

AN ABSTRACT OF THE THESIS OF

Louis L. Downs for the degree of Doctor of Philosophy in Counseling
presented on July 22, 1997. Title: A Comparison of Two Diagnostic
Models Using The Diagnostic and Statistical Manual of Mental Disorders:
Toward the Development of a Teaching Paradigm for Counselor Education

Abstract approved: _____

Redacted for privacy

James Firth

The present study was conducted to examine the effects of early orientation of counseling related students to the two most prevalent paradigms of psychodiagnostic decision-making on first, the integration of the model, and second, on the ability to make proficient diagnostic decisions while in training.

Using an experimental, pretest posttest design, 60 participants from two higher educational sites were randomly assigned to two treatment groups. Participants in each group were oriented to one of two treatment conditions -- a binary decision tree model or a problem-solving model (multiple competing hypotheses). Participants were then introduced to DSM Axis II diagnostic categories utilizing a computer assisted learning laboratory.

Results suggested that participants learned diagnosis during the experiment. However, no significant difference in diagnostic proficiency occurred as a result of the two treatment conditions.

Additional analyses raised questions about use of case studies as a means of assessing diagnostic proficiency. Item difficulty appeared to be linked to diagnostic clusters and individual diagnoses. Item difficulty factors influenced the internal consistency and validity of test instruments. The assumption of the unidimensional weight of syndromes in the construction of assessment instruments is suspect. Considering the preponderance of case study use for counselor training assessment, caution during instrument construction and use is advised.

Evidence also existed that treatment groups responded differently to particular DSM diagnostic clusters and items. This suggested that cluster and item difficulty may be important to consider for instruction of diagnosis in the classroom. Results also suggested that as diagnoses become more complex, problem-solving diagnostic decision-making may become more important.

Secondary analysis of computer assisted learning resulted in significant evidence that nonsequential, user-friendly computer assisted instruction may overcome teaching-study style mismatch, resulting in more even distribution of learning over the sample population.

©Copyright by Louis L. Downs
July 22, 1997
All Rights Reserved

A Comparison of Two Diagnostic Models Using
The *Diagnostic and Statistical Manual of Mental Disorders*:
Toward the Development of a Teaching Paradigm for Counselor Education

by

Louis L. Downs

A THESIS
submitted to
Oregon State University

in partial fulfillment of
the requirements for the
degree of

Doctor of Philosophy

Completed July 22, 1997
Commencement June 1998

Counseling Doctor of Philosophy thesis of Louis L. Downs presented on
July 22, 1997

APPROVED

Redacted for privacy

Major Professor, representing Counseling

Redacted for privacy

Director of School of Education

Redacted for privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for privacy

Louis L. Downs, Author

ACKNOWLEDGMENTS

I would like to thank a number of people who made valuable contributions to the development of this project. Brooke Collison gave the Counselor Education Course assignment that sparked my interest in the topic, resulting in this thesis. Mary Lou facilitated my development of counselor educator skills. Warren Suzuki provided mentorship for research and statistical skills. Brooke Collison provided valuable advice for this project.

The computer program was authored and graciously contributed by David Patterson, along with continuous correspondence of technical support. Isha Downs contributed the videotaping of orientation sessions with perseverance and expertise. Emotional support, tireless editing, and organizational aid were provided by Rita Downs. Validation of case studies was provided by Neil Williamson, Bernie Legner and John Gram. Diagnostic decision making orientation videos were viewed for potential bias by Annie Larrabee, Connie Harris and Susan Berg.

TABLE OF CONTENTS

	<u>Page</u>
INTRODUCTION.....	1
Introduction to the Research.....	1
Rationale for Research.....	1
Statement of Problem.....	4
Definitions.....	5
Null Hypotheses.....	8
REVIEW OF LITERATURE.....	9
Introduction.....	9
Adoption of DSM as a Standard.....	9
Historical Development of DSM Theory.....	11
Contemporary State of Psychodiagnostic Theory.....	12
The Bayesian Diagnostic Model.....	13
Competing Diagnostic Models.....	15
The Problem-Solving Diagnostic Model.....	16
The Structured Clinical Interview Diagnostic Model.....	21
Comparison of SCID to the Problem Solving Model.....	28
Review of Computer Assisted Learning Literature.....	29
Attitudes of Students Toward Computer Assisted Training.....	30
Skills Building Research.....	31
Components of Effective Computer Assisted Learning Software.....	32
Summary.....	33
DESIGN AND METHODOLOGY.....	34
Introduction.....	34
Assumptions Regarding Psychodiagnosis.....	34

TABLE OF CONTENTS (Continued)

Purpose of the Study.....	34
Research Questions.....	35
Research Question 1.....	35
Research Question 2.....	35
Research Methods and Procedures.....	35
Sample Population.....	36
Research Design.....	39
Research Procedure.....	39
Test Instrument.....	42
Analysis.....	45
Summary.....	47
FINDINGS AND DISCUSSION.....	48
Introduction.....	48
Preliminary Analysis of Instrument.....	48
Reliability Analysis of Three Point Scale.....	50
Reliability Analysis of One Point Scale.....	53
Hypothesis 1: Learning Effects.....	60
Hypothesis 2: Main Effect of Treatment Conditions.....	63
Unhypothesized Findings.....	65
Summary.....	70
Explanatory Variables.....	71
The Instrument.....	73
Association of Explanatory Variables with Diagnostic Learning.....	75

TABLE OF CONTENTS (Continued)

Association of Explanatory Variables with Treatment Effect.....	80
Summary.....	82
Item Content Analysis.....	82
Participant Attitudes Toward Use of Hyperaxis II.....	92
Summary.....	95
CONCLUSIONS, IMPLICATIONS, AND LIMITATIONS.....	98
Introduction.....	98
Summary of Study.....	98
Summary of Discussion.....	99
Limitations.....	101
Recommendations for Future Research.....	104
BIBLIOGRAPHY.....	108
APPENDICES.....	123
Appendix A Letters to Participants and Informed Consent Forms..	124
Appendix B Pretests and Posttests.....	134
Appendix C Proctor Scripts for Experiments.....	170
Appendix D Structured Posttest Interviews.....	178
Appendix E Problem-Solving Handout.....	180
Appendix F Video Scripts.....	184
Appendix G Interitem Correlations Using Three Point Scale.....	210
Appendix H Interitem Correlations Using One Point Scale.....	214
Appendix I Correlation of Scores to Explanatory Variables.....	218
Appendix J Item Difficulty Tables.....	224

LIST OF FIGURES

<u>Figure</u>		<u>Page</u>
1	Mean Three Point Scale Pretest and Posttest Scores.....	61
2	Mean One Point Scale Pretest and Posttest Scores.....	62
3	Mean One Point Scale Pretest Scores by Test Version and Number of Personality Disorder Cluster A Items Present.....	83
4	Mean One Point Scale Posttest Scores by Test Version and Number of Personality Disorder Cluster A Items Present.....	84
5	Mean One Point Scale Pretest Scores by Test Version and Number of Personality Disorder Cluster B Items Present.....	84
6	Mean One Point Scale Posttest Scores by Test Version and Number of Personality Disorder Cluster B Items Present.....	85
7	Mean One Point Scale Pretest Scores by Test Version and Number of Personality Disorder Cluster C Items Present.....	85
8	Mean One Point Scale Posttest Scores by Test Version and Number of Personality Disorder Cluster C Items Present.....	86

LIST OF TABLES

<u>Table</u>	<u>Page</u>
1 Gender and Education Levels of Site 1.....	39
2 Gender and Education Levels of Site 2.....	39
3 Item Assignment to Three Test Versions.....	44
4 Item Reliability in Pretest Version 1 Using Three Point Scale.....	51
5 Item Reliability in Posttest Version 1 Using Three Point Scale.....	51
6 Item Reliability in Pretest Version 2 Using Three Point Scale.....	52
7 Item Reliability in Posttest Version 2 Using Three Point Scale.....	52
8 Item Reliability in Pretest Version 3 Using Three Point Scale.....	53
9 Item Reliability in Posttest Version 3 Using Three Point Scale.....	53
10 Item Reliability in Pretest Version 1 Using One Point Scale.....	54
11 Item Reliability in Posttest Version 1 Using One Point Scale.....	55
12 Item Reliability in Pretest Version 2 Using One Point Scale.....	55
13 Item Reliability in Posttest Version 2 Using One Point Scale.....	56
14 Item Reliability in Pretest Version 3 Using One Point Scale.....	56
15 Item Reliability in Posttest Version 3 Using One Point Scale.....	57
16 Pretest Three Point Scale Mean Scores by Test Version.....	58
17 Posttest Three Point Scale Mean Scores by Test Version.....	58
18 Pretest One Point Scale Mean Scores by Test Version.....	59
19 Posttest One Point Scale Mean Scores by Test Version.....	59
20 Three Point Scale Posttest Mean Scores by Treatment Condition.....	63

LIST OF TABLES (Continued)

