ was a statement about one side of a stimulus card and $q$ a statement about the other side. Four stimulus cards, corresponding to $p$, $\neg p$, $q$ and $\neg q$ were provided. The subjects’ task was to indicate those cards - and only those cards - which had to be turned over in order to determine if the rule was true or false. The only cards that could falsify the rule however are $p$ and $\neg q$. Since the $\neg q$ card was almost never selected, the results indicated again the strong proclivity to seek confirmatory rather than disconfirmatory evidence. This bias for selection of confirmatory evidence has proved remarkably difficult to eradicate as indicated earlier (see Wason and Johnson-Laird, 1972, pp.171-201).

This experiment not only highlighted the disinclination to select disconfirmatory evidence (namely, ignoring $\neg q$) but by choosing $q$ instead subjects seemed to follow an assumed symmetry in the problem of the form: If $p$ implies $q$, then $q$ implies $p$. Although this assumed symmetry is clearly a logical fallacy (i.e. showing $q$ is true does not establish either the truth or falsity of $p$), the choice of $q$ in addition to $p$ indicates that the subjects did not perceive it as such. The relevance of this observation for understanding how contingency judgements are made is obvious. In fact Jenkins and Ward (1965) state that subjects in contingency tasks tend to reason as follows: "If there were a contingency, favourable results would occur, since favourable results did occur, there was a contingency."
In a later replication of Wason's (1968) study Einhorn and Hogarth (1978) selected 28 statisticians known to have been trained in examining possible disconfirmatory evidence. Although they used Wason's (1968) experimental paradigm they chose a stimulus related to checking predictive ability in the hope of reducing the abstractness - they chose the relationship between forecasting market rises and subsequent movements. The results indicated that a mere five subjects out of twenty-three subjects emitted the correct response but interestingly none committed the logical fallacy implied in choosing \( p \) and \( q \) as did Wason's undergraduates. Thus skilled statisticians, trained to make use of disconfirmatory information did conduct more disconfirmatory experiments (almost half did evidence this) but half of the subjects chose to examine the same piece of confirmatory evidence, namely \( p \). Like Mahoney's scientists, the statisticians found the task puzzling and inordinately difficult.

On the Strong Proclivity of all Humans to Favour Confirmatory Evidence

Only one study designed on inference behaviour in a setting chosen to resemble the conditions under which actual science is done is reported in the literature. In this study Mynatt, Doherty and Tweney (1977) found that: first, substantial evidence of failure to consider alternative hypotheses, suggesting confirmation bias remarkably similar to that found by Wason (1960) and secondly, subjects could use falsifying data once they got it. That is, if an initial hypothesis was not entertained falsifying evidence led to selection of correct or partially correct hypotheses. Anecdotal
evidence from Mitroff's (1974) large-scale non-experimental study of NASA scientists, reported that a strong confirmation bias existed among many members of this group. He cites numerous examples of the scientists' verbalisations of their own and other scientists' obduracy in the face of data in favour for this conclusion.

Although it may seem implausible to those trained in scientific method and experimental design that people do not seek disconfirmatory evidence when testing hypotheses, the relative novelty of thinking in this manner should not be overlooked. As Einhorn and Hogarth point out, the concept of a control group which illustrates the necessity of non-occurrence came late in the history of thought; as did the notion of equating experimental and control group prior to treatment. Certainly the evidence of Greenwald (1975, cited above) documents scientists' persistence in their claims of confirmation (cf. Popper, 1959, 1963, 1972).

There is much evidence to suggest that the structure of the judgemental task promotes the confirmatory bias and the 'illusion of validity', but additionally, all the above evidence suggests that a bias in favour of confirmatory evidence may be a general characteristic of human reasoning. The structure and process of human thought fosters and promotes the ready and willing adoption of confirmatory strategies for hypothesis testing. Considerable research on concept formation and utilisation indicates that people prefer and use positive instances of concepts over negative ones (see, e.g. Hovland and Weiss, 1953). Moreover confirming instances generally have more impact than do disconfirming instances (see e.g. Gollob, Rossman and Abelson, 1973) and as we have seen,
covariation between positive instances leads to estimates of greater relationship than does covariation between negative or mixed instances (e.g. Jenkins and Ward, 1965; Smedslund, 1965). Furthermore, in judgements of similarity, individuals preferentially look for common features rather than distinctive features (e.g. Tversky, 1977).

Confirmation and Self-Fulfilling Prophecies

To the extent that the individual clinician believes that hypothesis-confirming behaviours are typical of the client's activities, he or she may consider it not unreasonable to confine or at least focus the assessment to those areas in which the client can provide the most informative and meaningful facts vis-a-vis the clinician's hypotheses. Accordingly, the clinician may unwittingly use the assessment interaction as a means of collecting preferential evidence that confirms the hypothesis under scrutiny. Such a preferential evidence-gathering procedure may generate a sample of evidence in which hypothesis-confirming evidence will be over-represented and hypothesis-disconfirming evidence will be under-represented: There is every reason to believe that our clients will be 'generous' in providing specific instances of hypothesis-confirming actions. And, as Regan, Straus and Fazio (1974) have shown, behaviour consistent with one's expectations of another tend to be perceived as due to dispositional properties of the other, whereas inconsistent behaviours are perceived to be due to the situation.
To the extent that clinicians chronically formulate and enact confirmatory strategies for assessing the accuracy of their hypotheses about the relation between behavioural signs or symptoms and dispositional attributes, they may create for themselves a world in which hypotheses became self-perpetuating beliefs (see, e.g., Snyder and Swann, 1978; Snyder and Uranowitz, 1978). From this perspective, it becomes clear how erroneous beliefs about clients are formed and why they are so stubbornly resistant to change.

ILLUSORY CORRELATES

Unfortunately, there are other sources of error that plague our attempts to determine relationships between symptoms and personality characteristics in addition to the failure to think probabilistically and seek and utilise disconfirmatory evidence. A line of research begun some years ago by Chapman (1967) has identified an extremely robust and resistant type of error that appears to be particularly widespread in psychological assessment situations.

Initially, Chapman studied a relatively simple verbal learning task in which subjects were shown pairs of words projected on a screen for a couple of seconds and were later asked to recall how many times particular pairs had been shown. For example, one series consisted of 12 word pairs, each shown ten times giving a total of 120 presentations. The important finding was that when asked to recall how frequently each of the 12 word pairs had been presented in the series, subjects significantly overestimated the number of times those with strong verbal associations had been shown (i.e. bacon-egg, lion-tiger were examples of the pairs overestimated).
In a follow-up study Chapman and Chapman (1967) attempted to extrapolate the finding to the clinical assessment situation. Specifically they felt that it had some applicability to "One of the most puzzling and distressing problem that confronts clinical psychology today ... the persistent report by many psychodiagnosticians of clinical observations which, by objective evidence, clearly appears to be erroneous" (p.193). To be more precise, the Chapman's were concerned with the continued reports that certain responses to projective tests were correlated with certain 'problems' such as homosexuality or depression, when the research evidence appeared to indicate no relationship. The resultant "report by observers of the correlation between two classes of events which, in reality, (a) are not correlated, or (b) are correlated to a lesser extent than that reported, or (c) are correlated in the opposite direction from that which is reported" is called illusory correlation by Chapman (1967, p.151). The over-estimation of strong associates in the Chapman and Chapman study suggested the systematic errors are produced by variables inherent in the stimuli observed and based on verbal associative connections of the test sign to the symptom rather than on valid observations. As such, "... entirely naive observers who view psychodiagnostic material would report the same erroneous correlates of patients' symptoms" (1967, p.193-194).

To check this Chapman and Chapman (1967) reported a series of studies which demonstrated illusory correlation in judgements from human figure drawing. After determining that practising
psychodiagnosticians agreed with each other beyond a chance level concerning which Draw-a-Person figures were correlates for various symptoms, a set of figure drawings by 'psychotics' and 'normals' were assembled. From a list of six symptom statements, two statements were paired with each drawing in such a way that each of the six statements occurred equally often with each drawing characteristic. Undergraduates were exposed to the drawing-symptom combinations and were then asked to report which drawing characteristics occurred most often with various symptoms. The results showed that subjects produced massive illusory correlation and that the correlates they reported were very similar to those previously reported by the experienced psychodiagnosticians. For example, "... 'eyes' are a stronger associate to 'suspiciousness' than to any other symptom" (p.200). The Chapmans also determined that the systematic misperceptions were related to the strength of associative connections between various drawing characteristics and symptom statements. Further while the illusory correlations phenomenon could be reduced slightly by increasing the subject's motivation to attend to the task and by allowing unlimited time to study the stimulus materials, the misperceptions were markedly resistant to modification.

In a later study Chapman and Chapman (1969) demonstrated that illusory correlation was present when subjects were asked to judge homosexuality from Rorschach responses. They found that expert diagnosticians, on the basis of their clinical experience, and naive judges, on the basis of their observations of random materials, tended to report the same Rorschach signs as being valid indications of homosexuality. This was found to be the case even though these
signs had not been demonstrated in the previous research literature to relate to the purported problem of homosexuality, nor did they have, in the experimental task, a non-random association with the purported problem of homosexuality.

Consistent with previous studies, Chapman and Chapman (1969) found that subjects picked those percepts that had the highest verbal associative connection with the symptom, in this case homosexuality. In addition, the Chapmans demonstrated that subjects still tended to report the popular invalid sign disproportionately even when it was paired randomly with the symptom "homosexuality" and the unpopular valid sign was paired 100 per cent of the time with the same symptom.

A Robust and Resistant Clinical Phenomenon

In an extensive replication of the Rorschach study Golding and Roser (1972) came to the same conclusion. They found little change occurred even when the non-illusory ('valid') cues were paired 100 per cent of the time with the symptom of homosexuality and when the illusory ('invalid') cues had a randomly paired relationship with the symptom of homosexuality. In addition they found that while the illusory correlation did drop significantly over a large number of trials, the phenomenon was remarkably resistant to modifications even under conditions designed specifically to modify it. Starr and Katkin (1969) applied the Chapman method to the Rotter Incomplete Sentence Blank. Clinical and non-clinical graduate students and naive undergraduate observers all displayed illusory correlation after exposure to sentence completion response-symptom combinations.
While the illusory correlation phenomenon has been demonstrated with three different projective techniques Dowling and Graham (1976) demonstrated massive amounts of illusory correlation in the reports of both graduate and undergraduate judges in an objective technique, namely MMPI sub-scale names. Lastly, Kurtz and Garfield found, in spite of the fact that attempts were made to influence the illusory correlation by providing a simulated training session for the subjects, it was not possible to reduce the illusory correlation. Under neither the 100 per cent nor the 50 per cent presentation conditions was training more effective than the non-training conditions in reducing the effect of this phenomenon. This finding is consistent with that of Chapman and Chapman (1969) and Golding and Roser (1972) with respect to the resistance to modification and extinction of the phenomenon, and when coupled with the other research, attests to the robustness of the phenomena. In all of these studies, the erroneous beliefs about relationships between test responses and particular psychosocial problems, that is, the illusory correlations, were interpreted in terms of pre-existing associative links between aspects of the test response and connotations of the psychosocial problems.

In his original article, however, Chapman (1967) postulated a second basis for illusory correlations. In addition to finding over-estimation of the frequency of word-pairs for which there was a pre-existing associative link (lion-tiger), Chapman also found that the frequencies of word pairs made up of atypically long words (e.g. blossoms notebook) were also over-estimated. Chapman interpreted the later finding as being due to the "distinctiveness" of pairing two long words in a series in which most of the word-pairs were two short words or one short and one long word.
Recently Tversky and Kahnemann (1973) postulated a judgemental heuristic or informal decision-making criterion to explain both sources of illusory correlations. They pointed out that the judgement of how frequently two events co-occur could be based on the strength of the associative bond between them. When the association is strong, one is likely to conclude that the events have been frequently paired. Consequently, strong associates would be judged to have occurred together frequently. The ease with which instances of a particular class of events can be brought to mind has been termed availability by Tversky and Kahnemann (1973), who pointed out that "Availability is an ecologically valid clue for the judgement of frequency because, in general, frequent events are easier to recall or imagine than infrequent ones" (p.209). Repetition, however, is not the only influence affecting the availability of a particular class of events. Anything that makes the particular class more salient will make that class more available. But since availability is often poorly correlated with frequency or probability, systematic error and bias in judgement inevitably result.

Availability and Illusory Correlates

According to this view, the illusory correlation between, say suspiciousness and peculiar drawing of the eyes found by Chapman and Chapman (1967), is due to the fact that suspiciousness is more readily associated with the eye than with any other part of the body. To return to the earlier study (Chapman, 1967) by way of an additional example, in judgements of the frequency with which pairs of words had been presented, subjects erred by over-estimating the
frequency of pairs such as 'Bacon-Egg', for which there were a pre-existent association and by over-estimating the frequency of pairs composed of atypically long words. Thus, rather than implicating two different bases for the formation of illusory correlates, Tversky and Kahnemann (1973) pointed out that these data may be viewed as indicating that:

"... illusory correlation is due to the differential strength of associative bonds. The strength of these bonds may reflect prior associations between the items or other factors, such as pair distinctiveness, which facilitates the formation of bonds during learning. Thus the various sources of illusory correlation can all be explained by the operation of a simple mechanism - the assessment of availability or associative strength." (p.224)

We could speculate that representations of the client behaviourally confirming the clinician's hypothesis are more cognitively available than representations of the client violating the hypothesis. Thus there is every reason to believe that the clinician will over-estimate the likelihood that the client will, in fact behave in ways that confirm the hypothesis. Consequently, by virtue of contemplating the forthcoming interaction with the client in the light of the hypothesis, the clinician not only will find it easier to think of the client confirming the hypothesis but will believe that these hypothesis-confirming actions will occur in greater numbers and that these hypothesis-confirming behaviours will be representative of the target's 'true' personal nature (Snyder and Swann, 1978).

In a very impressive series of papers Tversky and Kahnemann (1971, 1973, 1974; Kahnemann and Tversky, 1972, 1973) have elegantly shown how other every-day heuristics of inference may
bias the judgements of lay persons and clinicians alike. These investigators have demonstrated that intuitive predictions and judgements deviate markedly from the dictates of the conventional statistical model. In addition to the availability heuristic, the adjustment (Tversky and Kahnemann, 1974) and the representativeness (Kahnemann and Tversky, 1973) were advanced as additional heuristics governing intuitive prediction and judgement.

The Representativeness Heuristic

The representativeness heuristic, for instance, is reflected in the tendency to predict that outcome which appears more representative of salient features of the evidence while ignoring conventional statistical criteria. That is, individuals tend not to detect bias in their quantitative estimates because they lack a code for such judgements. They are prejudiced by a 'representative fact': they take a resemblance between A and B as the basis for inferring a relationship between them. A pervasive bias of this kind holds for most individuals over a wide range of statistical concepts. The employment of this approach to the judgement of probability leads to serious errors because similarity, or representativeness, is not influenced by several factors that should affect judgements of probability. The representativeness heuristic leads one to predict incorrectly extreme values and low probability events when they happen to resemble what one is trying to predict.