<u>Table</u>	<u>Page</u>
21 Three Point Scale Posttest Mean Scores by Treatment Condition and Test Version.....	64
22 One Point Scale Posttest Mean Scores by Treatment Condition.....	64
23 One Point Scale Posttest Mean Scores by Treatment Condition and Test Version.....	65
24 Ethnicity of Participants by Treatment Group and Test Version.....	72
25 Correlation of Variables to Posttest Scores by Test Version.....	74
26 Correlation of Variables to Difference between Pretest and Posttest Scores by Test Version.....	76
27 Correlations of Multiple Hypotheses Written on Posttest to Interview Ratings of Participant Multiple Hypothesis Use.....	79
28 Correlations of Explanatory Variables and Treatment Effects.....	81
29 Internal Consistency of Test Versions by Personality Disorder Cluster Using Three Point Scale.....	87
30 Internal Consistency of Test Versions by Personality Disorder Cluster Using One Point Scale.....	88
31 Frequency of Correct Responses by Treatment Condition.....	91
32 Hyperaxis II Ease of Use.....	93
33 Preference for Computer Assisted Learning.....	93
34 Pretest Individual Item Mean Scores Using Three Point Scale.....	210
35 Posttest Individual Item Mean Scores Using Three Point Scale.....	211
36 Interitem Correlations Using Three Point Scale.....	212
37 Pretest Individual Item Mean Scores Using One Point Scale.....	214
38 Posttest Individual Item Mean Scores Using One Point Scale.....	215

LIST OF TABLES (Continued)

<u>Table</u>	<u>Page</u>
39 Interitem Correlations Using One Point Scale.....	216
40 Correlations of Scores to Explanatory Variables Using Three Point Scale.....	218
41 Correlations of Scores to Explanatory Variables Using One Point Scale.....	221
42 Frequency of Correct Diagnosis by Item and Test Version.....	224
43 Reliability Coefficients by Item Difficulty.....	226
44 Percent Frequency of Correct Responses by Test Version and Diagnostic Cluster.....	227

To the memory of my uncle, William Ziegs

A Comparison of Two Diagnostic Models Using The Diagnostic and Statistical Manual of Mental Disorders: Toward the Development of a Teaching Paradigm for Counselor Education

CHAPTER I INTRODUCTION

Introduction to the Research

The statement of the problem, regarding creation of protocol for psychodiagnosis using the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994) and a teaching model for counselor education, is discussed in this chapter. The rationale for this study of teaching psychodiagnostic philosophy and definitions of terms used are submitted. Finally, the hypotheses upon which this research is founded, are stated.

Rationale for the Research

The field of counseling is in a state of evolution from guidance and school counseling toward a broader base of specialties. Originally based in the philosophy of Frank Parsons -- personality, interest and aptitude testing (Zunker, 1994) -- the field now includes specialties in disorder-based areas -- community, gerontology, marriage and family, chemical dependence and mental health counseling (CACREP, 1995). Zimpfer (1996) published two surveys of specialty fields in the counseling profession, reporting that private practice, which made up 13.6% of the population in 1986, represented 30% of the population in 1991. At the same time, higher education settings dropped from 25% to 22.8% of the population, and K-12 representation made only moderate gains, from 12.4% to 13.3%. As the counseling field has broadened to include new specialties, mental health practice in agencies (which require

diagnostic accountability) have suffered drops in population from an overall 25.4% to 16.6%.

School counselors are also being faced with student problems more in line with Roe's (1956) needs-based and Super's (1990) Segmental Model of guidance counseling rather than the more traditional trait and factor assessments (Holland, 1985; Strong & Campbell, 1974). To illustrate, Oregon Benchmarks, which outlines the educational and community goals of the State of Oregon, shows dramatic increases in drug use among students (Oregon Progress Board, 1994) as early as sixth grade. In 1994, 11% of sixth graders had used alcohol within the last month, and 6% were already smoking cigarettes. Twenty-three percent of eighth graders had used alcohol in the last month in 1990, 26% in 1992 and 30% by 1994, while 19% were smokers. Further, diagnosed incidence of Attention Deficit Disorder has risen to the point that the State of Oregon has established protocol for school monitoring of medications (Oregon Regulatory Statute No. 336.650, 1992).

Currently, Oregon educational institutions do not rely on school counselors to respond to these problems other than classroom observation and referral to school psychologists (P. Holstead, Oregon Department of Education, personal communication, October, 11, 1996), but have turned to community professionals through the establishment of integrated community services at school sites (Senate Bill 1099, 1991). Retention of outside professionals may suggest that school counselors may be considered only partially competent to handle nonacademic student problems.

Concurrently, counselors appear to be suffering from a lack of credibility (Ritchie, 1990) as school districts continue to question the efficacy of retaining counselors on staff (B. B. Collison, personal communication, November 29, 1995), and mental health agencies decrease the number of counselors, favoring therapists from other orientations (Zimpfer, 1996). West, Hosie and Mackey (1988) surveyed full-service mental health clinics and discovered that master's level counselors were hired by agencies in equal numbers to master's level psychology majors and hired to positions in 15% less agencies than social workers.

Mental health counselors tended to be assigned to client psychotherapy as well as to normal treatment team functions, but not to diagnostic duties except in regard to formal testing. Intelligence, personality, education, and vocational testing were not heavily emphasized, and less than half of the agencies assigned counselors to the task (West, Hosie, & Mackey, 1988). A survey of mental health administrators determined that the most important knowledge and skills for which administrators screened were counseling skills, understanding of client developmental levels and needs, ability to write case notes, ethical knowledge, cultural competence, group theory and diagnostic skills. Administrators considered counselors to be competent in all areas except those of case documentation and diagnostic skills (Cook, Berman, Genco, Repka, & Shrider, 1986). Again, West and associates' (1988) survey of mental health administrators found that master's level counselors working in mental health institutions were considered to be competent in direct delivery of services but needed strengthening of knowledge in psychotropic medication, psychopathology, individual assessment, and use of the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994).

Niles and Pate (1989) suggested, after comparing training standards for members of the American Mental Health Counselors Association to those of the National Career Development Association and the American Psychological Association, that there is a need to increase skills of AMHCA members, particularly their diagnostic skills. In fact, only one survey (Ginter, 1991) suggested that counselors need not increase psychodiagnostic skills. The 12 subjects in the survey were counselor educators or professional administrators of the American Mental Health Counselors Association. Further, the only suggestions the survey produced were to increase program evaluation and test interpretation skills. These areas are deemphasized in mental health counseling. In school counseling the trend is toward testing by psychologists, except to determine degree of developmental disorders and attention based disorders for purposes of establishing services eligibility (Grob, 1991).

Statement of the Problem

Counselor education literature consistently suggests the adoption of the Diagnostic and Statistical Manual of Mental Disorders as the standardized psychodiagnostic instrument in counselor education and in professional practice (Niles & Pate, 1989; Wilcoxon, 1990; Waldo, Brotherton, & Horswill, 1993; Seligman, 1993; Ritchie, 1990). Additionally, validity and reliability research has begun for use of the DSM in school settings (Sinclair, Forness, & Alexson, 1985), although little evidence has been presented to suggest potential applications.

Nevertheless, development and institutionalization of curricula to teach DSM has been slow. Ritchie, Piazza, and Lewton (1991) surveyed counselor education programs and discovered that at the end of 1980s only 27% of counselor education programs had a distinct course in DSM diagnosis. In addition, efforts appeared to be singular and disorganized. Hohenshil, (1993) delineated reasons for the use of DSM in training and reviewed issues of concern, but only vaguely suggested that a medical model of training be adopted to increase expertise. The author did not elaborate. Fong (1993) has attempted to provide a framework for teaching diagnostic skills, though no research has yet provided evidence of effective results with counselor trainees.

Further, psychological diagnosis in counselor education has been approached with suspicion (Cook, Warnke, & Dupuy, 1993; Velasquez, Johnson, & Brown-Cheatham, 1993), and although counselors have been trained in the use of the DSM no training model has been tested for efficacy with counselor trainees. The purpose of this study, then, is to investigate and compare psychodiagnostic models in common use. Further, the study is intended to compare models as instructional applications and to study the development of diagnostic proficiency with each model.

Definitions

Axis I: A means “for reporting all the various disorders or conditions in the Classification except for the Personality Disorders and Mental Retardation,” using the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994, p. 25).

Axis II: A means “for reporting Personality Disorders and Mental Retardation” using the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994, p. 26).

Card: An individual “page,” “screen,” or “field” in any Hypermedia, hypertext or Hypercard software program (Patterson & Yaffe, 1993, p. 268).

Computer assisted instruction (CAI), computer assisted learning (CAL): “The presentation of lesson material and related questions through the use of a computer” (MacFadden, 1990, p. 29).

Counselor: “A counselor who has received a master’s degree or higher from an entry-level program in counselor education or a closely related field, preferably matching the standards outlined by CACREP” (CACREP, 1994, p. 103).

Decision tree: A binary approach to psychodiagnostics in which the clinician asks, either to self or to client, a structured series of yes or no questions intended to confirm or rule out a single, suspected diagnosis (Williams, Gibbon, First, Spitzer, Davies, Borus, Howes, Kane, Pope, Rounsaville, & Wittchen, 1992, p. 630).

Diagnostic proficiency: The ratio of helpful diagnostic decisions in relation to the total number of diagnostic decisions made (Berven & Scofield, p. 398).

DSM Decision Tree: The specific decision tree published in the Diagnostic and Statistical Manual of Mental Disorders (American Psychiatric Association, 1994, pp. 689-701).

Heuristic: "Each piece of information requested...should be related to a plan...for solving the problem...no diagnostic hypothesis should be more specific or more general than the evidence...there should always be at least two or three competing hypotheses under consideration...and whenever...a new hypothesis emerges information previously collected should be reviewed..." (Elstein, et al., 1978, p. 253).

Hypercard: "A Hypermedia program" with "total informational content of the software available to the user. A stack of cards, with each card holding some information. Hypermedia is unlike a book or movie in which users encounter information in a predetermined linear fashion....Hypermedia makes possible nonsequential information access" (Patterson, Lee, & Evers, 1992, p.4).

Hypothesis: "A proposition about relationships set forth to explain a given set of phenomena" (McGuire, 1985).

Nosology: "The study and classification of diseases. Discovery of symptoms and consequent grouping into syndromes is the main area of concern. The delineation and definition of diseases perform four major functions in the...field: classification of terminology, the categorizing of names and codes within each classification, the establishing reliable and specific procedures for collecting information, and the operationalizing of rules for making classifications. The three main purposes or uses of information gathered by nosology are to serve as a guide to selection of treatment, to make prognoses, and to function as administrative devices" (Brenner, 1985, p. 762).

Parsimony: The preference or tendency of "trying to fit all the patient's symptoms and signs into one disorder." (Reid & Wise, 1989, p. 13).

Personality Disorder: “[Traits] characterized by deeply ingrained, generally lifelong maladaptive patterns of behavior that are usually recognizable at adolescence or earlier” (Kaplan & Sadock, 1990, p. 5).

Problem-solving: a multiple hypothesis approach to diagnostics in which the clinician entertains simultaneously as many diagnoses as are suggested by each sign or symptom of the client until compiled data rule out any one diagnosis or diagnoses (Elstein, Shulman, Sprafka, Allal, Gordon, Jason, Kagan, Loupe, & Jordan, 1978).

Signs: “...signs are phenomena which the psychiatrist can observe more or less objectively” (Sullivan, 1954, p. 183).

Stack: A group of cards focused on a single subject or related subjects, accessible in a given Hypercard, hypertext or Hypermedia program (Patterson & Yaffe, 1993, pp. 268-269).

Symptoms: “Any condition accompanying or resulting from a [disorder] and serving as an aid in diagnosis” (McKechnie, 1979, p. 1849); “Symptoms must be reported by the patient; in other words, only the patient experiences the symptoms” (Sullivan, 1954, p. 183).

Taxonomy: “The science of classification, that is, the assignment of individuals to groups within a system of categories distinguished by a priori characteristics. In essence, one may form homogeneous groups by assessing “similarities” or assign to mutually exclusive groups by assessing “dissimilarities.” As will be noted, typology is a special case of systematic classification: namely, the study to types” (Prentky, 1994, p. 507).

Null Hypotheses

The purpose of this study is to compare the Structured Clinical Interview diagnostic model to the problem solving diagnostic model as instructional foundations and to measure the resultant diagnostic proficiency of counseling students. The null hypotheses are as follows.

1. There is no significant difference between psychodiagnostic proficiency of students before and after receiving training in DSM taxonomy.

2. There is no significant difference between psychodiagnostic proficiency of students who learn DSM taxonomy after receiving prescribed training in Structured Clinical Interview (single hypothesis decision tree) and problem-solving technique (multiple, competing hypotheses).

CHAPTER II REVIEW OF LITERATURE

Introduction

This chapter presents a review of literature. An overview of the universality of the Diagnostic and Statistical Manual of Mental Disorders, theoretic foundations for development of the DSM, and the theoretic influences on the evolution of the diagnostic instrument are explored. The significance of diagnostic theory is presented, and developmental deficiencies in the counseling field toward establishment of a hegemony foundational to a proposed theoretic interpretation of DSM taxonomy is discussed. Because this study compares two distinct approaches to teaching psychodiagnostics -- a theory-based approach to DSM diagnosis instruction and computer assisted learning -- the literature of both fields will be reviewed.

Adoption of the DSM as a Standard

The DSM has become the most widely used instrument for assessment of psychological disorders in existence. Maser, Kaelber, & Weise (1991) compared the use of the International Classification of Diseases, the ICD-9 (World Health Organization, 1987), to the DSM internationally, and discovered that the DSM has taken precedence over the ICD for some clinical purposes and is used conjointly in most others. The authors also discovered that 95% of recipients used the DSM for teaching, 97% for research, and 81% for clinical practice. Sample errors weaken inference generalization, but this most ambitious of surveys thus far achieved suggests strongly that influence of the DSM is growing steadily. A study of British psychiatrists in training (Macaskill, Geddes, & Macaskill, 1991) reported that 93% of curricula included some DSM training and 73% used the instrument extensively in training.

A study of social work curricula (Raffoul & Holmes, 1986) reported that 33% of surveyed programs offered a specific course in DSM, and that course

content devoted specifically to DSM ranged from 15% to 100%. Further, DSM training occurred in 53% of in-service agency training, 16% of agency workshops, and 7% of faculty workshops.

Another survey -- of practicing social workers (Kutchins & Kirk, 1988) -- reported that, although as high as 69% of social workers did not believe that the DSM was of major importance for analysis of client behavior, 81% reported that the DSM was very important for third party payment and 92% considered it of some importance for third party payment. There is little doubt that the DSM is becoming the universally accepted standard for psychodiagnostics, whether for nosological, case management, research, instructional, or mercenary purposes (Kutchins & Kirk, 1988).

The diversity of purpose for and attitude toward the DSM demonstrates a variety of orientations toward the instrument. Diagnosis has ramifications of stigmatization, particularly in an age of managed care (Shore, 1992). Insurance companies presently require detailed reports from practitioners and have no ethical obligation to confidentiality. Thus, the cavalier attitude toward taxonomy supposed to exist in social work may be more an issue of ethics than simply an adaptation. A field that has not thoughtfully considered client welfare during adoption of a taxonomical system provides little foundation for exploration, particularly in light of research that gives evidence of the importance of counselor attitudes and therapeutically meaningful relationships with clients (Atkinson, Worthington, Dana, & Good, 1991; Barrett-Lennard, 1962; Cummings, Martin, Hallberg, & Slemon, 1992; Smith & Glass, 1977). Psychiatry, however, has been embroiled in the nosological debate for over a century. Therefore, investigation of theoretic foundation should occur within the historical confines of the field most likely to present an astute appraisal of taxonomy. The historical development by the psychiatric community offers a rich tradition of theoretic debate as a potential model for development of counseling nosological theory.

Historical Development of DSM Theory

The DSM-I was the product of a major change in thinking in the psychiatric community (Grob, 1991). Before World War II, the Association of Medical Superintendents of American Institutions for the Insane, the precursor organization to the American Psychiatric Association, struggled between debate over specific diagnoses and the stated belief that no practical nosological system could be developed without first understanding the pathology or etiology of the disorders (Rosenberg, 1979).

Attempts at a nosological system were, until World War II, the interest of the Bureau of Census for statistical and policy purposes (Grob, 1991). The War produced 2400 Army-assigned psychiatrists, more than doubling the ranks of the American Psychiatric Association (APA). These new APA members tended to be psychodynamic in orientation and influenced by experience with soldiers. Since most military clients had been suffering from stress related disorders, the new psychiatrists were interested in nonpsychotic syndromes, unlike their more traditional colleagues. Military psychiatrists had also discovered that soldiers treated in supportive environments, located near their units, returned to duty more quickly. So, upon reentering civilian careers, the "young Turks" (Grob, 1991, p. 427) formally espoused a social activist stance. William Menninger -- one of the most influential of the new APA members, serving after the War as Surgeon General -- called for a revolution within the APA to establish a philosophy of concern for prevention as well as for social needs and problems.

An outcome of the movement was the publication of the DSM-I in 1952. The intention of the document was to reify nonpsychotic disorders at a level with psychosis. However, the document was not intended to be theoretical in philosophy (Nelson-Gray, 1991; Frances, First, Widiger, Miele, Tilly, Davis, & Pincus, 1991). The objective of the original document was to supply descriptive dimensions (Frances, et. al., 1991) which would be flexible (Grob, 1991) and multidimensional (Widiger, Frances, Pincus, Davis, & First, 1991). In other words, no assumption of singular diagnosis nor of classification rigidity was intended. Rather, dimensions and patterns of behaviors were to be the

assumption of the APA. As Grob (1991, p. 421) stated, "Classification systems are neither inherently self-evident nor given. On the contrary, they emerge from the crucible of human experience; change and variability, not immutability, are characteristic. Indeed, the ways in which data are organized at various times reflect specific historical circumstances. Empirical data, after all, can be presented and analyzed in endless varieties of ways."