Consider, for example, the effects of base rates. There is a growing literature indicating that people, including clinicians, ignore base rate information in making probabilistic judgements
(e.g. Lyon and Slovic, 1976; Tversky and Kahnemann, 1974). According to normative statistical theory, this tendency can result in large mistakes. The implication of the persistence of the illusions of validity are equally serious. By way of an example, consider the clinician who, on the basis of Rorschach test data of in-patients diagnosed as schizophrenic noted that his/her diagnosis agrees with the staff diagnosis on about 70 per cent of the cases. Without knowledge of the base rate, this may seem to be indicative of accurate judgement. However, if the base rate was 70 per cent (i.e. the number of cases diagnosed as schizophrenic was 70 per cent) then the agreement rate of 70 per cent would not look so impressive. Hence his/her diagnostic work did not exceed the base rate for diagnosis of schizophrenia, and diagnosing every additional patient as schizophrenic, would have been as accurate as diagnosis derived from the Rorschach examination. The lesson is that accuracy of judgement should be evaluated as marginal increase in the 'positive hit rate' over the base rate. If the clinicians choose not to use the marginal rate evaluating their judgements they are likely to over-estimate the judgement ability. As Einhorn and Hogarth (1978) pointed out in an extreme example - if the base rate was .75 with judgement ability and treatment effects both of zero, then the positive hit rate would be .75 (equal to the base rate).

The representativeness heuristic seems to be widely applied when evaluating the probability of obtaining a particular result in a sample drawn from a specific population. In their research Kahnemann
and Tversky (1972) have shown that most subjects are dominated by the same proportion and are essentially unaffected by the size of the sample, which plays a crucial role in determining the probability of an outcome. Most judges assess the probability of an outcome to be the same in small and large samples, presumably because these events described by the same statistic and are therefore equally representative of the general population. In contrast, Tversky and Kahnemann (1974) pointed out that the sample size is crucial in determining the probability of obtaining a given result drawn from that population. Similarly intuitive judgements are dominated by the sample proportion in estimation of posteriori probability, that is, of the probability that a sample has been drawn from one population rather than from another.

Rosenthal and Gauto (1963) have demonstrated that research psychologists appear to misunderstand the statistical basis of hypothesis testing. In their study the psychologists claimed that they would place greater 'confidence' in experimental results with a larger sample size than with a small sample size when the probability value of rejecting the null hypothesis was held constant. This is a mistake because the probability of rejecting the null hypothesis automatically increases as a function of the number of observations. Hence, rationally, more 'confidence' should have been attributed to the experimental results with a small sample size. The probability value is evidently wrongly interpreted as a measure of 'confidence' rather than reflecting on a prior decision on which to reject the null hypothesis. This error is not inconsistent with Tversky and Kahnemann's emphasis on the importance of large samples in research.
One of the most obvious applications of the representative heuristic is to shortcomings in dealing with problems of regression. Regression effects occur when there is an imperfect relationship between judgements and criterion values. People expect and predict behaviours and outcomes on variable B to be as 'distinctive' or deviant from the norm as the predictor of variable A, and they are surprised and often disturbed by the phenomenon of 'regression to the mean'. In fact, they are prone to invent spurious explanations for events that, in reality, are simple regression phenomena (Kahnemann and Tversky, 1973). For example, when actions are given to extreme groups (as measured by some A), outcome (B) will be regressive with respect to A. However, unless one understands the regressive nature of the environment, it is easy to incorrectly attribute outcomes to actions. Furthermore, regression effects are symmetric, that is, A is also regressive with respect to B. Again, Tversky and Kahnemann (1974) emphasised that inadequate understanding of regression is not restricted to non-scientists. Further, when actions are based on judgements, the regression phenomenon produces several non-intuitive outcomes (Einhorn and Schacht, 1977).

The representative heuristic also helps to account for other common judgemental embarrassments, which are found in the literature on clinical inference - reviewed earlier. Indeed the 'illusion of validity' (Kahnemann and Tversky, 1972, p.249) arises and persists because the very factors that enhance the judges' subjective confidence - such as the consistency and extremeness of the data - often are in fact correlated negatively with the accuracy of
predictions, creating the paradox of confident predictors who persist in practices that are objectively unjustifiable. In fact, Kahnemann and Tversky (1973) have shown that people are most confident in judgement when information is consistent and/or extreme, even though these factors should induce them to decrease confidence in judgement. Indeed they stated:

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

The Adjustment Heuristic

The last heuristic is referred to as the adjustment heuristic and its use leads one to make estimates and predictions by 'adjusting' either some salient initial value or the result of some partial computation procedure. Such adjustments, however, are rarely sufficient, and the result is typically an 'anchoring effect'. Tversky and Kahnemann argued convincingly that over-estimation of likelihood for conjunctive events (i.e., the likelihood of $A$ and $B$ and $C$ all occurring) and under-estimation for disjunctive events (i.e. the likelihood of at least one of $A$ or $B$ or $C$ occurring) are further results of the intuitive statistician's failure to adequately adjust preliminary or partially computed estimates.

It appears that in our drive to confirm our causal model we easily twist data to make them fit poor models and are reluctant and slow to revise the models themselves (Tversky and Kahnemann,
1979). These judgement frailties and particularly tendencies to
be overly influenced by the representative heuristic and under-
influenced by such considerations as sample size and base rates,
are not limited to wide-eyed 'gut' clinicians and naive lay
persons: Tversky and Kahnemann (1971) demonstrated similar
tendencies among sophisticated mathematical psychologists. If any
general conclusion can be drawn at this point it is that even highly
educated adults rely upon intuitive considerations rather than the
appropriate abstract principles. In the absence of the appropriate
code, individuals rely upon their own experiences, derived from
familiar content, in making an inference or judgement. Hammond
(1978) has correctly pointed out that as one leaves the experimental
'mode of inquiry':

"Inability to hold certain variables constant, and
to manipulate other variables leaves the question
of causal directions ambiguous. As a result,
interdependent variables must be disentangled
sheerly by cognitive activity, that is, by reaching
a judgement about what the results of disentangle-
ment might be ... for the disentanglement of causal
relations by (passive) cognitive instead of (active)
experimentation is subject to a variety of
psychological factors, such as memory loss, information
overload, and recency and primacy effects, to mention
only a few." (p.16)

What we have seen is that the difficulty of disentangling variables
by unaided judgement is important, but equally important is the
fact that intuitive judgement is frequently based on everyday
heuristics of inference that almost invariably bias the judge.
PREMATURE CONCLUSIONS IN CLINICAL ASSESSMENT

In addition to the errors that can be introduced into the assessment and interpretation of our clients discussed above, other variables have been delineated that influence clinical judgement. For example, a number of studies have indicated that premature conclusions about what another person is 'really' like can interfere with appropriate utilisation of new information about the person. Dailey (1952) concluded that "... premature conclusions on the basis of small amounts of data can apparently prevent the observer from profiting as fully from additional data as he (she) would without the premature decision" (p.142). Recall that research has indicated that clinicians form a firm and relatively resistant image of their clients within one to four interviews. If the premature decisions were based on unimportant information - due, say, to impression management - then the detrimental effect of premature conclusion would be exacerbated. Importantly, Dailey concluded that, in situations "... involving repeated evaluation by observers of the same persons, validity may be lost by the cumulative tendency of understanding to affect itself adversely; to diminish, as it is expressed" (p.151). This point will be returned to below, but first a recent study that builds on Dailey's conclusions merits discussion. Snyder and Urbanowitz (1978) demonstrated that social stereotypes (clinical labels?) exert a retrospective as well as prospective influence on our interpersonal perception and information processing. In their study subjects reconstructed the events in a case history in ways that supported and bolstered a later arrived at stereotype. They demonstrated that individuals selectively retrieve information
that reinforces current stereotyped interpretations about another person, and these reconstructive processes may serve to perpetuate acceptance of widely held but essentially inaccurate social stereotypes. No doubt a parallel phenomenon could be shown in the clinical assessment situation, for in the clinical-psychotherapeutic situation the clinician does make repeated observations and evaluations of the same person over time. Information revealed in later sessions may induce reconstruction of earlier information in ways that reinforce the current construction.

Long-term Memory 'Drift'

According to D'Andrade (1974) situations in which repeated observations and evaluations of the one subject are taken over time are "... subject to a special effect analogous to that of an illusion; an effect in which there is a reliable and systematic distortion of judgements" (p.161). D'Andrade's argument is that, when an observer relies on his or her long term memory in making assessments of a number of different people each of whom performed a number of different behaviours, the characteristics and behaviours that the observer considers to be similar are likely to be recalled as applying to the same person. That is, there is a tendency for observers to remember similar behaviours as having been performed by the same actor under such circumstances, then over a few months the therapist will build up partially erroneous pictures of various clients.
D'Andrade (1974) in an analysis of the patterns of correlations among the various attributed behaviours in both the observers' and participants' memories for what occurred in group interactions revealed a systematic bias—reflecting a tendency to recall similar behaviours as having been performed by the same person, when in fact they were not. Shweder (1975) in a re-analysis of Newcomb's (1929) study of extroverted and introverted social behaviour among boys at a summer camp graphically evidenced the tendency of human memory to 'drift' in the direction of pre-existing conceptual schemata. As D'Andrade (1974) pointed out, the implication here is that, "With this type of memory error, any attempt to discover how human behaviour is organised into multi-behaviour units ... which is based on data consisting of long-term memory judgements will result in conclusions which primarily reflect the cognitive structure of the rater" (pp.175-177). Unfortunately, as Jones (1977) pointed out, one way of conceptualising the task of psychodiagnosticians and psychotherapists is to discover how an individual's behaviour is organised into multi-behaviour units.

In the foregoing discussion some of the important pitfalls that interfere with accurate coding and information processing in personality assessment have been examined. Hopefully, enough has been said to make the fallibility of the clinician patent. When this is coupled with the major problems revealed in the psychiatric system of classifying abnormal behaviour and the dubious validity of many assessment devices, the traditional approach to psychodiagnosis appears to be, if not untenable, at least fraught with difficulties. However, such a conclusion may be premature; the implications of the research presented to date will be discussed in the next chapter.
CHAPTER IX: CONCLUSION

Notwithstanding the ostensibly overwhelming negative evidence that attests to the inadequacy of the traditional psychometric approach to psychodiagnosis, research suggests that to a considerable extent the clinical practitioner still operates within a paradigm based upon the 'essence' concept of personality. To briefly recapitulate, according to this view, mental and personality structures and processes are considered to be the determinants of human behaviour. Despite their diversity most personality theories share a common assumption in that they conceive of personality as consisting of certain relatively stable and interrelated motives, characteristics and dynamics that underlie and are responsible for the person's overt actions. In order to fully understand why an individual behaves in a particular way, then, one needs to obtain a comprehensive understanding of the underlying dynamics.

Most psychometric tests and assessment devices are designed to be measuring instruments to index the internal traits and personality structure of the individual, and, importantly, many clinicians indicate that they use psychological tests precisely for that reason. For example, Wade and Baker (1977) found that clinicians select and use tests for the information they provide concerning personality structure. Thus, not only do clinicians view as relatively unimportant the two criterion measures of test performance in research (i.e. behavioural prediction and diagnostic accuracy) but they appear to operate within a superseded paradigm or at least one which is becoming increasingly unpopular from the academic researcher's perspective (Ekehammar, 1974).
The central aim of this essay was not only to highlight the substantial disparity between psychodiagnostic practice and the body of empirical data evidencing its shortcomings, but to delineate plausible explanations for the failure of clinical psychologists to adapt themselves to the imperatives of the literature. By way of an example, consider the clinicians failure to adequately respond to one of the oldest and most convincingly corroborated generalisations to emerge from psychological research, namely, the relative specificity of the structure of psychological traits. The finding that correlations between various manifestations of any psychological trait are substantially lower than the reliability of those manifestations (Shapiro, 1979, p.211) antedates Mischel's (1968) expose by at least three decades. For example, Spearman (1932) was one of the first to discuss specificity; it was embodied in his formulation that all the various manifestations of intelligent behaviour involve a general factor and a very large number of factors specific to each manifestation. Similarly, Hartshorne and May (1930) demonstrated that the intercorrelations between various reliable tests of honesty in young children approached zero.

The failure of clinical psychologists to adapt themselves to this situation is an outstanding anomaly. Clinicians continue to try to confine their assessment operations to general traits, and in so doing they concern themselves therefore with only a relatively small part of the reliable variance of behaviour and experience. When one places the great variety of human phenomena in the context of relative specificity, one is led to expect a unique pattern of dysfunctions and related variables in each individual client. The implications of this assertion have not made a significant impact on clinical practice.
Conventional psychometric approaches which aim to measure common variance only, and not specific variance, cannot therefore be expected to meet our clinical needs and will inadvertently serve to maintain what Mischel (1979) identified as the 'prevalent form of clinical hostility' (p.245).

In the earlier chapters several other important research/practice anomalies were discussed along with a delineation of the factors and processes that maintain what appears to be untenable practices. However, before considering what may be done to change the situation, as this appears desirable, it may be timely to discuss briefly what factors influence the clinicians' initial orientation to psychodiagnostic assessment.

Influences on Clinician's Testing Practices

In their recent survey Wade and Baker (1977) found that 70.7% of their respondents first learned to use assessment instruments in graduate training, 27.7% in undergraduate training and only 1.6% in postgraduate employment. These figures strongly suggest that initial training instructions are to be held, at least partly responsible for the current situation. This is augmented by other findings, for example, only 20.6% of respondents indicated that they systematically collected and analysed data regarding their own testing practice. Additionally, although 54.7% indicated that they read at least several articles relevant to testing every six months, only 25% stated that the studies critical of tests seemed accurate. The majority of respondents indicated that these studies employed inappropriate criteria or questionable methods, that they over-generalised or that they reported conflicting findings.
From the above and other research (see e.g. Garfield and Kurtz, 1974; Shemberg & Keeley, 1970) it appears that the initial training of clinicians in Universities and Colleges fails in at least two important respects. Firstly, instructors continue to teach students to administer and interpret tests of dubious validity, and secondly, they fail to develop behaviours compatible with the ideals embodied in the concept of 'scientist-practitioner'. It is beyond the scope of this essay to undertake an analysis of the factors that could explain this situation but it should be reiterated that the presentation of negative evidence is invariably insufficient to dissuade neophytes from using the 'offending' tests especially when they are faced with the role expectations of mental health colleagues. A viable alternative must be taught to a high level of competence if clinical training is to have a marked effect on assessment practices. Additionally, however, a repertoire of social skills (i.e. assertion, diplomacy, etc.) and cogent rationale may be equally important if the new graduate is to implement, sustain and evaluate an alternative approach to psychodiagnostic assessment.

In the light of the importance of initial training of clinicians for later assessment practices, this conclusion focuses specifically on the implications of the earlier chapters for the teaching of clinicians. The specific aims are to: Firstly, recommend ways of improving clinical assessment conducted within the dominant approach. This will be intentionally brief as it is the author's view that the paradigm is not the 'best' available. It has been argued that the model is not only conceptually and empirically inadequate but also facilitates the medicalisation of deviance and supports a 'blame-the-victim' ideology (Ryan, 1971).
Secondly, an already available alternative assessment paradigm, namely, behavioural assessment, will be briefly introduced. Again its presentation will be brief as it is considered to be predicated on an equally limited and truncated image of humans. However, it will be argued that many of the problems associated with the attribute approach - delineated in the earlier chapters - are avoided when the behavioural model is applied. Thus it will be argued that it is a preferable approach.