Since 1952, the DSM has been revised four times and has reflected Grob's statement not so much for its emphasis on phenomenology but on "historical circumstances." Medical explanations of pathologies, influenced by psychotropic medications and physiological studies, have redirected impetus from working nosology toward taxonomy (Brenneis, 1994). As various DSM revisions evolved they increasingly reflected the influence of biomedical theory.

Contemporary State of Psychodiagnostic Theory

The field of psychiatry has been convinced throughout the century of nosology debate that the cognitive process by which a diagnostician considers, decides and applies taxonomy is critical to healing (Grob, 1991). Two schools of thought presently exist which emphasize the thought process of the diagnostician and the effects of that process on proficiency: Bayesian (McGuire, 1985) and problem-solving (Elstein, et al., 1978). A third school of thought is based in binary decision tree (American Psychiatric Association, 1994, pp. 689-701) and is best represented by the Structure Clinical Interview for the DSM, SCID (First, Spitzer, Gibbon, & Williams, 1996).

Two of the models are currently in wide use, binary decision tree and problem-solving. Although these two models compete for preeminence as models for diagnosis, the third, Bayesian, continues to be considered periodically and so deserves attention in review of the literature.

The Bayesian Diagnostic Model

The Bayesian cognitive model of diagnosis, the most recent hypothesis to explain proficient diagnostic thinking, was first advanced by Warner and associates (1961) in relation to physiological diagnosis. The postulate that proficient diagnosticians base taxonomical decisions on the existing probability of incidence of any given disorder, rather than by some linear method, was reiterated by Jacquez (1964). It was then used as a theoretic base for the singular research done in psychodiagnostics based on the theory. Smith (1966), during a critique of methodology for research of interrater reliability for DSM diagnosis, used meta-analysis in an attempt to give evidence that interrater reliability is much higher than reported by the studies he analyzed.

The author was able to show, using the Bayesian Model to reanalyze data from several reliability studies of DSM, that interrater reliability could be achieved at no less than .87 and, with one added assumption, .97. However, the methodology was based in suppositions which stretch the credibility of the model beyond the acceptability of the psychiatric community. The first assumption upon which Smith based analysis was that all psychiatric symptoms are independent. Second, he postulated that all disorders are mutually exclusive -- if a client can be classified into one syndrome, then no other syndrome may be present and no symptoms from other syndromes exist. Ultimately, he assumed (although he admitted that this third assumption was impossible, albeit necessary to his model) that the probability of occurrence of any diagnosis was equal to that of any other. This third assumption, oddly, seems incompatible with "conditional probability decision model" (Smith, 1966, p. 529) described by the author. Finally, because he could only achieve .87 with this most implausible set of assumptions, he achieved a .97 reliability by disclaiming disagreement between psychiatrists. The author restated the diagnoses of individual clinicians he studied, generalizing responses of individual raters until separate criteria, divergent symptomology and distinct diagnoses overlapped.

Later, Lusted (1968) argued for the Bayesian model, but without research evidence. Elstein and colleagues (1978) integrated the Bayesian concept of incidence probability of disorders into the the knowledge base required to attain diagnostic proficiency using problem-solving methodology, although Bayesian information was deemphasized beyond its role as a part of the milieu of baseline diagnostic information. The model appeared in medicine infrequently, surfacing again in an article by McGuire (1985), who used the premises of the Bayesian model as the basis for a critique of the problem-solving model of psychodiagnosis. The author renamed the Bayesian Model as "decision analysis." Citing diverse studies into college student decision-making, McGuire suggested that no one model of cognition could explain diagnostic decision-making, and reiterated the claim of Smith (1966) that proficient diagnosticians choose diagnoses based on probability of occurrence.

Another application of Bayesian theory (Hogarth, 1974) attempted to establish the model within psychology to increase diagnostic accuracy. Basing diagnostic strategy on forward, stepwise regression, the author reanalyzed the DSM validity research of Einhorn (1972) using a statistical model which weighted certain cases and clusters, and which added variables (unmeasured in the original study and estimated by Hogarth) including number of cue stimuli present at any time during interview, cost of each diagnosis, mistakes made during formulation, amount of time needed to establish diagnosis and number of thought stages needed for a conclusion. The author concluded that the model would increase reliability to 1.0 with a small sample and suggested that psychodiagnosticians apply the formula to each diagnosis mathematically.

More recent studies applied Bayesian decision analysis beyond mathematical models. Shamian (1991) discovered that decision analysis trained nursing students, practicing physical diagnosis, made clinical decisions more consistent with expert decisions than did the control group ($p < .0001$). Interrater reliability of the experimental group was .999, while interrater reliability of the control group was .729. However, while the experimental group received four hours of diagnostic training, controls received none -- watching four hours of diagnostically irrelevant movies instead. Another study (Friedman,

Massaro, Kitzis, & Cohen, 1995) compared the Bayesian diagnostic model to four other diagnostic approaches, with medical students diagnosing physiological rather than psychological disorders. It was discovered that subjects tended to diagnose more congruently with probabilities but not as accurately as in any other model studied. The Friedman and associates (1995) conclusions are important to apply to the most recent study (Cosmides & Tooby, 1996). Working with medical students to compare the Bayesian diagnostic model to a problem-solving model, researchers presented evidence that Bayesian subjects increased congruence with probability of occurrence of disorders by 16% over problem-solving subjects, decreased occurrence of single diagnosis by 20%, were less confident of diagnosis and were less concerned about false positives because they were able to apply Bayesian statistical principles. A study by Clay and associates (1995) was able to reduce misdiagnosis of depressive symptomology in physically injured patients using Bayesian analysis in the place of a structured interview protocol.

Although there is merit in consideration of probability-based decision-making for research purposes, the model must assume some knowledge on the part of the clinician regarding incidence of various disorders. This suggests that diagnosticians gain expertise and proficiency with added knowledge base or experience. Conversely, research gives convincing evidence that experience and knowledge does not correlate with diagnostic skill (de Mesquita, 1992; Elstein, et al., 1978). Also, the assumption that with a complicated mathematical model, human judgment may be replaced would need more than reanalysis of gathered data; the hypothesis would need rigorous field testing to validate. Currently, no evidence exists to substantiate the hypothesis, and it has been relegated to obscurity in the field of psychiatry.

Competing Diagnostic Models

The psychiatric community has long been interested in nosology. Throughout the history of psychiatry, debate over diagnosis has occurred because of the concern that, without a taxonomical order, chaos would develop

(Grob, 1991). Since World War II, two distinct nosological models have competed for dominance. The earliest, a medical model, attempted to establish a taxonomical system which was challenged by the Menninger associates, who held that individual diagnosis is phenomenological and systemic. Consequently, no distinct category may describe a client's behaviors. Although the Menninger associates were responsible for publication of the first DSM (American Psychiatric Association, 1952), the DSM was presented as prototypical rather than categorical. In other words, flexibility in both the practice of diagnosis and in the interpretation of the instrument were considered appropriate.

The Problem-Solving Diagnostic Model

As the DSM began to show evidence of increased influence by the medical model, independent research began to study the thought process intrinsic to diagnosis. In response to the increasing emphasis in medical education on process over content, research attempted to establish a model to increase proficiency in diagnosis. Researchers at Michigan State University (Elstein, et al., 1978) began a five year study of diagnostic process -- the most extensive to date. The studies were methodologically designed to generalize across the medical fields and so included psychiatry.

The authors coined the name "problem-solving" and "patient-management problems" (Elstein, et al., 1978, pp. 3, 122) to describe the research track, which developed through several studies; the names held. The first study used a qualitative approach, selecting a criterion group from physician nominations of colleagues, respected for diagnostic ability, from several specialties. A noncriterion group was chosen randomly from the regional medical community and invited by deception: potential noncriterion subjects were told that they were chosen because of diagnostic expertise. Hypotheses about thought process were generated by observing physicians using case simulations utilizing live actors, identifying thought processes and coding them, then building matrices for comparison of groups. The first

evidence suggested that experience has little to do with proficiency: proficient diagnosticians began accurate diagnosis early in their careers, and physicians who were not proficient appeared not to develop skills over time. Researchers also discovered that proficient diagnosticians began to entertain hypotheses during the first few minutes of a patient interview, entertaining as many as were cued by the patient's signs and symptoms which regulated the interview. Efficient, generalized "search units" (Elstein, et al., 1978, p. 56) and continuous addition and rule out of hypotheses throughout the interview were hallmarks of the criterion group, producing a significantly more accurate diagnosis, with fewer overinterpretations, underinterpretations or misinterpretations, sooner than the noncriterion group.