Lastly, an approach to assessment predicated on a recognition of the transactional nature of the person-environment relationship - a perspective requiring that the person be viewed at the psychological level as a component of the complex-system of which he/she is a part - will be advocated. This interactionalist perspective does not represent a rapprochement between the opposing theoretical camps represented in the two alternatives listed above (i.e. personologism and situationalism) but a synthesis of the dialectic. It represents a dialectical synthesis that should result in the eventual abandonment of both prior positions. The recognition that complex human behaviour tends to be influenced by many determinants and reflects the almost inseparable and continuous interaction of a host of variables both in the person and in the environment, has deep implications for psychodiagnostic assessment.

WITHIN PARADIGM RECOMMENDATIONS

To this point the essay has been negative in the sense that it has delineated and discussed the numerous conceptual and empirical weaknesses of the 'attribute' model. However, it is possible to convert some of the critical findings into recommendations that
may help clinicians avoid some of the identified shortcomings. Although this may amount to little more than 'mending a thread bare' paradigm it cannot be disputed that the attribute model remains the dominant assessment model and as such, until a viable alternative is accepted, the onus is on trainers and researchers to develop strategies that diminish the errors and biases in diagnostic assessment conducted within the paradigm. Thus in this section several recommendations for what should be included in the training course for students in psychodiagnostic assessment will be made. They include suggestions on how to improve the clinician's ability to learn from experience, on how to sensitise clinicians to the importance of the cognitive structure and strategies they bring to the assessment situation, on how to sensitise the clinicians to the influence of their values and biases in clinical judgement, on ways to sensitise clinicians to unfamiliar people, and on how to give up the luxury of global labelling.

**Improving the Clinician's Ability to Learn from Experience**

In Chapter VIII the many information processing pitfalls in psychological assessment were enumerated. The difficulties centre to a very large degree on the disinclination to think probabilistically and make use of disconfirmatory evidence. Although the fallibility of the clinician is now well documented an important question that remains largely unanswered is the degree to which these shortcomings can be alleviated or at least attenuated through systematic training.
An obvious exception, however, is the research on the effects of training and the offering of inducements in the search for and utilisation of disconfirmatory evidence. The research with regard to this is disheartening. Certainly formal training in experimental design, teaching the logic of central groups and base line predictions, and so on, would seem a reasonable suggestion but as we have seen it has been demonstrated that statistically sophisticated subjects make similar mistakes to those without training (Einhorn and Hogarth, 1978). Further, even when explicitly instructed to do so and when rewards were made contingent upon disconfirmatory strategies, little improvement was found (Wason and Johnson-Laird, 1972). The prospects for overcoming this tendency appear bleak. It represents a real hurdle for those instructors committed to moving their students towards the ideals of the scientific practitioner model.

A second major factor in the clinician's inability to fully utilise feedback from their assessment endeavours is their lack of awareness of environmental effects on outcome. In order to overcome this deficit instructors should explore ways of increasing awareness of effects like regression, base rates, selection ratios and treatment effects. Einhorn and Hogarth (1978) suggest that the use of a model of the environment, such as that advocated by Hammond and his colleagues (Hammond, 1971; Hammond, Wilkins and Todd, 1966) may prove useful for this purpose. Such a model draws attention to the structure of the environment and the manner in which structure affects outcome.
Lastly, methods of improving the clinician's coding, storing and retrieving of outcomes should be included. Several years ago Goldberg (1968) suggested that a simple tallying of data relevant to a particular hypothesis (i.e. keeping a 'box score') may be useful. This may serve a dual purpose if disconfirmatory instances can be included. In the study of Jenkins and Ward (1965) the only condition that substantially improved the subject's ability to accurately infer causal relationships was when they were given the data in a 2 x 2 contingency table. However, when the table was given after subjects had experienced the set of instances and non-instances the table failed to improve their performances.

On the basis of their explorations with decision-making heuristics, Tversky and Kahneman (1974) suggested that people should attempt to encode events not by their substantive content but by judged probability. They point out that when events are grouped in this manner it is possible to keep a tally of the extent to which judged probabilities match subsequent empirical relative frequencies. Again they highlighted the inadequacy of feedback indicating whether an event did or did not occur for indicating one's ability to make probabilistic judgements. The most obvious outcome of this brief review is that it is a fruitful area for research. At this stage little in the way of empirically corroborated strategies, found to be efficacious in inoculating the clinician against the many pitfalls to information processing in psychological assessment, can be offered.

Sensitising the Clinicians to the Importance of Their own Cognitive Structures

It was argued earlier that the clinician and the client are faced with the same task, namely to discern, describe and apply
systematic covariation between events. To sensitise clinicians
to this and thus to the need to examine the cognitive structures
which they bring to bear on the assessment situation introduction
to the theoretical perspectives of George Kelly and/or Egon
Brunswik is recommended by Jones (1977). These two offer useful
perspectives on clinical judgement within which the evidence on
judgement fallibility makes sense and within which there is explicit
appreciation of the importance of the clinician's own beliefs and
expectations.

The basic philosophical position underlying Kelly's (1955)
perspective which is termed 'constructive alternativism', is very
compatible with the teachings of Popper and Kuhn and thus may
present the clinical instructor with a useful link between philosophy
and practice. Kelly asserts that:

"... whatever nature may be, or however the quest
for truth will turn out in the end, the events we
face today are subject to as great a variety of
constructions as our wits will enable us to contrive ...
Events do not tell us what to do, nor do they carry
their meanings engraved on their backs for us
to discover. For better or worse, we ourselves
create the only meanings they will ever convey
during our lifetime" (Kelly, 1966, cited in
Bannister and Mair, 1968, pp.6-7)

Kelly assumed that in the attempt to bring order to the events
surrounding them, persons including clinical psychologists, will
develop their own idiosyncratic, hierarchically organised system of
personal constructs. A construct is simply a dichotomous abstraction
that the person has made of the similarities and differences among
events in the onrushing stream of stimuli surrounding them. The key,
of course, is that different people make different abstractions from
the 'same' events. To say that a person's system of constructs is hierarchically organised means that some constructs are more important than others and it is the interrelationships among constructs that are of central concern. It is important to note the similarities between Kelly's notion of a hierarchically organised set of personal constructs and the notion of an implicit theory of personality. The relationships among constructs define a network of interrelated expectations or anticipations that is essentially the same as the correlation matrix with its underlying dimensional structure. It is equally important to note the idea that one's perceptions are tentative and subject to reconstruction is basic to Kelly's argument.

A related perspective on clinical judgement has its origins in the work of Brunswik (1952, 1956). Like Kelly, Brunswik's focus was on the fact that different people may see the 'same' situation in different ways. His basic point about perception was that the perceiver's knowledge of the environment is of a probabilistic nature. His system or theoretical stance is referred to as 'probabilistic' functionalism and interestingly a revival of interest in Brunswik's 'lens' model will not doubt follow its promotion by Petrinovich (1979) as an alternative paradigm for psychology – an alternative to the ideal of the linear process experiment paradigm.

It is beyond the scope of the essay to fully introduce either of these theoretical perspectives or to outline others (e.g. Sarbin, Taft and Bailey, 1960); merely to suggest the utility of these perspectives for sensitising clinical trainees to the salience of their own cognitive structures in any assessment.
situation. A real advantage of both viewpoints is that the models apply equally well to understanding the clinician and client in that both are represented as utilising similar cognitive strategies in organising their experiences of people and, hence, developing certain implicit expectations and beliefs. To phrase this in Brunswikian terms, the clinician develops a set of cue-utilisation coefficients, which may or may not match the ecological validities of the cues and may or may not correspond to the cue-utilisation coefficients of the client with whom he or she is interacting. In Kelly's terms, both the clinician and the client have systems of personal constructs, which may or may not be organised in the same way.

Scientific Knowledge and Social Values

In line with Kelly's notion of 'constructive alternativism' students of clinical assessment should not harbour illusions about the veracity of their clinical inferences or the 'realness' of the constructs. I believe that through an exposure to the ideas of Popper, Kuhn, Lakatos, Mitroff and other important philosophers of science (e.g. Feyerabend, Bartley) students of clinical assessment may learn to be more circumspect and conditional about their knowledge and hopefully learn to show considerable tolerance for statements that convey relativity and tentativeness.

As Cohen and Wartofsky (1978) have so aptly highlighted:

"... modern philosophy of science has turned out to be a Pandora's box. Once the box was opened by critical independent minds, puzzling monsters appeared: not only was the neat structure of classical physics changed (and partly by philosophical analysis within physics), but a variety of wide-ranging questions were let loose. Philosophy of science could not help but become
epistemological and historical, and could no longer avoid metaphysical questions, even when these were posed in disguise. Once the identification of scientific method with that of physics had been queried, not only did biology and psychology come under scrutiny as major modes of scientific inquiry, but so did history and the social sciences, particularly economics, sociology and anthropology. This trend raises anew a much older question, whether the conception of science is to be distinguished from the wider conception of learning and inquiry? Or is 'science' to be more deeply understood as the most adequate form of learning and inquiry, whose methods reach every domain of rational thought and action? Is modern science to be seen as matured reason, or is it simply one historically adapted and limited species of reasoning, and of western reasoning at that?" (p.2)

Exposure to some of the questions raised in the above quote may diminish the tendency to justify ones beliefs by recourse to claims as to their scientific status. Further, by exploring the alternative paradigms available to psychology (i.e. the Natural Science Model versus the Historical Model for example) trainees may, in a non-threatening way be introduced to the view that 'facts' and 'truths' emerge within 'specific places and times'. Thus, that sociohistorical influences are intrinsic to all knowledge, including our clinical judgements and beliefs may reduce resistance to the issue of the politics of clinical judgements. Just as Merton (1967) pointed out that the values embodied in the Protestant Ethic 'pervaded the realm of scientific endeavour [such that] the why and wherefore of science bore a point-to-point correlation with the Puritan teachings on the same subject' (pp.575-576), so Draguns (1974) has highlighted a similar set of values is reflected in the typology used to classify problematic behaviour.

Understanding the intrinsic fallibility of our attempts to know the world scientifically could well be a prerequisite for acknowledging our personal fallibilities and attempting to apply methodologies to improve the usefulness of our own attempts to generate knowledge.
Perceiver and Perceived Relationship

In a recent discussion of the literature on behaviour classification, Phillips and Draguns (1971) pointed out that research on the clinician's diagnostic activity seem to converge in suggesting that social distance may well be the most salient variable in systematic biases and judgement error. As we have seen the middle-class diagnostician is likely to employ different categories for lower class clients and is likely to note few symptoms in the process of assigning these clients to those categories. Many factors have been suggested to account for this phenomenon (see e.g. Adinolfi, 1971) but irrespective of these hypothesised mediators, the person perception literature does seem to have demonstrated the strong relationship between being actually similar to, or highly familiar with, the perceived person and accurately perceiving him/her.

Adinolfi (1971) suggests that '. . . notwithstanding contrary indications in person perception research and demonstrations in clinical relationships, there seems to be a persistent attempt to explain, categorise and characterise others in the face of pervasive unfamiliarity' (p.171). In practice clinicians may simply not have the requisite cognitive skills to comprehend those who are unfamiliar to them.

Even though it may prove to be inadequate, Adinolfi suggests that clinical trainees who plan or wish to do assessment or therapeutic work with those different from themselves should imbue themselves in the experiential products of members of the unknown groups. That is, rather than read the intellectual-conceptual
exercises of middle-class observers, students should go to expressive works (poetry, art, novels, cinema, etc.) by the people they are trying to understand.

The literature on the impact of other client irrelevant influences on assessment (i.e. sex, colour, political beliefs, etc.) evidence that awareness of their impact may well lead to a diminution of their effect. Consequently, exercises that are designed to confront trainees with their own biases may be helpful in reducing undesired systematic errors. It will be important to emphasise that the goal is not to achieve the impossible state of being value-free but to explicitly incorporate desirable values and biases.

Giving up the Luxury of Labelling

If the primary function of psychodiagnosis is to describe the client's problematic behaviour in ways that facilitate the selection of an efficacious treatment regime then it is doubtful that the currently available psychiatric labels have sufficient utility to justify their usage. Additionally, as Mischel (1968) repeatedly claimed, to do justice to the complexity of those we are trying to comprehend, clinical psychologists will have to divest themselves of the luxury of categorisation which too readily leads to an unwarranted extrapolation of correlative variations.

Failing this, instructors should highlight the pragmatic, provisional and metaphoric nature of classificatory typologies. The diagnostic labels should be 'de-reified' and the all important issue of 'who defines' should be omnipresent in any course of psychological assessment. Here too, the teachings of contemporary philosophers of science may prove useful. That is, following the logic of Popper it can be asserted that no one classification can
be said to equal the original data. In consequence, therefore, it would seem important to have more than one set of categories at one's disposal, to avoid confusing any one category with the unique event.

No doubt, more recommendations could be made but in the limited space available the aim was not to be all inclusive but suggestive. Fortunately, for the instructor who decides to operate within the attribute model the vast research literature throws up many suggestions that may lead to the development of more sensitive, humane and efficacious clinical psychologists.

BEHAVIOURAL ASSESSMENT

In an attempt to simplify recent developments in personality assessment, researchers have distinguished between three different conceptions of human behaviour that seem to reasonably represent the contemporary scene. Although heterogeneous and to some extent overlapping, the three theoretical positions (personologism, situationalism and interactionism) can be readily distinguished. It should be noted that there are differences in points of view within each model and similarity between the models. For the present purpose, the typical and in some sense a caricatured position for each model is presented, with the focus on their differing points of emphases and the implications of these for assessment.

The traditional attribute model under the above typology, is an example of personologism. In contrast situationalism can be regarded as the antithesis of personologism with its emphasis on
environmental (situational) factors as the main source of behavioural variation. Although individual differences are recognised, situationalism in practice tends to focus almost exclusively on observable events, current behaviour and situational determinants (O'Leary, 1979). This almost exclusive focus on the immediate environment, particularly in the late 50's and early 60's was possibly in reaction to the personologism.

Contemporary social learning theories (see e.g. Bandura, 1969, 1977; Mischel, 1968, 1979) redirected the focus from the situation and the hypothetical underlying attributes in the organism to what the organism does in particular situations. The concept of personality from this viewpoint now becomes more or less a shorthand term for summarising the sum total of an individual in his/her social environment. Thus an individual does not have a personality, but rather the concept 'personality' is an abstraction that one may make after observing a person interacting in a comprehensive sampling of situations. From within a behavioural framework, personality is viewed as a set of learned capacities (cf. Wallace, 1968). Mischel's (1973) recent cognitive approach to personality represents probably the most useful behavioural approach to the concept.

Although at the theoretical level social learning theorists have emphasised the reciprocal interaction of the person and the situation, rarely has this commitment to reciprocal causality manifested itself in practice (an outstanding exception is the recent work in self-control). In practice a rather static, unidirectional model of causality where the individual is construed as relatively passive seems to dominate. Parenthetically, this
parallels the difficulties with the trait concepts that Hogan et al (1977) claim are a function of the individuals who apply them rather than being attributable to inherent qualities of the concept.

The Nature of Behavioural Assessment

Although the recent spate of books devoted to behavioural assessment (Ciminero et al, 1977; Cone and Hawkins, 1977; Haynes, 1978; Hensen and Bellack, 1976; Mash and Terdal, 1976) and the even more recent launching of two specialist journals, may be taken as indicators of the rise and widespread acceptance of a new approach, behavioural assessment is not a new field at all. Because assessment and continuous monitoring of the process of change are essential to the behaviour therapy model, behaviour assessment is as old as behaviour therapy. In a more general sense it is not new since at its heart, the goals of behaviour assessment are to identify meaningful response units and their controlling variables for the purpose of understanding and altering behaviour. Recent developments merely represent the emergence of behavioural assessment as a defined area of professional interest in its own right.

However, surprisingly, even at this stage no definition or conceptualisation of behavioural assessment has been widely adopted. Most attempts at definition have been put forth in relation to, and usually by way of comparison with the attribute model. Comparisons have been made with respect to the different underlying assumptions regarding the concept of personality and test construction, interpretation and use. This is understandable as the legitimate work of criticism between paradigms. This amounts
to offering an alternative viewpoint, one which is an adequate representation of the phenomena and which is discontinuous in the sense of being incompatible with the original position.

For example, in discussing 'behavioural personality assessment' Goldfried and Sprafkin (1974) state:

"... the point of divergence between traditional and behavioural approaches lies in the assumptions underlying the construction and interpretation of these assessment methods. A delineation of the contrasting assumptions should provide ... the general principles of behavioural personality assessment. As indicated by Goldfried and Kent (1972) a comprehensive distinction between traditional and behavioural personality assessment must involve the assumptions underlying (a) what is basically involved in the definition of 'personality', (b) the selection of the items to be included within the test, and (c) the approach taken to the interpretation of the individual's responses." (p.3)

The differential assumptions relating to each of these areas are elaborated by Goldfried and Kent (1972). Bersoff (1971) in discussing 'psychosituational assessment' states:

"... to obtain a better sample of typical functioning it may be necessary to pursue techniques more appropriate to that end. These techniques can be considered within the framework of what the author calls a 'psychosituational assessment' - defined as an assessment designed to measure and interpret behaviour as it is elicited through interaction with specific stimulus situations. A number of specific methods have been constructed to accomplish such a goal. (p.392)

According to Bersoff (1973) these methods include primarily the systematic observation of behaviour in its usual context and the design of assessment situations which resemble the more natural settings in which behaviours occur. These methods attempt to minimise the amount of inference necessary in generalising assessment information to an individual's usual life circumstances. In
a recent paper O'Leary (1979) suggested that historically there have been three important hallmarks of behavioural assessment, viz., emphases on observable events, on current behaviour and as situation-al determinants of problematic behaviour. He stresses that the hallmarks should be seen as 'emphases', not rigid structures that may confine creativity in assessment situations.

Because behaviour therapy and behavioural assessment are part of one and the same process (Franks, 1978; Yates, 1971) assessment must be capable of identifying the environmental variables - internal as well as external, self-imposed as well as imposed by others - that are currently maintaining the individual's problematic thoughts, feelings and behaviour. This requires analysing both the specific environmental events controlling the problematic response repertoire and the broader social learning and biological histories that are mediating or otherwise influencing the way in which the environment modifies behaviour.

To think within the behavioural model requires that the clinical psychologist radically alter his/her concepts and role. Firstly, as Franks (1976) points out he/she must come to a view of assessment not merely as a prerequisite to therapy - the customary role of traditional testing - but as an ongoing process interwoven with the process of therapy. Secondly, the clinical psychologist must come to see him/herself not as a team member performing a circumscribed task, first testing then therapy, but rather as a professional for whom assessment and interaction are intrinsically integrated.

However, as Nelson and Haynes (1979) correctly point out, the novelty of the field of behavioural assessment lies not so much in its goals and strategies "... but rather in its deliberate attempt
to improve the identification and measurement of dependent variables, to increase the probability of selecting successful treatment techniques, and to re-define the evaluation of those intervention procedures" (p.2). The strength of the behavioural assessment approach is not to be found in any single element, but rather in the relationships and similarities of concepts, methods, and practices that characteristically occur. As Mash (1979) points out, this notion parallels the concept of 'family resemblance' as discussed by Wittgenstein, and to use his analogy, "the strength of the thread does not reside in the fact that some one fibre runs through its whole length, but in the overlapping of many fibres (Wittgenstein, 1968, p.32, cited by Mash, 1979).

Central Aspects of Behaviour Assessment

In the view presented above, there are no rigid boundaries for what differentiates behavioural assessment from alternative approaches. Although continued efforts at definition are likely to do with the methodology and practical validity of behaviour assessment as a potential contributor to both theory and the solution of applied problems. Several of these issues will be highlighted in the brief comments below.

Rejection of traditional diagnostic systems: The eshewing of psychiatric labels by behavioural assessors was recently shown to have important implications in practice. The fact that labels create sets that influence subsequent perception has been already documented, but recently Langer and Abelson (1974) studied the effects of labels on clinician's judgements in a 2 x 2 factorial design. Clinician's representing two different schools of thought,
behavioural and analytic, viewed a single videotaped interview between a man who had recently applied for a new job. Half of the group was told that the interview was with a 'job applicant', while the remaining half was told that he was a 'patient'. At the end of the videotape, all clinicians were asked to complete a questionnaire evaluating the interviewee. The important finding was that the behaviour therapists were apparently immune to the biasing effects of the label itself, i.e. patient versus job applicant. Although this finding is open to alternative interpretation it is in line with the substantial literature on person perception and does suggest a possible solution to many of the problems discussed in Chapter VI.

The rejection of traditional diagnostic systems within behavioural assessment has created a need to develop an alternative, because some classification framework is required as a guide for theory, research and practice. The need for alternative classification systems was recognised early in the development of behaviour therapy (see e.g. Ferster, 1965) and has received continued recognition (e.g. Adams et al, 1977; Dengrove, 1972). In an obvious attempt to develop a diagnostic and classification system for behavioural assessors that serves a function similar to those of the Diagnostic and Statistical Manuals of the American Psychiatric Association, Cautela and Upper (1975) produced a total of 283 behavioural classifications on their Behavioural Coding System. A client would be diagnosed as having certain responses needing alteration, and would be classified by listing the Behavioural Coding System numbers of these behaviours.
A behavioural diagnostic system might eventually include information as to the relative effectiveness of various intervention strategies. The recent article by Kanfer and Grimms (1977) makes a start in this direction.

**Behavioural approach to abnormality:** From the behavioural perspective, behaviours traditionally called abnormal are no different, either qualitatively or quantitatively, in their development and maintenance from other behaviours (Ullman and Krasner, 1975). That is, no behaviour is ipso facto abnormal or an indicator of 'mental illness'. From this perspective human behaviour is not dichotomous but rather can be understood and dealt with through a single set of principles.

Behaviour that is reacted to as 'abnormal' is seen to be the reasonable outcome of past and present person-environment interactions. As Masserman (1976) pointed out: "In effect many of the amazingly diverse patterns of behaviour we elicited [in the laboratory] could be labelled 'experimentally neurotic', 'addictive', 'sociopathic', or even 'psychotic', yet could also be considered adaptionally 'normal' in the cogent sense that, with due regard to genetic proclivities, they were the product of the unique-experiences of the individual" (p.ix)

**Identification of Controlling Variables:** Besides identifying the behaviour to be changed, a second goal of behavioural assessment is to identify the variables that control the occurrence of that behaviour. The reasons for doing so are to understand and explain the occurrence of the behaviour and to use the controlling variables that can be altered for therapeutic gains.
The current working model within behavioural assessment identifies two major classes of controlling variables: Current environmental variables (antecedent and consequent stimuli) and organismic variables (individual differences produced by physiology and by past learning). The components to be considered within behavioural assessment have been neatly summarised by Goldfried and Sprafkin's acronym (1976) of SORC (stimulus-organism-response-consequence). In any particular case, the controlling variables needing further assessment are frequently suggested by the nature of the problematic responses (or the diagnostic label assigned to covarying responses). For example, schizophrenic behaviour may lead to inquiries about family history of similar disorders or to questions about this individual's own history with regard to social withdrawal. Again, the nature of the problem behaviour or the diagnostic label provides only nomothetic suggestions of controlling variables for the clinician to consider in an idiographic assessment.

Situational specificity and temporal consistency: The case for the specificity of behaviour has been put and need not be repeated here. However, as Epstein (1979) indicated it is common to find a good deal of temporal consistency to behaviour. The implications of these two issues for assessment is that consistency across situations or across time cannot be assumed. Thus the technique of behavioural assessment must include careful consideration and description of the relevant stimulus situation, in order to fully assess the response to interest. Without independent empirical support, conclusions should not be generalised from one situation to another, nor from one response type to another, nor from one time to another. Similarly because the assessment technique itself is part of the
situation, generalisations are not automatically made from one assessment technique to another.

Triple Response System: Behavioural assessments have also emphasised the need for multiple sources of information, especially in light of the non-convergence of different types of assessment such as verbal report, observation, and situational tests. Today, most behaviourists consider all forms of organismic activity to be 'behaviour'. Thus behavioural assessment also includes the measurement of physiological-emotional behaviour and cognitive-verbal behaviour. Although these three types of behaviour may covary, it is not assumed that such covariation occurs. In fact, when contemporaneous measurements are taken from all three response systems, the correlations are often only moderate. There are, however, few guidelines for how such diverse and sometimes discrepant information sources can be integrated in formulating intervention programs. Questions regarding the relative weight to be given to different types of assessment information in decision-making are important, and in fact, the generation of decision-making models for the use of assessment information with specific populations is needed.

Although there are other 'threads' in the make up of behavioural assessment, there is insufficient space to discuss this issue further. Similarly, there is insufficient space to even briefly discuss the variety of behavioural assessment methods. Certainly direct observation in the natural setting is the *sine qua non* of behavioural assessment. However an elaborate set of procedures utilising self-
report data, self-monitoring, analogue and physiological measures have been developed. Detailed descriptions are available in any of the behavioural assessment handbooks cited earlier.

Before moving on to a discussion of the interactionalist position, however, two issues merit brief discussion. Firstly, the utility of behavioural assessment seems to be a persistent problem for practitioners. Careful functional analysis involving direct observation in natural circumstances, or utilising multiple settings and informants is not easily achieved. This was noted in a recent study by Wade, Baker and Hartmann (1979), which also showed that assessment was a small part of the activities of behaviour therapists, and that many continued to use traditional assessment devices. It would appear that the growth of behaviour assessment in principle, as evidenced by books, journals, and research, may be greater than the growth of behavioural assessment in practice. Developing a usable technology is a high immediate priority.

On this issue Haynes (1979) has argued that while an idiographic approach has certainly resulted in substantial improvements in conceptualisation of behavioural determinants of individual target problems and in identifying factors associated with treatment outcomes, it is costly and frequently unnecessary. Thus comprehensive pre-interaction behavioural assessment, when evaluated by a cost/benefit analysis, may be only warranted under some conditions. The utility of a comprehensive pre-intervention assessment is probably a function of two factors: (1) the clinical significance of the targeted behaviour and (2) the power of the intervention. However,
because behavioural interventions have most frequently been applied in research settings, there has been minimal pressure to develop cost/efficient assessment strategies that are adaptable to clinical settings.

Secondly, several prominent spokespersons have recently called for a rapprochement between traditional psychometrics and social learning theory (Evans and Nelson, 1977; Staats, 1975). They see the strict situational approach, personified in the functional analysis of the operant kind, too narrow and necessarily limited in application and suggest the dismissal of the psychometric testing tradition and diagnostic labels would be premature. Arthur Staats (1975) demonstrates how such a rapprochement could be made in a chapter called 'social behavioural psychometrics' in his recent book.

However, in the next section it will be argued that what is required is not a rapprochement between the opposing theoretical perspectives but a dialectical clash that should result in an eventual abandonment of both prior positions. This reflects a view that both situationism and personologism are the products of historical and cultural perspectives that are no longer relevant (Cronbach, 1975; Riegel, 1978; Sampson, 1978).

A PROLEGOMENON FOR INTERACTIONAL ASSESSMENT

So far we have discussed the two alternative paradigms that are seen by many to present antithetical views on the nature of people. Although fundamentally different, they are similar in sharing an emphasis on either the individual or the social environment as a passive force in the process of human development and change.
Personologism retains the 'Lockeian' notion of a passive environment but introduced the individual as an active agent. In contrast situationalism places the passive organism in an active environment. These views about the nature of people conveyed by these two perspectives require critical examination on conceptual or social grounds for as we have seen what we believe people to be affects which aspects of human functioning we study most thoroughly and which we disregard. Premises thus delimit research and are, in turn, shaped by it. As knowledge gained through study is put into practice, the images of people on which social technologies rest have even vaster implications.

In this section it will be shown, through the examination of interactionism, that both personologism and situationalism present a truncated and inadequate image of people. As Bandura (1977) cogently argues:

"... people are neither driven by inner forces nor buffeted by environmental stimuli. Rather psychological functioning is more adequately explained in terms of a continuous reciprocal interaction between personal behaviour and environmental determinants." (pp.11-12)

Although it is widely accepted among psychologists that the behaviour of a given person in a particular environment is a joint function of the relevant behavioural determinants within the person and those within the environment, few researchers appear to operate within the paradigm. As Mischel (1977) reminds us:

"... in the abstract this recognition seems as bland and obvious as a cliche, and one wonders if a focus on interactionism and multiple determinism may not be little more than a substitution of a new slogan for old varieties." (p.246)
It will be argued that when examined more concretely, this recognition has profound implications for psychology in general, as well as specifically for psychodiagnosis. For, if it is useful to construe behaviour as a joint function of person and situation determinants, then it follows that psychological assessment should systematically take account of both person and situational variables and their transactions. It need hardly be said that this is rarely the case in contemporary assessment practice.

**Interactionist Conceptions**

Although interactionist thought can be traced back to Aristotle, in Psychology explicit statements from this perspective were cogently put by Kantor (1924; 1926); Lewis (1936) & Murray (1938). Whereas the classical interactionist views were usually formulated within comprehensive personality themes, most often without empirical support, the more recent conceptualisations have usually been professed in the absence of any elaborate theories, but often with some empirical support. Some exceptions to this common picture are the recent attempts by Bowers (1973) and Mischel (1973) to base interactionist formulations of personality on empirical evidence. Interestingly, Bowers (1973) arrived at his defence of the interactionist view by arguing against the situationist position whereas Mischel (1973) and others (e.g. Argyle and Little, 1972) came to the same conclusion (i.e. the defence of interactionism) from a critical examination of the personologist view.