Hypotheses developed during the qualitative study were tested quantitatively, again using case simulations and practicing physicians. Once accurate diagnosticians were identified from the study, statistical analysis discovered a trend toward more thorough data collection among physicians who entertained competing hypotheses than among physicians entertaining single hypotheses (range from $p < .10$ to $p < .05$), fewer mistakes made by accurate diagnosticians during formulation ($p < .05$), and fewer cues used ($p < .05$). Authors concluded that proficient diagnosticians asked fewer, more general questions (based on multiple hypotheses, generated early and throughout the interview) made less errors, and were quicker to diagnose than less proficient physicians.

A series of studies followed, testing primarily the developing model, and later, investigating teaching methods. Research, used to develop the problem-solving model, first studied complicated case simulations, discovering that proficiency tends to be related to conceptualization and that patient psychological variables produced the most errors. Another study of fixed-order problems discovered a high correlation between cue consistency and diagnostic specificity (range from .80 to .92, $p < .05$) and concluded that the number of serious hypotheses entertained by proficient diagnosticians was higher ($p < .01$). Filmed interviews tested initial problem formulations and discovered that 35 % of subjects developed hypotheses within 32 seconds and

53% began to generate hypotheses with the first presentation by the patient. Verbal cues predominated, but nonverbal cues were used for formulation as well. Authors concluded that direct associative retrieval rather than strategy-guided search typified proficient diagnosticians (Elstein, et al., 1978, p. 199).

Several studies to develop the underlying skills of proficient diagnosticians in medical students have been undertaken. The first, using simulated cases, tested process and outcome feedback, reported that outcome feedback was significantly more effective than process and outcome feedback or process feedback alone ($p < .01$). A second study discovered that subjects using a think-aloud protocol during practice developed more hypotheses ($p < .04$) and continued to entertain new hypotheses longer ($p < .05$). Finally, the heuristic formula of Elstein and colleagues (1978) was tested with professionals and medical students for accuracy and cost and reported a significant effect of heuristics on cost and accuracy ($p < .012$). The mean interrater reliability for subject performance accuracy was .91. The authors concluded that they had discovered an effective model for instruction of diagnostics for physicians.

Continued research in problem-solving medical education has produced four studies of importance to counseling. The first study (Smith, Galdes, Fraser, & Miller, 1991) considered the problem of multiple solution problems. The assumption that multiple hypotheses may not narrow to a single diagnosis suggested to the authors that more than one solution may also be in order, whether one or multiple diagnoses exist. To increase safety for clients, a model was produced from a case study with the following four procedures, (a) using data patterns to simplify the problem, multiple hypotheses equal to all potential interventions necessary to solve problems, assuming an equal number of single-solution problems, (b) using a mixture of data and hypothesis-driven processes to counteract bias, (c) using confirming and rule out processes to provide varied evidence, and (d) using models for error to think about hypotheses as they develop.

Another study (Boshuizen, Machiels-Bongaerts, Schmidt, & Hermans, 1995), this time including students in behavioral science, studied a complex

problems examination. Participants were medical students having received problem-solving training versus traditional Dutch training. Problem-solving subjects scored significantly higher on tests over three years than did controls ($p < .0001$). Some of the same authors (Schmidt, Machiels-Bongaerts, Hermans, Cate, Venekamp, & Boshuizen, 1995) compared problem-solving training to traditional training and to an integrated model (problem-solving and combined traditional methodologies including decision tree), reporting that no difference occurred between the integrated and problem-solving models. The authors reported that a significant difference occurred between the two models and traditional models ($p < .0001$), that clinical judgment improved over time with problem-solving and integrated models ($p < .0001$), and that problem-solving and time interacted ($p < .001$).

Finally, Goss (1996) compared problem-solving training taught at various levels of medical education in an educational case study. The author reported that students who began early orientation to problem-solving exhibited an increase of clinical judgment compared to later oriented students.

Concern with counseling skills and the potential of skills deterioration while learning diagnostic skills motivated Bögels (1994) to develop a model to integrate process, content, and cognitive skills into the diagnostic interview. The Diagnostic Interviewing Rating Scale in Mental Health Care was developed and tested over three years with psychologist trainees, giving evidence that patients were more satisfied with process skills ($r = .68$) but that diagnostic accuracy was not achieved ($r = .45$). Considering the results of research by Goss (1996) and Schmidt and colleagues (1995), it is not surprising that Bögels' subjects, advanced students who learned the model late, had difficulty with diagnosis.

Several studies (de Mesquita, 1992; Lambert & Meier, 1992; Janikowski, et al., 1989; Berven & Scofield, 1980; Hayden, 1990; Patterson & Yaffe, 1993) have been undertaken to investigate diagnostic training of counselor related skills using computer-assisted learning programs to allow students practice. Of these computer-assisted studies, four studied varied forms of problem-solving protocol on diagnostic skills of students. Foundational research with counselor

related fields (de Mesquita, 1992) compared professional and student educational psychologists and master's level counselor trainees and professionals, providing evidence that experience made no difference to diagnostic proficiency, but that cognitive process did. The author discovered that accurate diagnosticians requested less information ($p < .001$), required less time ($p < .0001$), gathered less information ($p < .02$), and entertained more hypotheses ($p < .0001$). The author also discovered that think-aloud protocols, designed to increase practice, resulted in better scores ($p < .0001$).

Lambert and Meier (1992) reported increased skills using a problem-solving CAI ($p < .05$) but discovered a more significant site effect ($p < .005$). The authors suggested that orientation affected learning of problem-solving skills.

Janikowski and associates (1989) introduced a patient problems program for rehabilitation counseling students and discovered a learning effect ($p < .05$). The authors also discovered that subjects had difficulty predicting client behavior.

Berven and Scofield (1980) studied counseling students and professionals using a patient problems program designed to simulate several months of interventions, and reported that diagnostic skills were learned early ($p < .05$). There was evidence of a relationship between training, experience and performance ($p < .001$), but the author did not investigate further.

Chan, Rosen, Wong, and Kaplan (1993) updated Berven and Scofield's (1980) rehabilitation counselor training model with the patient problem management CAI. The research group reported that over the course of five simulations the difference in diagnostic proficiency between professionals and students disappeared.

Interestingly, computer assisted learning of diagnostic skills, which relied solely on practice of taxonomy without orientation to a form of a problem-solving model, resulted in no improvement in skills compared to reading the DSM itself (Hayden, 1990; Patterson & Yaffe, 1993).

Problem-solving diagnostic research has developed a theory and praxis that assume psychological taxonomy is prototypical, not categorical. The theory is congruent with a field that has not, as of yet, discovered the pathology and etiology of disorders it treats and so is also congruent with both early pioneers and recent psychiatrists who have resisted rigid taxonomy. As Pliny Earl stated eloquently in 1888,

In the present state of our knowledge no classification of insanity [or other disorder] can be erected upon a pathological basis, for the simple reason that, with but slight exceptions, the pathology of the disease is unknown...Hence, for the most apparent, the most clearly defined, and the best understood foundation for a nosological scheme for insanity, we are forced to fall back upon the symptoms of the disease -- the apparent mental condition, as judged from the outward manifestations (Grob, 1991).

Cleghorn (1985, p. 504) restated the argument; it is the validity of "meaning" and "mechanism" that are weak in classification. In other words, etiology and phenomenology are the primary requisites of effective counseling. Further, systemic factors (i.e., culture, socioeconomics, family) preclude rigid taxonomy just as they preclude monolithic treatment (Brenneis, 1994).

However, problem-solving technique was developed not for philosophical but for practical reasons. Foundationally, the model was developed to adapt to changes in medical school training, from content to process. Criteria were established from medical clinicians renowned for expertise in diagnosis. Upon that basis, teaching technique was developed and tested to duplicate proficiency in medical students.

The Structured Clinical Interview Diagnostic Model

The decision tree diagnostic process developed from an entirely different foundation and with a dissimilar focus than did problem-solving. The DSM Decision Tree (American Psychiatric Association, 1994, pp. 689-701) was developed as a feature of the Diagnostic and Statistical Manual of Mental

Disorders for the original purpose of consultation when considering features of a case (Saigh, 1991). The American Psychiatric Association, however, has stated, "In DSM-IV there is no assumption that each category of mental disorder is a completely discrete entity with absolute boundaries dividing it from other mental disorders or from no mental disorder....It is suggested that the DSM-IV Classification be organized following a dimensional model rather than the categorical model" (American Psychiatric Association, 1994, p. XXII).

Nevertheless, the developing dominance of a distinct philosophy of psychopathology and the reintroduction of the medical model has increasingly influenced not only attempts to validate single syndrome diagnosis but has contributed to the development of theory and technique of interview and of teaching of diagnostic practice. The Structured Clinical Interview for DSM (Williams, et. al., 1992) has been developed to increase proficiency of single diagnosis and to study methods of increasing validity of the DSM as an taxonomical instrument.

Further, the advent of the computer has significantly affected psychodiagnostics. Concurrent with the progression of software intended to diagnose disorders for the practitioner (Sptizer & Endicott, 1968; First, Williams, & Spitzer, 1988; Furlong & Hayden, 1993), binary decision tree tests (Williams, et al., 1992) have been developed based strictly on the DSM Decision Tree. The latest addition to the binary decision tree software has been the adaptation of DTree software as a teaching program, DTree and Autoscid (First, et al., 1988). Furlong and Hayden (1993), writing in Counselor Education and Supervision, endorsed the use of the teaching instruments, although the authors describe the software inaccurately. So, binary decision tree has been introduced into counselor education, even though no research evidence of the effects of parsimonious training on proficiency of students of psychodiagnosis has been presented.

Conceptually, the clinician can ask a dichotomous question as to the existence of a particular, major symptom of a suspected syndrome. If the client shows signs or reports symptoms, the clinician may assume to be on the correct diagnostic track and ask the next question for confirmation of the prospective

diagnosis. A sequence of queries or observations, based on the prioritized symptoms of the potential diagnosis, ensues. So long as the clinician receives confirming responses, the diagnostic interview continues along the path of confirmation of diagnosis. At any point, if the clinician receives a negative sign or reported symptom, she or he moves to the next most logical diagnostic track, based on data already gathered. The outcome is a foregone conclusion. Once a diagnosis is confirmed, treatment of the disorder follows.

Advocates of the approach have declared, "Parsimony is a good tool" (Reid & Wise, 1989, p. 13). The principle of parsimony has implications pervasive throughout the field of psychotherapy. State regulatory agencies adopt case management rules based in parsimony.

For example, the Oregon Alcohol and Drug Abuse Division requires that all counselors working in a State licensed facility reevaluate each client's case, including diagnosis -- generally a single diagnosis with diagnostically supportive problem statements -- 30 days after the treatment plan is written and every 60 days thereafter (ORS #309-51-035 [7][a][b]).

In fact, parsimony is reflected in the strongly argued assumption that the single diagnosis, chemical dependence, requires a separate treatment field (Vaillant, 1981; Murphy, 1980; Krueger, 1982). The assumption pervades even research. For instance, a study of Alcoholics Anonymous members (Brown, 1985) reported that psychotherapy is ineffective for active addicts. The conclusion was based on evidence that 30% of the of the 75% of participants who had attended psychotherapy did not find it useful.

Managed care protocol is also built on a parsimonious principle. A client diagnosis is discussed, with supporting data, and a preagreed number of sessions allotted. Treatment type is approved based on diagnosis. Changes in diagnosis or in treatment, including referral to other professionals are approved on the basis of specific diagnosis and justifying evidence (Shore, 1992).

Some psychotherapeutic schools of thought, which do not espouse use of DSM taxonomy, base counseling on parsimonious factors. The primary assumption of Brief Therapy is parsimony. Cade and O'Hanlon (1993) state that encouragement of affective expression in counseling may not be

therapeutic. Emphasis is so strong on this assumption that the authors quote Schachter and Singer (1962, p. 381), "To the extent that cognitive factors are potent determinants of emotional states, it could be anticipated that precisely the same state of physiological arousal could be labeled 'joy' or 'fury' or 'jealousy' or any other of a great diversity of emotional labels depending on the cognitive aspects of the situation." Brief counseling protocol uses the assessment session to quickly develop objectives to resolve a predominant problem (Cade & O'Hanlon, 1993).

Psychiatric treatment protocol differs little from other psychological schools which espouse parsimonious diagnosis, assuming that medication for the dominant diagnosis would be generally adequate. No other diagnosis is warranted unless treatment is found to be unsuccessful. At the point that the client does not respond favorably to a planned intervention, a case management decision to reassess diagnosis is appropriate. Otherwise, stabilization and discharge are standard procedure, utilizing a behaviorist case management system -- the identification of specific problems with time limited objectives and specific interventions preagreed between client and counselor (Hersen & Bellack, 1981, pp. 14-29).

The influence of binary decision making appears to be strong in the field at large. In fact, binary diagnostics have been documented in research more than any other form of diagnosis.

Several of the authors of the DSM have developed the Structured Clinical Interview for DSM, SCID, which has become a standard in the psychodiagnostic field. The first question which must be asked is, "How reliable is the instrument?" The authors (Williams, et al., 1992) tested the SCID in a multi-site study, setting an acceptable kappa level of .5 for interrater reliability, although it was not consistently achieved. The highest mean reliability was .62 and the lowest mean reliability was .47. Item analysis of individual diagnostic categories resulted in a range of interrater reliability from .00 to .90 within sites. Test-retest results of individual diagnoses were reported to range from .25 to .85.

A follow-up, multi-site study (First, et al., 1995) reported kappa correlations ranging from .12 to .47. Although inpatient subjects showed an increase of reliability from the earlier study ($k=.53$), the control group was low ($k=.38$).

Some studies comparing the Structured Clinical Interview to other assessment instruments have not given strong evidence of reliability. Kendler and Roy (1995) compared family history interview to the SCID to predict depression and reported only modest agreement (kappa correlate range = .09-.29, mean kappa = .24). The authors stated that SCID overpredicted depression, probably due to inability to factor in pertinent individual data, and suggested multimethod approaches to diagnosis. Efficiency comparison with the Minnesota Multiphasic Personality Inventory, MMPI, and to the Millon Multiaxial Inventory, MCMI, (Hillis, 1995) found that the SCID correlated with the MCMI on 6 of 11 categories and with the MMPI on 8, while the MCMI correlated with the MMPI on 10 of 11. The MCMI was more sensitive and the MMPI more specific than the SCID. The SCID was reported to have higher correlation coefficient values compared to self report than did the MCMI for personality disorder occurring with eating disordered clients (Kennedy, Katz, & Rockert, 1995) with correlation ranges from .09 to .59. When compared to Axis II diagnoses given by clinicians, the SCID was reported to have poor reliability (Steiner, Tebes, Sledge, & Walker, 1995). Overall mean kappas equaled .25, with kappas ranging from .0 to .55. Jacobsen, Perry and Frances (1995) compared the SCID to the Personality Disorder Examination (PDE) with HIV positive clients and reported that the SCID overdiagnosed clients. The SCID diagnosed three times as many personality disorders as did the PDE and four times as many as are reported in literature to naturally occur with HIV positive individuals. In a comparison of MMPI and SCID assessment of diagnosed histrionic personality disorders, the SCID failed to identify 57 of 82 cases (Schotte, de Doncker, Maes, & Cluydts, 1993). Authors stated that the SCID-HST scale measures intraversion versus extraversion.

One study (Albanese, Bartel, Bruno, & Marcy, 1994) gave evidence that the SCID was significantly more sensitive than urinalysis at identifying

coexisting substance abuse disorders in psychotic patients. The results are not surprising, considering that urinalysis can only detect the drug present in the body at the moment, while psychodiagnostic instruments are designed to discover behavioral signs and symptoms beyond the moment. However, since an n of 178 subjects for the Albanese and associates study recorded six subjects who refused urinalysis, 21 who could not take the SCID due to psychotic episode interference and 62 subjects who were transferred before a urinalysis could be taken, methodology was weak.

Another study (Clarkin, Hull, Cantor, & Sanderson, 1993) reported high correlations between the SCID and the NEO-Personality Inventory (NEO-PI) for rating coexisting Axis II diagnoses in identified borderline personalities and found that they substantially agreed on neuroticism ($p < .01$) and agreeableness ($p < .01$). A study of unipolar depression in male alcoholics (Elwood, 1993) using SCID and MMPI reported that the SCID was sensitive to the coexisting disorder, while the MMPI was not. Perhaps the most consistent research discovery was that the SCID was able to differentiate between schizophrenia and dissociative identity disorder based on five specific dissociative symptoms (Steinberg, Cicchetti, Buchanan, & Rakfeld, 1994). Finally, a study of obsessive-compulsive personality disordered patients (Schotte, de Doncker, Maes, & Cluydts, 1991) reported that while the SCID was able to identify 24 cases of obsessive-compulsive disorder the MMPI found no significant difference between the obsessive-compulsive group and controls.

Results of reliability tests for the Structured Clinical Interview have been mixed, but the consistency of low kappa scores in field trials have lead Fisher and Martin (1981) and Saigh (1992) to state that although structured interviews have been used predominantly for research reliability has never been established. It seems, then, that the instrument itself may significantly affect outcome of studies of the DSM itself, yet the instrument was used in field studies of the DSM-IV (Widiger, et. al., 1991).