The conceptualisations of the classical theories of, for example, Kantor's early 'interbehavioural psychology', Lewin's 'field theory' and Murray's 'personology' are essentially the same
as the main ideas of the recent interactionist formulations by Argyle and Little (1972); Bowers (1973); Endler and Magnusson (1976) and Mischel (1973). They all have in common the view that human behaviour has to be seen as a function of the interaction between person and environment. In spite of the conceptual similarity between early and recent interactionists' thoughts, the latter seem to have been developed almost independently of the former (Ekehammar, 1974). In fact, Hogan et al (1977) rather disparagingly remarked that:

"It is also interesting to watch the discipline of personality research recycle itself. Thus Bem and Allen (1974) return to Allport's (1937) thesis that the variables of greatest importance in the study of personality are those that are most important to the subject. Similarly, Bowers (1973) reformulates Murray's (1938) view that overt behaviour is a function of the interaction between traits (i.e., psychogenic needs) and situational contingencies (i.e., environmental press). Mischel (1977) also seems to return to a Murray theme. His statement that people are "so complex and multifaceted as to defy easy classifications and comparisons on any single or simple common dimension" (p.253) sounds remarkably like Murray warning us about the overwhelming complexity of the human personality. This warning is puzzling as well as interesting, since as a historian friend once remarked, anyone who has thought for long about human affairs knows that they are complex; the question is what to do next." (p.262)

However, in a recent review Buss (1977) examined several prominent approaches to interactionalism and concluded that two radically different meanings of the term are often emphasised - the mechanistic paradigm and the organismic or system paradigm. Overton and Reese (1973) made a similar distinction between the mechanistic or reactive organism and the organismic or dynamic model.
The mechanistic model being concerned with unidirectional causality assumes that independent variables affect dependent variables. The organismic model is concerned with reciprocal interaction between environmental events and behaviour. The distinction is all important if a genuine, creative system-oriented paradigm is to emerge as a dilectical synthesis of personologism-situationism clash. Essentially the resulting perspective should depict an active organism transacting with an active environment (e.g. Mischel, 1977; Riegel, 1978).

As Bandura (1977) has pointed out interaction can be conceptualised in different ways reflecting alternative views of how causal processes operate. In the unidirectional notion of interaction, persons and situations are treated as independent entities that combine to produce behaviour. This approach is usually represented as $B = f(P,E)$ where $B$ signifies behaviour, $P$ the Person, and $E$ the environment. The validity of this commonly held view is questionable on several grounds, but particularly because personal and environmental factors do not function as independent determinants, rather they determine each other. Nor can 'persons' be considered causes independent of their behaviour. It is largely through their actions that people produce the environmental conditions that affect their behaviour in a reciprocal fashion. The experiences generated by behaviour also partly determine what a person becomes and can do which, in turn, affects subsequent behaviour.

A second conception of interaction acknowledges that personal and environmental influences are bidirectional, but retains a unidirectional view of behaviour. In this analysis, persons and situations are depicted as interdependent causes of behaviour as though it were only a product that does not figure at all in the
causal process \( B = f(P + E) \). As we have already seen, behaviour is an interacting determinant, not simply an outcome of a 'person-situation interaction'.

The Main Features of Modern Interactionism

Modern interactionism rejects unidirectional causality represented in the two model above. According to Endler and Magnussen (1976) the essential four features of modern interactionism can be summarised as follows:

1. Actual behaviour is a function of a continuous process of multidirectional interaction (feedback) between the individual and the situation that he or she encounters.

2. The individual is an intentional, active agent in this interaction process.

3. On the person side of the interaction, cognitive factors are the essential determinants of behaviour, although emotional factors also play a role.

4. On the situation side, the psychological meaning of the situation for the individual is the important determining factor.

1. Reciprocal Causality

A continuous dynamic process of interaction in which behaviour, other personal factors, and environmental factors all operate as
interlocking determinants of each other (B \rightarrow E) is the hallmark of the modern interactionist position. The relative influences exerted by these interdependent factors differ in various settings and for different behaviours. There are times when environmental factors exercise powerful constraints on behaviour, and other times when personal factors are the overriding regulators of the course of environmental events. Thus modern interactionism rests on a model of reciprocal causality. The distinction between the reciprocal causality model of interactionism and the reactive (Overton and Reese, 1973) and mechanistic (Buss, 1977) has been discussed by Pervin (1968) who suggests the term interaction for unidirectional causality and the term transaction for reciprocal causation.

2. Intentional and Active Agents

In the interactional dynamic process, the person is an intentional, active agent. Individuals interpret situations and assign meaning to them. Additionally, as Bandura (1977) has pointed out they activate the environment through their actions. Mischel (1977) reflects this perspective in his conjecture that our subjects are much smarter than many of us thought they were and with respect to psychological assessment he recommended that:

"It would be wise to allow our 'subjects' to slip out of their roles as passive 'asseses' or 'testees' and to enrol them, at least sometimes, as active colleagues who are the best experts on themselves and are eminently qualified to participate in the development of descriptions and predictions - not to mention decisions - about themselves." (p.245)
3. Cognitive Personal Variables

The personal variables that are central to modern interaction focus are primarily cognitive factors, for example selective attention and encoding relevance, rehearsal and storage processes, cognitive transformation and the active construction of cognitions and actions (see, e.g. Bandura, 1977; Mischel, 1973). For example Mischel (1973) argued that the following five cognitive factors and their interactions are important in explaining individual differences:

(a) constructive competencies (the ability to construct or generate particular cognitions and behaviours);

(b) encoding strategies and personal constructs;

(c) behaviour outcome and stimulus outcome expectancies in particular situations;

(d) subjective stimulus values; and

(e) self-regulatory systems and plans, rules and self-reactions for the performance and organisation of complex behaviour sequences. These social cognitive person variables develop ontogenetically in a social learning process on the basis of a given genetic disposition, and interact multidirectionally with situation variables in determining actual behaviour.
4. The Analysis of Environment

On the situation side of the person-situation interaction, the psychological meaning that a situation as a whole has for an individual is seen as the essential determinant of his or her behaviour. The emphasis on the psychological meaning of situations has important consequences for research. Very little empirical research in this area has been conducted. There has however, been a dramatic rise of interest in the environment as it relates to the person and as is true in most new fields, a first concern has been to try to classify them into a taxonomy. Depending on one's purpose, many different classifications are possible and useful (e.g. Magnusson and Ekehammar, 1973; Moos, 1975). But as Mischel (1977) points out

"To seek any single 'basic' taxonomy of situations may be as futile as searching for a final or ultimate taxonomy of traits: We can label situations in at least as many different ways as we can label people. It will be important to avoid emerging simply with a trait psychology of situations, in which events and settings, rather than people, are merely given different labels. The task of naming situations cannot substitute for the job of analysing how conditions and environments interact with the people in them."

(p.250)

In summary - from the perspective of modern dynamic interactionism the person is constructed as an active, aware problem-solver, actively constructing his or her psychological world, and influencing the environment. It views the person as so "complex and multifaceted as to defy easy classifications and comparisons on any single or simple common dimension, as multiply influenced by a host of interacting determinants, as uniquely
organised on the basis of prior experiences and future expectations, and yet as rule-guided in systematic, potentially comprehensible ways that are open to study by the method of science" (Mischel, 1977, p.253).

It is an image that is far removed from the paradigms of personologism and situationalism and its implications for psychodiagnostic assessment are substantial.

**Interactional Assessment**

Although much recent thinking and empirical evidence confirms the importance of studying not just the person and the situation separately but the interaction between them, little more than lip service to the position has occurred in the assessment literature. Textbooks of assessment - including both those of traditional and those of behavioural-orientation - concentrate almost exclusively on person evaluation, though a number of newer ones also include chapters on assessment of environmental variables. The only text that includes a chapter on interactional assessment is Sundberg (1977) which includes a chapter titled 'Assessment of Person in Contexts' (Chapter 5, pp.110-132).

The essence of interactional assessment is the systematic attempt to assess the individual-situation transaction. This is not simply the assessment of individuals and the assessment of situations, but rather the assessment of the two in an integrated fashion for a given individual, and to do so in a systematic, objective manner. In other words, to assay, "... for a particular person in a given situation, both the relevant intrapersonal and situational determinants for that person in that environment" (McReynolds, 1979, p.239). Although
the reciprocal sources of influence are separated for assessment purposes, in everyday life the two-way control operates concurrently and this should be recognised by the assessor.

Within the realm of systematic assessment, Murray's (1938) early work on needs and press is a rare example of assessment within the interactional orientation. In his theoretical formulation a press is anything outside a person that can do something to or for that person. Both press and need have organising and directional qualities. Murray (1938) clarified the concept press as follows:

"... a directional tendency in an object or situation. Like a need, each press has a qualitative aspect - the kind of effect which it has or might have upon the subject ... as well as a quantitative aspect, since its power for harming or of benefiting varies widely. Everything that can supposedly harm or benefit the well-being of an organism may be considered pressive, everything else inert." (pp.118-119)

Murray distinguished between two kinds of press - the alpha press that is the environmental force which objectively exists, as far as scientific inquiry can determine it, and the beta press that is the subject's own interpretation of the phenomena that he or she perceives. For instance, walking along a strange street at night, I might think a dark object ahead is a crouching person (beta press), but on getting closer I see it is only a shrub (now my beta press coincides with the alpha press).

In assessment, one also must relate personal needs to press. Murray, of course, used stories elicited by pictures on his Thematic Appreception Test to discern both needs and press. TAT analysts look for need-press patterns; the conjunction of need and press is
called a *thema*. It is doubtful, however, that many clinicians actually score the TAT for press, and in any event the use of this procedure does not involve the direct assessment of the subject's environment (i.e. alpha press).

More recently George Stern (1970) has carried Murray's ideas and the analysis of press further than anyone else in the field of assessment. Stern developed several instruments to measure need and press variables in the college context. Student needs are assessed by the Activities Index and characteristics of college-environments are evaluated by several Environmental Indexes. This project represents an example of the systematic assessment of the person combined with systematic but separate assessment of the relevant environment(s). This method assumes that relevant data could be obtained on the prospective target environment, and further, that these data could be combined to yield a specific prediction.

A more recent application of this methodology is to be found in the work of Bem and Funder (1978). These authors emphasise the need for the development of a common language for describing and measuring both the person and the target environment, and they propose a "template-matching" technique to accomplish this end. The essence of this proposal is that a situation can be described in terms of a set of templates, each of which is "a personality description of an idealised 'type' of person expected to behave in a specified way in that setting" (p.486). The assessor can then directly compare the personality of the subject with each of the situation templates in order to evaluate the likelihood of the subject behaving in that way.
in the given situation. Bem and Funder report three illustrative studies using Q-sort techniques to develop both the situation templates to indicate the probability of given behaviours occurring.

An approach that by-passes the problem of combining data and attempts to record the person-environment transaction is the direct behaviour sampling of persons-in-situations. Many sophisticated examples of this approach are to be found in the chapter on 'Direct Measurement in Natural Settings' in Cone and Hawkins (1977) and other equivalent handbooks (e.g. Ciminero et al, 1977; Hersen and Bellack, 1976; Mash and Terdal, 1976) and will not be outlined here.

Little more can be said at this stage since, despite Ekehammar's (1974) assertion that "if interactionism is not the zeitgeist of today's personality psychology, it will probably be that of tomorrow's" (p.1045), very little attempt has been made to explore the implication of the paradigm to psychodiagnostic assessment. Much more research is required at both the theoretical and technical levels. With respect to the former McReynolds (1979) notes that the interactionist formula is not, as such, a usable formula, but merely a point of view that will need a great deal of elaboration before being directly applicable in research and practices. At the technical level, McReynolds suggests that most of the available assessment techniques are "more suitable for research than for dealing with individual clients, hence, there is an excellent opportunity for innovativeness in the development of practicable, usable interactional assessment packages" (1979, p.245).

In the light of McReynolds (1979) conclusion, it appears as if the practitioner will not be significantly affected by the rising paradigm for many years. It was pointed out in the introduction that,
according to Kanfer (1977), we are currently witnessing the synthesis stage of the dialectic between personologism and situationism. However, it was also pointed out that this 'revolution' has hardly touched the practicing clinician. And it is unlikely to, until the technological by-products of the new paradigm are developed and taught to trainee clinicians. Thus, the onus is squarely on researchers and instructors to develop reliable and valid processes with demonstrable utility and to socialise a new generation of clinical psychologists or counsellors in their use.
ABRAMOWITZ, S.I., ABRAMOWITZ, C.V., JACKSON, C., & GOMES, B.

ABRAMOWITZ, C.V. & DOKIECKI, P.R.

ADAMS, H.E., DOSTER, J.A. & CALHOUN, K.S.
A Psychological Based System of Response Classification. In A.R. Ciminero et al, 1977, op cit

ADINOLFI, A.

ALBEE, G.W.

ALEXANDER, F.
Psychoanalysis and Medicine. Mental Hygiene, 1932, 16, 63-84

ALEXANDER, F. & SELESNICK, S.T.
Freud-Bleuler Correspondence. Archives of General Psychiatry, 1965, 12, 1-9

ALKER, H.A.
Is a personality situationally specific or intrapsychically consistent? Journal of Personality, 1972, 40, 1-16

ALLPORT, G.W.

ALLPORT, G.W.
Traits revisited. American Psychologist, 1966, 21, 1-10

ANASTASI, A.

ANASTASI, A.

ANDRESKI, S.

ARGYLE, M. & LITTLE, B.R.
Do personality traits apply to social behaviour? Journal of Personality, 1972, 40, 1-16

ARTHUR, A.Z.
AYER, A.J.  

BAILEY, P.  

BAKAN, D.  

BAKAN, D.  

BAKKER, C.B.  
Why people don't change. Psychotherapy: Theory, Research and Practice, 1975, 12, 164-172

BANDURA, A.  

BANESH, H.  

BARATZ, J. & BARATZ, S.  

BARBER, B.  
Resistance by Scientists to Scientific Discovery. Science, 1961, 134, 596-602

BARBER, T.X.  

BARITZ, L.  

BARLOW, D.H.  

BARNES, B.  

BARNES, B.  
BARTLEY, W.W. 


BECKER, H.S. 

BEGELMAN, D.A. 
Behavioural classification. In M. Hersen & A.S. Bellack op cit, 1976

BEM, D.J. 

BEM, D.J. & ALLEN, A. 
On predicting some of the people some of the time: The search for cross-situational consistencies in behaviour. Psychological Review, 1974, 81, 506-520

BEM, D.J. & FUNDER, D.C. 
Predicting more of the people more of the time: Assessing the personality of situations. Psychological Review, 1978, 85, 485-501

BENASSI, V., & LANSON, R. 
Survey of the teaching of behaviour modification in colleges and universities. American Psychologist, 1927, 21, 1063-1069

BENEDICT, R. 
Anthropology and the abnormal. Journal of General Psychology, 1934, 10, 59-80

BERG, I.E. 

BERSOFF, D.N. 

BERSOFF, D.N. 
Silk purses into sow's ears: the decline of psychological testing and a suggestion for its redemption. American Psychologist, 1973, 28, 892-899

BIERI, J. 
BIRLEY, J.L.T.
Coercion and care, cited in A. Clare, Psychiatry in Dissent.
London: Tavistock, 1976

BJORKMAN, M.

BLACK, H.

BLACKBURN, R. (ed)

BLASHFIELD, R.K. & DRAGUNS, J.G.
Evaluative criteria for psychiatric classification. Journal of Abnormal Psychology, 1976, 85, 140-50

BLATT, S.J.
The validity of projective techniques and their clinical and research contributions. Journal of Personality Assessment, 1975, 39, 327-43

BLOCK, J.
Advancing the psychology of personality: Paradigmatic shift or improving the quality of research. In D. Magnusson and N.S. Endler (eds), Personality at the crossroads: Current issues in interactional psychology. Hillsdale, New Jersey: Erlbaum, 1977

BLUM, J.D.

BLUM, J.M.

BORDUA, D.J.

BORING, E.G.

BOWERS, K.S.

BRAGINSKY, D.D.
BRAGINSKY, D.D. & BRAGINSKY, B.M.

BRAGINSKY, D.D. & BRAGINSKY, B.M.
Psychologists: High Priests of the Middle Class. Psychology Today, 1973, 7, 15-18

BRAGINSKY, D.D. & BRAGINSKY, B.M.

BRAGINSKY, B.M., BRAGINSKY, D.D. & RING, K.

BREGER, L.

BRIELAND, D.

BRILL, A.
Lectures on Psychoanalytic Psychiatry. London: Methuen, 1948

BRODY, B.
The Conventional Elite. Psychotherapy and Social Science Review, 1971, 5, 22-28

BROWN, W.R., & McGUIRE, J.M.
Current assessment practices. Professional Psychology, 1976, 7, 475-84

BRUNSWIK, E.
The conceptual framework of psychology. Chicago: Chicago University Press, 1952

BRUNSWIK, E.
Perception and the representative design of psychological experiments. Berkeley: University of California Press, 1956

BRY, I. & RIFKIN, A.H.

BUROS, O.I. (ed)

BURT, C.
Ability and Income. British Journal of Educational Psychology, 1943, 13, 83-98
BURT, C.

BURT, C.

BUSS, A.H.

BUSS, A.R.
The emerging field of the sociology of psychological knowledge. American Psychologist, 1975, 30, 988-1002

BUSS, A.R.

CAMPBELL, C.H.
Induced Delusions: The Psychopathy of Freudism. Chicago: Regent House, 1957

CAREY, A.
The Place of the Social Sciences in Contemporary Culture. Unpublished mimeographed paper, University of New South Wales, 1972

CAREY, A.
The Lysenko Syndrome in Western Social Science. Australian Psychologist, 1977, 12, 27-38

CARNAP, R.

CATTELL, J. McK.
Mental tests and measurements. Mind, 1890, 15, 373-381

CATTELL, R.B.
Description and measurement of personality. New York: World Book, 1946

CATTELL, R.B.

CAUTHEN, N.R., ROBINSON, I.E. & KRAUSS, H.H.

CHALMERS, A.F.
What is this thing called science?: an assessment of the nature and status of science and its methods. St. Lucia: University of Queensland Press, 1976
CHAPMAN, L.  

CHAPMAN, L. & CHAPMAN, J.  
Illusory correlation as an obstacle to the use of valid psycho-diagnostic signs. Journal of Abnormal Psychology, 1969, 74, 271-280

CHAPMAN, L.J. & CHAPMAN, J.P.  

CHAPMAN, L. & CHAPMAN, J.  
Genesis of popular but erroneous psycho-diagnostic observation. Journal of Abnormal Psychology, 1967, 72, 193-204

CHASEN, B., & WEINBERG, S.L.  
Diagnostic sex-role bias: How can we measure it? Journal of Personality Assessment, 1975, 39, 620-29

CHESLER, P.  

CHOMSKY, N.  
IQ tests. Building blocks for the new class system. Ramparts, 1972, 11, 24-30

CHU, F.D. & TROTTER, S.  

CIMINERO, A.R., CALHOUN, K.S. & ADAMS, H.E.  

CIOFFI, F.  

CIOFFI, F.  
"Was Freud a Liar?" The Listener, 1974, 91, 65-89

CLARE, A.  
Psychiatry in Dissent. London: Tavistock, 1976

CLEARY, T.A., HUMPHREYS, L.G., KENDRICK, S.A. & WESMAN, A.  
Educational Uses of Tests with Disadvantaged Students. American Psychologist, 1975, 30, 15-41

CLEVELAND, S.E.  
Reflections on the rise and fall of psychodiagnosis. Professional Psychology, 1976, 7, 309-18
CLINE, V.B.
Ability to judge personality assessed with a stress interview and sound-film technique. *Journal of Abnormal and Social Psychology*, 1955, 50, 183-187

CLINE, V.B.

CLINE, V.B. & RICHARDS, J.M.
The accuracy of interpersonal perception: A general trait? *Journal of Abnormal and Social Psychology*, 1960, 60, 1-7

COCH, L. & FRENCH, J.R.P.
Overcoming resistance to change. *Human Relations*, 1948, 1, 512-32

COLES, J.K. & MAGNUSSEN, M.G.
Where the action is. *Journal of Consulting Psychology*, 1966, 30, 539-543

CONRAD, P.

COUCH, A.S. & KENISTON, K.

COWEN, E.L.

CRAIK, K.H.

CROMWELL, R.L., BLASHFIELD, R.K. & STRAUSS, J.S.

CRONBACH, L.J.
CRONBACH, L.J. & GLESER, G.C.

DAILEY, C.A.
The effects of premature conclusions upon the acquisition of understanding of a person. Journal of Psychology, 1952, 33, 133-152

DANA, R.M.

DANA, R.H. & LEECH, S.

D'ANDRADE, R.G.

D'ANDRADE, R.G.

DAVENPORT, C.B.
Heredity in relation to eugenics. New York: Holt, 1911

DAVID, F.N.
Games, gods and gambling. London: Charles Griffin, 1962

DeMONBREUN, B.G. & MAHONEY, M.S.
The Effects of Data Return Patterns on Confidence in an Hypothesis. In M. Mahoney, Scientist as Subject: The Psychological Imperative. Cambridge Mass: Ballinger Publishing Co., 1976

DENGROVE, E.

DICKSON, C.R.

DI NARDO, P.A.

DIVOKY, D. & SCHRAG, P.
DOBZHANSKY, T.

DOBZHANSKY, T.
Differences are not deficits. Psychology Today, 1973, 7, 96-102

DOHRENWEND, B.P. & DOHRENWEND, B.S.

DOHRENWEND, B.P. & DOHRENWEND, B.S.

DOWLING, J.F. & GRAHAM, J.R.
Illusory correlation and the MMPI. Journal of Personality Assessment, 1976, 40, 531-38

DRAGUNS, J.G.
Values reflected in psychopathology: the case of the Protestant ethic. Ethos, 1974, 2, 115-136

DRAGUNS, J.G. & PHILLIPS, L.
Psychiatric Classification and Diagnosis: An Overview and Critique. New Jersey: General Learning, 1971

EDWARDS, A.D. & HARGREAVES, D.H.

EDWARDS, A.L. & WALSH, J.A.

EELLS, K., DAVIS, A., HAVINGHURST, R., HERRICK, V., & TYLER, R.
Intelligence and cultural differences. Chicago: Uni of Chicago, 1951

EFRON, C.

EINHORN, H.J. & HOGARTH, R.M.

EINHORN, H.J. & SCHACHT, S.

EKEHAMMAR, B.
Interactionism in personality from a historical perspective. Psychological Bulletin, 1974, 81, 1026-1048
ELKIN, D.
56-59

ELLENBERGER, H.F.
The Discovery of the Unconscious. New York, Basic Books, 1970

ENDLER, N.S. & HUNT, M. McV.

ENDLER, N.S. & HUNT, J. McV
Generalizability of contributions from sources of variance in the S-R inventories of anxiousness. Journal of Personality, 1969, 37, 1-24

ENDLER, N.S. & MAGNUSSON, D.
Towards an interactional psychology of personality. Psychological Bulletin, 1976, 83, 956-74

ENGEL, G.L.
Some obstacles to the development of research in psychoanalysis. Journal of American Psychoanalytic Association, 1968, 16, 195-229

EPSTEIN, S.

ESCALONA, S.

ESTES, W.K.
The cognitive side of probability learning. Psychological Review, 1976, 83, 37-64 (a)

ESTES, W.K.

EVANS, I.M. & NELSON, R.D.
A curriculum for the teaching of behaviour assessment. American Psychologist, 1974, 29, 598-606

EVANS, I.M. & NELSON, R.D.

EYSENCK, H.J.
EYSENCK, H.J.
The IQ Argument: Race Intelligence and Education. New York: Library Press, 1971

EYSENCK, H.J.

EYSENCK, H.J.
Behaviour Therapy - Dogma or Applied Science? In M.P. Feldman & A. Broadhurst (eds), Theoretical and Experimental bases of the behaviour therapies. London: John Wiley, 1976

EYSENCK, H.J.
The Case of Sir Cyril Burt. Encounter, 1977, 48, 19-24

EYSENCK, H.J. & WILSON, G.D.
The Experimental Study of Freudian Theories. London: Methuen, 1973

FARINA, P. HOLLAND, C.H., & RING, K.
Role of stigma and set in interpersonal interaction. Journal of Abnormal Psychology, 1966, 71, 421-428

FARREL, B.A.
Can psychoanalysis be refuted? Inquiry, 1964, 4, 16-36

FARSON, R.

FENICHEL, O.

FERNALD, N.E.
The burden of feeblemindedness. Journal of Psychoaesthenics, 1912, 17, 87-111

FERSTER, C.B.

FEYERABEND, P.K.

FEYERABEND, P.K.

FILER, R.N.
The clinician's personality and his case reports. American Psychologist, 1952, 7, 336
FISCHER, C.T.
Intelligence defined as effectiveness of approaches. Journal of Consulting Clinical Psychology, 1969, 33, 668-674

FISCHER, C.T.

FISH, J.M.

FISHER, S. & GREENBERG, R.P.

FISKE, D.W.
The limits for the conventional science of personality. Journal of Personality, 1974, 42, 1-11

FLOWERMAN, S.H.

FORD, D.M. & URBAN, H.B.

FOUCAULT, M.

FRANK, G.

FRANK, G.

FRANK, J.D.

FRANK, J.D.

FRANK, L.K.
Projective methods for the study of personality. Journal of Psychology, 1939, 8, 389-413

FRANKS, C.M.

FRANKS, C.M.
Forward. In E.J. Mash and L.G. Terdal, op cit, 1976
FREUD, E.L. (ed)  

FREUD, S.  

FREUD, S. (1910)  
A special type of choice of object made by men.  Standard Edition  
Vol. XI.  London: Hogarth, 1957

FREUD, S. (1916-1917)  
Introductory lectures on psycho-analysis.  Standard Edition  
Vols. XV-XVI.  London: Hogarth, 1963

FREUD, S. (1926)  

GALTON, F.  
Hereditary Genius (ed. 2).  London: MacMillan, 1892

GALTON, F.  
Inquiries Into Human Faculty (ed. 2).  London: Everyman, 1907

GARDNER, M.  
Fads and fallacies in the name of Science.  New York: Dover, 1957

GARFIELD, S.L. & KURTZ, R.M.  

GARFIELD, S.L. & KURTZ, R.M.  
A survey of clinical psychologists: characteristics, activities, and orientations.  The Clinical Psychologist, 1974, 28, 7-10

GATHERCOLE, C.E.  

GAURON, E.F. & DICKINSON, J.K.  
Diagnostic decision making in psychiatry.  Archives of General Psychiatry, 1966, 14, 225-232

GEERTZ, C.  

GIBBY, R.G.  

GILBERT, J.  

GILLIE, O.  
The life and deceptive times of the man who "invented" the eleven-plus: Sir Cyril Burt and the great IQ fraud.  New Statesman, 1978, 96, 688-694
GIOGRI, A.

GLOVER, E.

GODDARD, H.H.
The Binet Tests in Relation to Immigration. Journal of Psycho-asthenics, 1913, 18, 105-07

GODDARD, H.H.
The Kallikak Family. New York: MacMillan, 1912

GODDARD, H.H.

GODDARD, H.H.

GOLDFARBER, A.
Reliability of diagnostic judgements by psychologists. Journal of Clinical Psychology, 1959, 15, 392-396

GOLDFRIED, M.R.

GOLDFRIED, M.R. & SPRAFKIN, J.N.

GOLDFRIED, M.R. & LINEHAN, M.M.
Basic Issues in Behavioural Assessment. In M. Hersen & A.S. Bellack, op cit, 1976

GOLDING, S.L.
Flies in the ointment: Methodological problems in the analysis of the percentage of variance due to persons and situations. Psychological Bulletin, 1975, 82, 278-288

GOLDING, S.L. & ROVER, L.G.
Illusory correlation and subjective judgement. Journal of Abnormal Psychology, 1972, 80, 249-60
GOLLOB, M.F. ROSSMAN, B.B. & ABELSON, R.P.

GOODENOUGH, F.I.
Mental Testing. New York: Rinehart, 1949

GOUGH, H.G.

GOUGH, H.G.

GOVE, W.R.

GREENWALD, A.G.
Consequences of prejudice against the null hypothesis. Psychological Bulletin, 1975, 82, 1-20

GRIGG, A.E.
Experience of clinicians and speech characteristics and statements of clients as variables in clinical judgements. Journal of Consulting Psychology, 1958, 22, 315-319

GRINSTEIN, A. (ed)

GROLLMAN, E.A.

GROSZ, H.J. & GROSSMAN, K.G.

GULLIKSEN, H.

HABERMAS, J.
Knowledge and Human Interests. New York: Beacon Press, 1971

HALDANE, J.B.S.
Heredity and Politics. New York: W.W. Norton & Co., 1938

Haley, J. (ed)
The power tactics of Jesus Christ and other essays. New York: Grossman, 1969
HALLECK, S.L.

HALLER, M.

HAMMOND, K.R.

HAMMOND, K.R., WILKINS, M.M. & TODD, F.J.
A research paradigm for the study of interpersonal learning. Psychological Bulletin, 1966, 65, 221-232

HANSON, N.R.
Patterns of discovery: an inquiry into the conceptual foundations of science. Cambridge: Cambridge University Press, 1958

HARRIS, S., & MASLING, J.

HARTMANN, H.

HARTMANN, H.

HAYAKAWA, S.I.

HAYNES, S.N.

HAYNES, S.N.

HEBB, D.O.
What psychology is about. American Psychologist, 1974, 34, 71-79

HEIDER, F.
HENDERSON, P.

HENLEY, N.M.

HERRNSTEIN, R.J.
IQ. Atlantic Monthly, 1971, Sep, 43-64

HERSEN, M. & BARLOW, D.H.

HERSEN, M. & BELLACK, A.S. (eds)

HERSH, J.B.

HOBBS, N.
The Futures of Children. San Francisco: Jossey-Bass, 1975

HOBBS, N. (ed)
Issues in the Classification of Children. San Francisco: Jossey-Bass, 1975a

HOGAN, J.C.

HOGAN, R., DESOTTO, C. & SOLAND C.

HOLLAND, J.L.

HOLLAND, J.L. & RICHARDS, J.M. JR.
Academic and non-academic accomplishment: Correlated or uncorrelated? Journal of Educational Psychology, 1965, 56, 165-174

HOLLINGSHEAD, A.D. & REDLICH, F.C.
Social class and mental illness. New York: Wiley, 1958

HOLMES, D.S.

HOLMES, G.
Introduction to clinical neurology. Edinburgh, Livingstone, 1946
HOLT, R.R.