The American Psychiatric Association has not relied solely on the Structured Clinical Interview to validate the DSM-IV. There have been efforts to improve reliability and validity of the DSM as a psychodiagnostic instrument

throughout its history (Grob, 1991). Each revision has added more empirical methodology, and each revision has reflected more of the biomedical belief that valid taxonomy is possible. Follow-up and family studies, advisory committees made up of specialty researchers, field trials, and comparison with the ICD-9 and ICD-10 were established by the introduction of the DSM-III-R; and data reanalysis was added during development of the DSM-IV. In fact, committees debated many of the methodological suggestions of researchers and rejected any which did not seem to increase validity. For instance, the APA resisted the temptation to use meta-analysis, recognizing that the mixed criteria would not permit comparison of separate research (Widiger, et al., 1991).

However, since the SCID has been used during development of the DSM-IV, a review of results is in order. The research of Spitzer and Fleiss (1974) represents an example of DSM reliability studies before the advent of the Structured Clinical Interview. The authors, after a study of specific diagnostic categories -- the typical protocol for study of the DSM -- reported that a kappa of .61 was obtained for personality disorders, .53 for somatic disorders, and .49 for factitious disorders.

Results of field trials, using the SCID for the DSM-IV, have also proven to be mixed. A field trial of criteria for mood disorders (Keller, Klein, Hirschfield, Kocsis, McCullough, Miller, First, Holzer, Keitner, Marin, & Shea, 1995) resulted in an intrasite interrater reliability of .82 for the last month and .81 for lifetime, an intersite reliability of .44 for last month and .57 for lifetime, and a six month test-retest reliability of .56 for last month and .53 for lifetime. Recommendations to improve reliability which were not adopted upon publication emphasized cognitive and social/motivational symptoms.

Another DSM-IV field trial, this time of defiant and conduct disorders (Lahey, Applegate, Barkley, Garfinkel, McBurnett, Kerdyk, Greenhill, Hynd, Frick, Newcorn, Biederman, Ollendick, Hart, Perez, Waldman, & Shaffer, 1994), discovered that there was improvement in validation of criteria for diagnosis from DSM-III-R to DSM-IV, but that interrater reliability varied on oppositional defiant disorder from .25 to .60 and on test-retest reliability from .35 to .55, while

conduct disorder interrater reliability varied from .20 to .57 and test-retest reliability varied from .21 to .63.

DSM-IV validation was attempted in a study of criteria for bulimia nervosa (Thelen, Mintz, & Wal, 1996). The study measured sensitivity of criteria at .91, false-negative rate at .08, false-positive rate at .04, positive predictive value at .81 and negative predictive value at .98.

Generalization and simplification of criteria for diagnosis appears to have increased DSM-IV and validity and reliability to some degree, but reliability continues to be low, based on field trials using the SCID. Other validation procedures, however, may show evidence of low interrater and intersite reliability because professional orientation may affect agreement. In fact, Johnson, Horwath, & Weissman (1991) suggest that cultures, systems, professional orientation and training are all variables which affect reliability studies.

Comparison of SCID to the Problem-Solving Model

Although substantial literature exists describing and testing the two models independently, little exists which contrasts the two models, and then only indirectly. Research suggests that problem-solving models produce better results both in professional diagnosis and in development of proficient diagnosis among students. Yet, structured interview dominates research. Because both models are represented and endorsed in counseling literature, that a comparison of the two is in order.

Perhaps the reluctance of counselors to adopt the DSM is, to some degree, a reaction to rigid taxonomy, particularly considering that counseling has long accepted the idea of assessment of personality (Seligman, 1980), but not in conjunction with psychotherapy until recently. Counselor educators who have concentrated on study of diagnostics find nosology less uncomfortable within the framework provided by problem-solving diagnostic theory.

This reluctance may partially be responsible for an interesting but limited development in counselor education literature. Most counselor educators who have published regarding DSM instruction appear not to have read studies which attempt to test efficacy of use of the model. Fong (1993) and Seligman (1986) have written most extensively about teaching psychodiagnostics in counselor education, and neither have cited seminal research. Instead, both authors cited textbooks or general theoretical works to substantiate their models. This trend is typical. Cook and associates (1993) and Waldo and colleagues (1993) discuss bias of the DSM but offer recommendations based more on detachment from, rather than development of, a theoretical base for systemic use of psychodiagnostics.

The significance of the study presented in this thesis is to establish a comparison of the two models. Also, counselor education has an ethical responsibility: "Counselors who are responsible for developing, implementing, and supervising education programs are skilled as teachers and practitioners" (American Counseling Association, 1995, p. 14). Development of new areas of expertise requires development of new skills and new theoretic principles.

Review of Computer Assisted Learning Literature

Recent literature has suggested that computer applications enhance counselor education research by removing instructor bias from studies of learning theory (Anderson & Hornby, 1990; Berven, 1985; Chan, et al., 1993; Hornby & Anderson, 1994; Lambert, 1989; Lambert & Meier, 1992; Patterson & Yaffe, 1992; Sharf & Lucas, 1993; Shaw, 1979). Because the presentation by instructor adds potential confounding variables to a study of learning -- as opposed to instruction -- the introduction of computer assisted instruction, based on effective learning principles established in literature may control for these variables. So, a review of the literature pertaining to computer-assisted presentation of material is in order.

Attitudes of Students Toward Computer Assisted Training

Student attitude toward computer-assisted learning and resultant effects on learning has been prolific. In counseling related fields of instruction, subject opinion surveys have been used for studies to measure not only the effects of attitude on computer-assisted learning (Lambert & Lenthall, 1989; Rappaport, 1975; Jenkins & Danker, 1991; Poulin & Walter, 1990; Jones & McCormac, 1992; Abouserie, Moss, & Barasi, 1992), but also to measure the effectiveness of computer programs as teaching tools (Poulin & Walter, 1990; Goldberg-Wood & Middleman, 1987; Banyan & Stein, 1990; Thomas & Bostow; Maple, 1994; Patterson & Yaffe, 1993; Gentry, 1992; Polson, 1995; Lambert, 1989).

Jones and McCormac (1992) attempted to discover whether study subject opinion was a valid method of determining computer assisted learning effectiveness. The authors investigated nursing students rather than counselor trainees. Although generalized inference is not valid, it may be the only contribution to social sciences literature of its kind and so becomes important to discuss. Because of the dependence of many researchers upon the attitudes and opinions of subjects for the study of effectiveness of computer assisted learning and instruction, Jones and McCormac studied the effect of orientation on evaluation of computer programs. The authors discovered that when students were introduced to a particular program first they preferred it ($p < .001$) even though scores were higher using the other alternative program ($p < .01$). When both programs were introduced simultaneously, with care taken to avoid bias, ratings became mixed, with students rating each program higher in different areas. Because of the singular nature of this study, it is safer at this time to assume that computer assisted instruction evaluations should be based on subject performance measures rather than on attitude.

On the other hand, attitude may effect the application of computer assisted instruments in educational settings. Poulin and Walter (1990) reported that shy students preferred computer assisted instruction, but most students did not increase their comfort with computer use. Further study (Rappaport, 1975; Lambert & Lenthall, 1989) gave additional evidence that subject CAL anxiety

did not improve with use but produced convincing evidence that skills improved despite anxiety ($p < .001$ and $p < .005$, respectively). Another study (Abouserie, et al., 1992) found that males preferred ($p < .05$) and performed better ($p < .01$) with computer-assisted instruction than did females -- although performance results were suspect due to design errors. The authors also discovered that field-dependent learners became more dependent on CAL than did field-independent learners.

Skills Building Research

Several studies presented evidence that students understand CAL presented concepts (Brothen, 1995) and gain skills and experience through drill and practice (Alpert, 1986; Berven, 1985; Berven & Scofield, 1980; Chan, Rosen, Wong, & Kaplan, 1993; Gentry, 1992; Hayden, 1990; Janikowski, Berven, Meixelsperger, & Roedl, 1989; Lambert, 1989; Lambert & Meier, 1992; Patterson & Yaffe, 1993; Polson, 1995; Sexton-Radek, 1993).

In fact, studies have shown evidence that there is no significant difference between traditional instructional approaches and CAL. Patterson and Yaffe (1993) and Lambert (1989) found that subjects using CAL learned equally as well as did subjects using traditional learning sources. CAL software was used successfully for study of other variables related to counselor related education, including teaching empathy skills (Shaw, 1978; Kronk, 1985), interview skills (Sharf & Lucas, 1993), Rational Emotive response (Thomas & Bostow, 1991), goal focus skills (Maple, 1994), and assessment and diagnostic skills (Hayden, 1990; Patterson & Yaffe, 1993; de Mesquita, 1992; Polson, 1995). Studies of computer assisted learning programs in counseling related training appear to be as effective as other teaching techniques. The question then arises, "What does the literature suggest to be the optimum design for counselor training software?"