HOLZBERG, J.D.
The historical traditions of the state hospital as a force of resistance to the team. American Journal of Orthopsychiatry, 1960, 30, 87-94

HOVLAND, C.I. & WEISS, W.
Transmission of information concerning concepts through positive and negative instance. Journal of Experimental Psychology, 1953, 45, 175-182

HOWARD, J.A.

HUDSON, L.
Degree class and attainment in scientific research. British Journal of Psychology, 1960, 51, 67-73

HULL, C.L.
Aptitude testing. Yonkers: World Book, 1928

HULL, C.L.

HUNT, J. McV.
Intelligence and Experience. New York: Roland Press, 1961

HUXLEY, A.
Is Psychanalysis a Science? The Forum, 1925, 73, 316-17

IVNIK, R.J.
Clinical Psychology. Professional Psychology, 1977, 8, 206-213

JACKSON, G.D.
On the report of the ad hoc committee on educational uses of tests with disadvantaged students: Another psychological view from the Association of Black Psychologists. American Psychologist, 1975, 30, 88-93

JACOBY, R.

JAHODA, M.

JAMES, W.
The Principles of Psychology. New York: Holt, 1890

JASTROW, J.
The House that Freud Built. New York: Greenberg, 1932
JENKINS, H. & WARD, W.

JENSEN, A.R.
How Much Can We Boost IQ and Scholastic Achievement? Harvard Educational Review, 1969, 39, 1-123

JENSEN, A.R.
How Much Can We Boost IQ and Scholastic Achievement? In Genetics and Education (ed), Jensen. London: Methuen, 1972

JENSEN, A.R.
Kinship correlations reported by Sir Cyril Burt. Behavior Genetics, 1974, 4, 76-83

JENSEN, A.R.

JOHNSON, H.K.
Psychoanalysis - a critique. Psychiatric Quarterly, 1948, 22, 321-338

JOHNSON, H.K.
Psychoanalysis: Some critical comments. American Journal of Psychiatry, 1956, 113, 466-467

JOHNSON-LAIRD, P.N. & WASON, P.C.

JONES, E.E.

JONES, E.E.

JONES, E.E.
The Rocky Road from Acts to Dispositions. American Psychologist, 1979, 34, 107-117

JONES, R.A.
Labels and stigma in Special Education. Exceptional Children, 1972, 15, 553-564

JONES, R.A.

JONES, R.R., REID, J.B. & PATTERSON, G.R.
Naturalistic observation in clinical assessment. In P. McReynolds (Vol. 3), op cit, 1975a
JURJEVICH, R.M.
   The Hoax of Freudism: A Study of Brainwashing the American
   Professions and Layman. Philadelphia: Dorrance & Company,
   1974

KADUSHIN, C.

KAHNEMAN, D. & TVERSKY, A
   Subjective probability - a judgement of representativeness.
   Cognitive Psychology, 1972, 3, 430-54

KAHNEMAN, D. & TVERSKY, A
   On the psychology of prediction. Psychological Review, 1973, 80,
   237-57

KAMIN, L.J.
   The science and politics of IQ. New York: John Wiley, 1974

KANFER, F.
   Personal control, social control, and altruism. American
   Psychologist, 1979, 34, 231-239

KANFER, F.H. & SASLOW, G.
   Behavioural diagnosis. In C.M. Franks (ed), Behaviour therapy:

KANTOR, J.R.
   1924

KANTOR, J.R.
   The Aim and Progress of Psychology and Other Sciences. Chicago:
   Principia Press, 1971

KAPLAN, A.
   The Conduct of Inquiry. San Francisco: Chandler, 1974

KARIER, C.J.
   Testing for order and control in the corporate liberal state.
   Educational Theory, 1972, 22, 154-80

KARPMAN, B.
   The Sexual Offender and His Offenses: Etiology, Pathology,

KAZDIN, A.E. & WILSON, G.T.
   Evaluation of behaviour therapy: Issues, evidence and research

KAZIN, A.
   The Freudian Revolution Analysed. In B. Nelson, Freud and
   the Twentieth Century. London, 1958

KEEN, S.
   Psychological establishment - the Freudian Mafia. Psychology
   Today, 1972, 5, 44
KELLY, E.L.

KELLY, G.

KELLEY, H.H.

KELLEY, T.L.

KESSEL, N. & SHEPHERD, M.

KITTRIE, N.N.
The right to be different: deviance and enforced therapy. Baltimore: Penguin, 1971

KOCH, E.
Psychological Science versus the science-humanism antimony: Intimation of a significant science of man. American Psychologist, 1961, 16, 629-639

KOFFKA, R.

KOSTLAN, A.

KOSTLAN, A.
A reply to Patterson. Journal of Consulting Psychology, 1955, 19, 486

KRANTZ, D.L.
The separate worlds of operant and non-operant psychologists. Journal of Applied Behaviour Analysis, 1971, 4, 61-70

KRATCHEWILL, T. (ed)
KUBIE, L.S.
Psychoanalysis and Scientific Method. Journal of Nervous and Mental Disease, 1960, 131, 495-512

KUHN, T.S.

KUHN, T.S.

KURIANSKY, J.B., DEMING, E. & GURLAND, B.J.

KURTZ, R.M. & GARFIELD, S.L.

LAING, R.D.

LAING, R.D. & ESTERSON, A
Sanity, Madness and the Family. London: Tavistock, 1964

LAKATOS, I.

LAKATOS, I. & MUSGRAVE, A. (eds)

LANGER, E.J. & ABELSON, R.P.
A patient by any other name ... clinician group difference in labelling bias. Journal of Consulting Clinical Psychology, 1974, 42, 4-9

La PIERRE, R.T.
The Freudian Ethic. Des Moines, Iowa: Duell, Sloan and Piece, 1959

LAWLER, J.M.
IQ, Heritability and Racism. London: Lawrence & Wishart, 1978

LAYZER, D.
Science or Superstition? (A Physical Scientist Looks at the IQ Controversy). In C. Senna (ed), The Fallacy of IQ. New York: Third Press, 1973

LEARY, T.
Interpersonal diagnosis of personality. New York: Ronald Press, 1957
LEE, S.

LEHMANN, H.C. & WITTY, P.A.
Faculty psychology and personality traits. American Journal of Psychology, 1934, 44, 486-500

LEIFER, R.
The Psychiatrist and tests of criminal responsibility. American Psychologist, 1964, 19, 825-830

LEIFER, R.
In the name of mental health. New York: Science House, 1969

LEIFER, R.

LEVI-STRAUSS, C.

LEVITT, E.E.

LEVY, M.R. & FOX, H.M.
Psychological testing is alive and well. Professional Psychology, 1975, 6, 420-34

LEWANDOWSKI, D.C. & SACCUZZO, D.P.
The Decline of Psychological Testing. Professional Psychology, 1976, 7, 177-184

LEWIN, K., LIPPITT, R. & WHITE, R.
Patterns of aggressive behaviour in experimentally-created 'social climate'. Journal of Sociology, 1939, 10, 271-299

LEWIS, K.

LEWIS, K.

LEWIS, M. & FREEDLE, R.

LEY, P.
The reliability of psychiatric diagnosis: some new thoughts. British Journal of Psychiatry, 1972, 121, 41-43
LICK, J.
Statistical Vs. clinical significance in research on the outcome of psychotherapy. International Journal of Mental Health, 1973, 22, 26-37

LINDNER, R.M.

LINDNER, R.M.
Rebel without a cause - the hypno-analysis of a criminal psychopath. New York: Grune and Stratton, 1954

LINDER, R.
Diagnosis: Description or prescription. Perceptual and Motor Skills, 1965, 20, 1081-1092

LIPPMANN, W.

LITTLE, K.B. & SCHNEIDMAN, E.S.
Congruencies among interpretations of psychological test and anamnestic data. Psychological Monographs, 1959, 73 (6, Whole No. 476)

LIVINGSTON, S.A.
Psychometric techniques for criterion - referenced testing and behaviour assessment. In J.D. Cone & R.P. Hawkins (eds) op cit, 1977

LORR, M. & MCNAIR, D.M.

LUBIN, R., WALLIS, R.R. & PAINE, C.

LUFT, J.
Implicit hypotheses and clinical predictions. Journal of Abnormal and Social Psychology, 1950, 45, 756-759

LYON, D. & SLOVIC, P.

MACCORQUODALE, K. & MEEHL, P.E.
On a distinction between hypothetical constructs and intervening variables. Psychological Review, 1948, 55, 95-107

MacDOUGALL, C.D.
Hoaxes. New York: Dover, 1968

MACKINNON, D.W. & DUCKES, W.F.
McClelland, D.C.

McCully, R.S.

McLemore, C.W. & Benjamin, L.S.
Whatever happened to interpersonal diagnosis? A psychosocial alternative to DSM-III. American Psychologist, 1977, 34, 17-34

McReynolds, P.

McReynolds, P.

McReynolds, P.

McReynolds, P.

McReynolds, P.

McReynolds, P.
The case for Interactional Assessment. Behavioural Assessment, 1979, 1, 237-247

Magee, B.

Magnusson, D.
An analysis of situational dimensions. Perceptual and Motor Skills, 1971, 32, 851-867

Magnusson, D. & Ekehammar, B.

Mahoney, M.J.
MAHONEY, M.J.

MAHONEY, M.J.

MAHONEY, M.J. & KIMPER, T.P.

MALLESON, A.
Need your Doctor be so Useless. London: Allen & Unwin, 1973

MALONEY, M.P., & WARD, M.P.

MANNHEIM, K.

MARCUSE, H.

MARKS, R.

MARMOR, J.

MARMOR, J.

MARTHE, R.

MARTIN, A.R.

MARTIN, B.
MARUYAMA, M.

MASH, E.J. & TERDAL, L.G.
Behaviour Therapy assessment: Diagnosis, design and evaluation. Psychological Reports, 1974, 35, 587-601

MASLING, J.
The effects of warm and cold interaction on the administration and scoring of an intelligence test. Journal of Consulting Psychology, 1959, 23, 336-341

MASLING, J.
The influence of situational and interpersonal variables in projective testing. Psychological Bulletin, 1960, 57, 65-85

MASLING, J. & HARRIS, S.

MASSERMAN, J.

MASSERMAN, J.H.

MATZA, D.

MAY, R.

MEDAWAR, P.B.

MEDAWAR, P.B.

MEEHL, P.E.

MEEHL, P.E.
MEEHL, P.E.

MEEHL, P.E.

MEEHL, P.E.
Theory-testing in psychology and physics: A methodological paradox. Philosophy of Science, 1967, 34, 103-115

MERCER, J.R.

MERCER, J.R.
Sociocultural factors in educational labeling. In M. Begab and S. Richardson (eds), The mentally retarded and society's: A social science perspective. Baltimore: University Park Press, 1975

MERCER, J.R.
Test "Validity", "Bias", and "Fairness": An Analysis from the Perspective of the Sociology of Knowledge. Interchange, 1978-79 9, 1-16

MERCER, J.R. & BROWN, W.C.

MERTON, R.K.
Social Theory and Social Structure (ed. 2). New York: Free Press, 1967

MERTSON, R.K.
Behavior patterns of scientists. American Scientist, 1969, 57 1-23

MICHOTTE, A.
The Perception of Causality. London: Methuen, 1963

MINTZ, J.
Survey of student therapists' attitudes toward psychodiagnostic reports. Journal of Consulting and Clinical Psychology, 1968, 32, 500

MISCHSEL, W.
Personality and Assessment. New York: John Wiley and Sons, Inc. 1968

MISCHSEL, W.
Continuity and change in personality. American Psychologist, 1969, 24, 1012-1018
MISCHEL, W.

MISCHEL, W.
Toward a cognitive social learning reconceptualisation of personality. Psychological Review, 1973, 80, 252-283

MISCHEL, W.

MISCHEL, W.
On the Interface of Cognition and Personality: Beyond the Person-Situation Debate. American Psychologist, 1979, 34, 740-754

MITROFF, I.I.
The subjective side of science. New York: Elsevier, 1974

MITROFF, I.I. & FEATHERINGHAM, T.R.
On systematic problem-solving and the error of the third kind. Behavioural Science, 1974, 19, 383-393

MOORE, G.H., BOBBITT, W.E. & WILDMAN, R.W.

MOOS, R.H.

MOOS, R.H.
Sources of variance in response to questionnaires and in behaviour. Journal of Abnormal Psychology, 1969, 74, 405-412

MOOS, R.H.

MORRISON, J.R.

MOWRER, O.H.

MURRAY, H.A.
Explorations in personality. New York: Oxford University Press, 1938
MYNATT, C.R., DOHERTY, M.E. & TWENEY, R.D.

MYRDUL, G.

NAGEL, E.

NATTALL, R.L. & FOZARD, T.L.
Age, Socioeconomic Status and Human Abilities. Aging and Human Development, 1970, 1, 161-169

NEARY, J.
A scientist's variation on a disturbing racial theme. Life, 1970, 68, 64

NELSON, R.O. & HAYES, S.C.
Some current dimensions of behavioural assessment. Behaviour Assessment, 1979, 1, 1-16

NELSON, R., HAY, L. & HAY, W.
Comments on Cone's "Relevance of Reliability and Validity for Behavioral Assessment". Behavior Therapy, 1977, 8, 427-431

NEULINGER, J.

NEWCOMB, T.M.
The consistency of certain extrovert-introvert behavior patterns in 51 problem boys. Contributions to Education, 1929, 382, 446-455

NISBETT, R.E. & WILSON, T.D.

NYSTEDT, L.
A modified lens model: A study of the interaction between the individual and the ecology. Perceptual and Motor Skills, 1972, 34, 479-498

OBERNDORF, C.P.
O'LEARY, K.D.
Behavioural Assessment. Behavioural Assessment, 1979, 1, 31-36

OLWEUS, D.
A critical analysis of the "modern" interactionist position. In D. Magnusson and N.S. Endler (eds); Personality at the crossroads: Current issues in interactionist psychology. Hillsdale, New Jersey: Erlbaum, 1977

OLWEUS, D.

O'NEIL, W.M.
The Margaret Austin Memorial Lecture: The Study of the Person. Australian Psychologist, 1972, 7, 72-89

ORLANSKY, H.
Infant Care and Personality. Psychological Bulletin, 1949, 46, 1-48

ORNE, M.T.
On the social psychology of the psychological experiment: With particular reference to demand characteristics and their implications. American Psychologist, 1962, 17, 776-783

OSKAMP, S.

OVERTON, W.F.

OVERTON, W.F. & REESE, H.W.

PALMER, J.O.

PARKER, C.
As a clinician thinks. Journal of Counselling Psychology, 1958, 5, 253-262

PARSONS, T.
Illness and the role of the physician: a sociological perspective. American Journal of Orthopsychiatry, 1951, 21, 452-460

PERVIN, L.A.
PETE ISON,  D.R.

PHILLIPS, L., DRAGUNS, F.G. & BARTLETT, D.P.

PIaget, J. & INhelder, B.
The Child's Conception of Space. New York: Humanities Press

PINCKNEY, E.R. & PINCKNEY, Cathey

PLANCK, M.

POLANYI, Michael
Personal knowledge: Towards a post-critical philosophy.
Chicago: Chicago University Press, 1958

POPPER, K.R.

POPPER, K.R.

POPPER, K.R.
Objective knowledge: An evolutionary approach. London: Oxford University Press, 1972

POPPER, K.R.
Unended Quest. Glasgow: Fontana/Collins, 1976

POPLESTONE, J.A. & McPHERSON, M.W.
The prolonged avoidance of intellectual behaviour. Psychological Record, 1974, 24, 549-557

PRICE, R.H.
The taxonomic classification of behaviours and situations and the problem of behaviour-environment congruence. Human Relations, 1974, 27, 567-585

PUNER, H.W.
Freud: His Life and His Mind. New York: Howell, Soskin, 1947

RABKIN, J.G.

RAINEY, V.C. (ed)
RAMZY, I.

RAMZY, I.

RAPAPORT, A.
Psychoanalysis as science. *Bulletin of Menninger Clinic*, 1968, 32, 1-20

RAPAPORT, D.

RAPAPORT, D., GILL, M., & SCHAFER, R.
Diagnostic psychological testing (2 Vols). Chicago: Year Book Publications, 1946

RAPAPORT, D., GILL, M.M. & SCHAFER, R.

RATNER, C.

RAUSH, H.L.

RAUSH, H.L., DITTMANN, A.T. & TAYLOR, T.S.
Person, setting and change in social interaction. *Human Relations*, 1959, 12, 361-378

REGAN, D.T., STRAUSS, E. & FAZCO, R.
Liking and the attribution process. *Journal of Experimental Social Psychology*, 1974, 10, 385-397

REYNOLDS, W.M. & SUNDBERG, N.D.

RICOEUR, P.

RIEGEL, K.

ROETHLESBERGER, F.J. & DICKSON, W.J.

ROGERS, C.R.
ROMMETREIT, R.
Selectivity, intuition and halo effects in social perception.
Oslo: Oslo University Press, 1960

ROSE, M. & ROSE, S.
Science and Society. London: Allen Lane, 1969

ROSE, H. & ROSE, S.
Social responsibility (III): the myth of the neutrality of science. Impact of science and society, 1971, 21, 137-149

ROSE, H. & ROSE, S.

ROSE, H. & ROSE, S.

ROSEN, G.

ROSENHAN, D.L.
On being sane in insane places. Science, 1973, 179, 250-257

ROSENTHAL, R.
On the social psychology of the psychological experiment: the experimenter's hypothesis as an unintended determinant of experimental results. American Scientist, 1963, 57, 268-83

ROSENTHAL, R. & GAITO, J.
The interpretation of levels of significance by psychological researchers. Journal of Psychology, 1963, 55, 33-8

ROSENWALD, G.C.
Psychodiagnostics and its discontents: A contribution to the understanding of professional identity and compromise. Psychiatry, 1963, 26, 222-240

ROSENZWEIG, S.P. & HARFORD, T.
Correlates of therapists' initial impressions of patients in a psychiatric day center. Psychotherapy: Theory, Research and Practice, 1972, 9, 126-129

ROSS, A.O.
Diagnostic testing in a behaviourally oriented clinical training program. The Clinical Psychologist, 1974, 28, 16-18

ROSS, L.D.
ROTHMAN, D.J.
The discovery of the asylum. Boston: Little, Brown, 1971

ROTTER, J.B.
Some implications of a social learning theory for the prediction of goal directed behaviour from testing procedures. Psychological Review, 1960, 7, 301-316

RUBIN, H.
The Minnesota Multiphasic Personality Inventory as a diagnostic aid in a veterans hospital. Journal of Consulting Psychology, 1948, 12, 251-254

RUSS, S.N.

RUSSELL, B.

RUSSELL, B.
The Impact of Science on Society. New York: Simon and Schuster, 1953

RYAN, W.
Blaming the Victim. New York: Random House, 1971

RYCHLAK, J.F.

RYCHLAK, J.F.

SALTZMAN, L.

SAMPSON, E.E.

SAMRA, K.A.
A classic study in censorship. Schizophrenia, Newsletter of the American Schizophrenia Association, 1969, 3, 1-10

SAMUELS, H.

SANDERS, R. & CLEVELAND, S.E.
The relationship between certain examiner personality variables and subjects' Rorschach scores. Journal of Projective Techniques, 1953, 17, 34-50
SANDIFER, M.G., HORDERN, A., TIMERRY, G.C. & GREEN, L.M.

SANFORD, N.

SARBIN, T.R.

SARBIN, T.R., TAFT, R. & BAILEY, D.

SARGENT, W.

SATTE, J.M.
Racial "experimenter effects" in experimention testing, interviewing and psycho-therapy. Psychological Bulletin, 1970, 73, 137-160

SAWYER, J.

SCARBROUGH, H.E.
The nature of the dialogue with the obsessive compulsive. Psychotherapy: Theory, Research and Practice, 1966, 3, 33-35

SCHAFER, R.

SCHOFIELD, W.

SCHUR, E.M.

SCHWARTZ, M.L.
Validity and reliability in clinical judgments of C-U-S protocols as a function of amount of information and diagnostic category. Psychological Reports, 1967, 20, 767-774
SCHWARTZ, M.S.
Functions of the team in the state mental hospital. *American Journal of Orthopsychiatry*, 1960, 30, 100-102

SCOTT, D.M.

SEARS, R.R.

SEITZ, P.F.D.

SHADER, R.I., KELLAM, S.G. & DURELL, J.
Social field events during the first week of hospitalisation as predictors of treatment outcome for psychotic patients. *Journal of Nervous and Mental Disease*, 1967, 145, 142-153

SHAKOW, D.
*Clinical Psychology as Science and Profession.* Chicago: Aldine, 1969

SHAPIRO, M.B.
An experimental approach to diagnostic psychological testing. *Journal of Mental Science*, 1951, 97, 748-764

SHAPIRO, M.B.
The single case in fundamental clinical psychological research. *British Journal of Medical Psychology*, 1961, 34, 255-262

SHAPIRO, M.B.
The single case in clinical-psychological research. *Journal of General Psychology*, 1966, 74, 3-23

SHAPIRO, M.B.

SHARP, E.
The IQ Cult. *New York: Coward, McCann and Geoghegon*, 1972

SHEMBERG, K., & KEELEY, S.
Psychodiagnostic training in the academic setting: past and present. *Journal of Consulting and Clinical Psychology*, 1970, 34, 205-211

SHERMAN, T.M. & CORMIER, W.H.

SHOCKLEY, W.
SHWEDER, R.A.
The between and within of cross-cultural research. *Ethos*, 1973, 1, 531-43

SHWEDER, R.A.
How relevant is an individual difference theory of personality? *Journal of Personality*, 1975, 43, 455-484

SHWEDER, R.A.

SIEGEL, J.M.

SILVERMAN, L.H.

SIMON, B.
*Intelligence Testing and the Comprehensive School*. New York: Lawrence and Wishart, 1953

SIMON, B.

SIMON, J.L.

SINES, L.K.
The relative contribution of four kinds of data to accuracy in personality assessment. *Journal of Consulting Psychology*, 1959, 23, 483-492

SKINNER, B.F.

SLATER, E.

SMART, R.
The importance of negative results in psychological results. *Canadian Psychologist*, 1964, 5a, 225-232

SMEDSLUND, J.

SMITH, M.B.
SNYDER, C.R., SHENKEL, R.J. & LOWELY, C.R.

SNYDER, M. & SWANN, W.B.

SNYDER, M. & URANOWITZ, S.W.

SOSKIN, W.F.

SOSKIN, W.F.
Influence of 4 types of data on diagnostic conceptualisation in psychological testing. Journal of Abnormal and Social Psychology, 1959, 58, 69-78

SPANDS, N.P.

SPEARMAN, C.
"General intelligence" objectively determined and measured. American Journal of Psychology, 1904, 115, 201-292

SPENCER, M.

STAATS, A.

STAINBROOK, E.

STARR, B.J. & KATKIN, E.S.
The clinician as an aberrant actuary: Illusory correlations and the incomplete sentences blank. Journal of Abnormal Psychology, 1969, 74, 670-75

STRUPP, H.H.

STUART, R.B.
SUNDBERG, N.D.

SWAN, G.E. & MACDONALD, M.L.

SZASZ, T.

SZASZ, T.

SZASZ, T.S.

SZASZ, T.S.

SZASZ, T.S.

SZASZ, T.S.
The Manufacture of Madness. New York: Dell, 1971

SZASZ, T.

SZASZ, T.
Medicine, Cure and Control. Audiotape, Big Sur Recording, 1974

TAMBIAH, S.J.

TEMERLIN, M.
Suggestion effects in psychiatric diagnosis. *Journal of Nervous and Mental Disease*, 1968, 147, 349-353

TEMERLIN, M.K. & TROUSDALE, W.K.
The social psychology of clinical diagnosis. *Psychotherapy: Theory Research and Practice*, 1969, 6, 24-29

TERMAN, L.M.

TERMAN, L.M.
Feeble-minded children in the Public Schools of California. *School and Society*, 1917, 5, 161-165

TERMAN, L.M.
The Intelligence Quotient of Francis Galton in Childhood. *American Journal of Psychology*, 1917a, 28, 209-15
TERMAN, L.M.
Intelligence Tests and School Reorganisation. New York: World Book Co., 1923

TERMAN, L.M.
The conservation of talent. School and Society, 1924, 19, 363

TERMAN, L.M.

TERMAN, L.M.

TERMAN, L.M.

TERMAN, L. & MERILL, M.

THELEN, M.H., VARBLE, D.L. & JOHNSON, J.
Attitudes of academic clinical psychologists toward projective techniques. American Psychologist, 1968, 23, 517-521

THORESEN, C.

THORNDIKE, E.L.
Principles of Teaching. New York: Seiter, 1906

THORNDIKE, E.L.

THORNE, F.C.

TOWBIN, A.
When are cookbooks useful? American Psychologist, 1960, 15, 119-123

TRACHTMAN, J.P.

TRYON, W.W.

TRYON, W.W.
A System of Behavioural Diagnosis. Professional Psychology, 1976, Nov, 495-506
Tversky, A.
Features of similarity. Psychological Review, 1977, 84, 327-352

Tversky, A. & Kahneman, D.
Belief in the law of small numbers. Psychological Bulletin, 1971, 76, 104-10

Tversky, A. & Kahneman, D.

Tversky, A. & Kahneman, D.

Tversky, A. & Kahneman, D.

Tyler, L.T.
The psychology of human differences. New York: Appleton-Century-Crofts, 1965

Ullman, L.P. & Krasner, L.

Ullman, L.P. & Krasner, L.

Valentine, C.W.
The Psychology of Early Childhood. London: Methuen, 1942

Vernon, P.W.

Wachtel, P.L.

Wade, T.C. & Baker, T.B.

Wager, W.W.
WALLACE, J.

The Psychiatric nomenclature. Archives of General Psychiatry, 1962, 7, 198-205

WARD, W.C. & JENKINS, H.M.

WASON, P.C.
On the failure to eliminate hypotheses in a conceptional task. Quarterly Journal of Experimental Psychology, 1960, 12, 129-40

WASON, P.C.

WASON, P.C.

WASON, P.C. & JOHNSON-LAIRD, P.N.

WATSON, P.

WATSON, R.I.

WATZLAWICK, P., BEAVIN, J. & JACKSON, D.

WECHSLER, D.
Intelligence Defined and Undefined: A Relativistic Appraisal. American Psychologist, 1975, 30, 135-139

WEIMER, W.B.

WEIMER, W.B.
Notes on the methodology of scientific research. Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1979

WEINER, I.B.
Does psychodiagnosis have a future? Journal of Personality Assessment, 1972, 36, 534-46
WELLS, F.L.

WELSH, G.A.
A two dimensional personality model for research in social science. Research Previews (Institute for Research in Social Science, Chapel Hill, N.C.), 1972, 19, 14-23

WENGER, D.I. & FLETCHER, C.R.
The effect of legal counsel on admissions to a state mental hospital: a confrontation of professions. Journal of Health and Human Behaviour, 1969, 10, 66-72

WHYTE, L.L.

WHYTE, W.H., Jr.

WILCOX, R. & KRASNOFF, A.

WILDMAN, B.G. & ERICKSON, M.T.
Methodological problems in behavioural observation. In J.D. Cone and R.P. Hawkins, op cit, 1977

WILLIAMS, R.
Scientific racism and IQ: The silent mugging of the black community. Psychology Today, 1974, 7, 32

WILLS, T.A.

WILSON, G. & GRYLLS, D.

WINKLER, R.C.
Psychology as a social problem: How value-free is 'objective psychology'. Australian Psychologist, 1975, 8, 120-127

WINKLER, R.C.

WOODS, P.J.
A Taxonomy of Instrumental Conditioning. American Psychologist, 1974, 29, 584-597

YATES, A.J.
YATES, A.J.

YOUNG, R.M.
Evolutionary biology and ideology: then and now. Science Studies, 1971, 1, 177-206

YOUNG, R.M.
The historiographic and ideological contexts of the nineteenth century debate on man's place in nature. In M. Teich and R.M. Young (eds), Changing Perspectives in the History of Science: Essays in Honour of Joseph Needham. London: Heinemann, 1973

ZIGLER, E. & PHILLIPS, L.

ZILBOORG, G. & HENRY, G.W.
A history of medical psychology. New York: Norton, 1941

ZIMAN, J.M.

ZUBIN, J.

ADDITIONAL REFERENCES

BANDURA, A.

BANNISTER, D. & MAIR, J.M.

CAUTELA, J.R. & UPPER, D.

COHEN, R.S. & WARTOFSKY, M.W.

GOLDFIELD, M.R. & KERT, R.N.
HAMMOND, K.R.

HARTSHORNE, M. & MAY, M.S.

HAYNES, S.
Behavioural variance, individual differences, and trait theory in a behavioural construct system: A reappraisal. Behavioural Assessment, 1979, 1, 41-49

KANFER, F.M. & GRIMM, L.G.
Behavioural Assessment - selecting target behaviour in the interview. Behaviour Modification, 1977, 1, 7-29

LANGER, E.J. & ABELSON, R.P.
A patient by any other name: Clinician group differences in labelling bias. Journal of Consulting and Clinical Psychology, 1974, 42, 4-9

MASH, E.J.
What is Behavioural Assessment. Behavioural Assessment, 1979, 1, 23-29

MASH, E.J. & TERDAL, L.G.
Behaviour Therapy Assessment: Diagnosis, design and evaluation. New York: Springer, 1976

NELSON, R.O. & MAYLES, S.C.
Some current dimensions of behavioural assessment. Behavioural Assessment, 1979, 1, 1-16

OVERTON, W.F. & REESE, H.W.

PERVIN, L.A.

PETRINOVICH, L.

SHAPIRO, M.B.
Assessment interviewing in clinical psychology. British Journal of Social and Clinical Psychology, 1979, 13, 211-218

SPEARMAN, C.

STERN, G.G.
SUNDBERG, N.D.

SWAN, G.E. & MacDONALD, M.L.

WADE, T.C., BAKER, T.B. & HARTMANN, D.P.
Behaviour therapist's self-reported views and practices. Behaviour Therapist, 1979, 9, 2-6