Components of Effective Computer Assisted Learning Software

Studies in counseling related education suggest that presentation of CAL can be enhanced, as can learning, if particular formats are applied to development of the CAL program. Programs that are comprehensive to the subject (Sexton-Radek, 1993; Brothen, 1995; Alpert, 1986; Maple, 1994; Patterson & Yaffe, 1993), are easy to learn and use (Jones & McCormac, 1992; Kronk, 1985; Patterson & Yaffe, 1993), interactive (Thomas & Bostow, 1991; Kronk, 1985; Jenkins & Dankert, 1981; Sharf & Lucas, 1993; Rappaport, 1975; Shaw, 1975; Maple, 1994; Hayden, 1990; Patterson & Yaffe, 1993; de Mesquita, 1992; Gentry, 1992; Polson, 1995; Lambert & Meier, 1992; Janikowski, Berven, Meixelsperger, & Roedl, 1989; Berven & Scofield, 1980; Chan, et al., 1993), nonsequential (Jenkins & Dankert, 1981; Banyan & Stein, 1990; Sharf & Lucas, 1993; Patterson & Yaffe, 1993; Polson, 1995), and that provide immediate feedback (Rappaport, 1975; Shaw, 1975; Hayden, 1990; Patterson & Yaffe, 1993; Janikowski, et. al., 1989; Berven & Scofield, 1980; Chan, et al., 1993) result in better performance and less anxiety on the part of test subjects. Also, Banyan and Stein (1990) as well as Gentry (1992) discovered that voice over text increased memory recall ($p < .002$). In other words, learning style preferences (Pask, 1976) may be more easily matched utilizing a computer program that can be adapted to these student preferences than can a structured classroom setting.

The evidence is convincing that CAL is capable of not only teaching linear and experiential information to students of counseling, but that CAL software allows for planned, triangulated presentation of material removing significant bias from research. It is logical, then, that CAL be used to inform and to research information integration for counselor educators.

Computer assisted learning also appears to have the capability, if designed according to the specifications discovered in research, to control for instructor bias and for potential learning style variables in research. Use of computer assisted instruction for diagnostic learning research may, consequently, decrease potentially confounding student variables. Inclusion of

these variables in computer assisted research is appropriate to gather evidence toward consideration of that potential control.

Summary

In essence, decision tree has become the most widely used psychodiagnostic method, but without significant evidence that the model is reliable. It has also received endorsement as an instructional philosophy from the ranks of counselor education, despite no evidence of effectiveness as a learning tool.

Studies of the problem-solving model have produced evidence, particularly in medical research, though a few counseling studies have found concurrent data, that diagnostic proficiency is related to the ability to generate multiple hypotheses throughout the diagnostic interview. What has not been accomplished is the comparison of the two models directly for purely psychodiagnostic purposes, nor has there been an attempt to study how early introduction of the philosophies affects learning.

Further, research has established that computer assisted instruction is an adequate if not potentially superior instrument for research in counselor education. Besides controlling for instructor bias, CAI has the adaptive capability to match student learning preferences, reducing the number of potentially confounding variables inherent to educational research. However, several other variables have been discovered by CAL studies which should be considered during the course of any further study. Particularly, CAI studies have suggested that results may be confounded by attitude (Jones & McCormac, 1992), computer skills (Rappaport, 1975), gender and preferred learning styles (Abouserie, 1992), and experience related variables (Chan, et al., 1993). So, it appears that while CAI inclusion in psychodiagnostic research may control for some confounding educational variables, other variables need to be considered in research design and analysis.

CHAPTER III DESIGN AND METHODOLOGY

Introduction

Chapter three will state the assumptions regarding psychodiagnosis, explain the purpose of the study, list the research questions, and examine the methods used to compare the outcome of students learning two protocols of diagnosis while studying DSM structure, Structured Clinical Interview and problem-solving. Methodology, instrumentation, sample population, design, data collection and analysis, and restatement of hypotheses will be presented in this chapter.

Assumptions Regarding Psychodiagnosis

1. The DSM will become the standard psychodiagnostic assessment instrument for counseling.
2. Comprehensive instruction of each of the two diagnostic models is not necessary to compare the influence of early orientation to the models on psychodiagnostic learning.
3. Case study is an appropriate assessment procedure for measuring diagnostic skills.

Purpose of the Study

This study was undertaken to examine the possible interconnection between early orientation to a model for psychological diagnosis, decision tree or problem-solving model, and the effect of early orientation on proficiency of diagnosis during diagnostic training of counseling students. The purpose of this study was to compare the ability of inexperienced diagnosticians to diagnose psychological syndromes using the DSM.

Research Questions

Research Question 1

Is there a relationship between diagnostic education and diagnostic proficiency?

Null Hypothesis 1. There is no significant difference between psychodiagnostic proficiency of students before and after receiving training in DSM taxonomy.

Research Question 2

Is the method for introducing single hypothesis decision tree or problem-solving models of diagnosis before instruction in DSM taxonomy related to resultant proficiency of diagnosis?

Null Hypothesis 2. There is no significant difference between psychodiagnostic proficiency of students who learn DSM taxonomy after receiving prescribed training in Structured Clinical Interview (single hypothesis decision tree) or problem-solving technique (multiple, competing hypotheses).

Research Methods and Procedures

An experimental method was used to assess possible change in skills as a function of exposure. First participants were exposed to a systematic model for diagnostic decision making and then to a DSM Axis II computer assisted learning program. With the research questions stated as reference points, an appropriate experimental design and measurement instruments were chosen. A pretest was administered to explain preknowledge, and a posttest was administered to determine the main effect of the treatments.

Sample Population

The sample population for this study consisted of volunteers solicited from two separate sites, Oregon State University (OSU) and Southern Oregon State College (SOSC). Sites were chosen because the Counseling Program is a CACREP approved program and the Psychology Program at SOSC is a CACREP equivalent program. Both institutions provide classes which meet criteria for NBCC certification and CPC licensure for Master's level graduates.

As incentive to participate, students received credit toward grades in classes from which they were solicited, and their names were entered in a drawing for a gift certificate at the local college book store for those who finished the experiment. Also, as a learning experience, purpose of the study and preliminary results were presented to participants after initial analysis was completed.

The sample at Site 1, OSU, consisted of 13 volunteers solicited from Master's level counseling students enrolled in the Winter Term, 1997, Special Topic: DSM-IV Diagnosis (Coun 580) course. This sample had experienced five weeks of training in the DSM including structure of DSM, criteria for diagnosis, and three weeks of exposure to Axis I major psychiatric disorders.

The sample at Site 2, SOSC, consisted of 49 students solicited from the undergraduate, postbaccalaureate, and graduate students enrolled in the Psychology Program who had not yet taken Abnormal Psychology I (Psy 479/579), and who had substantially completed Introduction to Psychology (Psy 102). These criteria assured some exposure to basic concepts of abnormal human behavior while controlling for preexposure to DSM, Axis II categories and preorientation to models of diagnostic decision making. Originally, a criterion was set, that students from SOSC would be upper divisional or postbaccalaureate; however, a population of underclasspersons who had completed Psy 101 and the portion of Psy 102 pertaining to abnormal human behavior were made available for the study, so they were included.

Some fundamental differences existed between sample populations. Students enrolled in the Site 1, OSU program were predominately oriented toward school counseling, either having been teachers, misassigned (uncertified) school counselors, or having worked minimally in agencies linked to schools. All students enrolled in Coun 580 had completed a counseling practicum course series and some had been placed in internship sites, predominately in K-12 school settings. Most had little exposure to mental health orientations and all intended to become school counselors or agency counselors linked to school programs. Requirements for admission to the Site 1 graduate program include achievement of an undergraduate 3.0 grade point average. Students may be admitted to the program without meeting the requirement only if faculty advocate for admission. The mean undergraduate grade point average for OSU participants was 3.35.

Participants from Site 2, SOSC, were predominately undergraduates. Participant grade point averages were not controlled and so were documented and analyzed post hoc. The mean undergraduate grade point average was 3.25 with three participants not reporting their grade point averages. Career goals varied more than those at Site 1, but included school counseling, agency counseling and, more often, mental health counseling. Freshmen and some of the sophomores in the sample population were considered pre-psychology majors and may not have been committed to a career in psychology. Few Site 2 students had been placed into internships or practica, nor had obtained any experience professionally. A criterion for participation in the study included that participants from Site 2 complete the reading of introductory material from the Diagnostic and Statistical Manual, fourth edition (American Psychiatric Association, 1994) to assure exposure to basic information about psychopathology and structure of the DSM. Site 1 participants had been exposed to DSM structural information in Coun 580.

An n of 13 students at site 1 volunteered for the study, and an n of 49 undergraduate, post baccalaureate and graduate students volunteered at site 2. Since the literature review had suggested that age (Chan, et al., 1993) and

cultural origin (Cook, et al., 1993) might be significant factors in diagnostic learning, demographic information was gathered from volunteers as variables to be included in the statistical analysis. Participants from Site 1 ranged in age from 23 to 51 years of age with a mean age of 37.36 years, while participants from Site 2 ranged in age from 18 to 57 years of age with a mean age of 25.51 years. Counseling participants included 10 anglo-caucasians, one Hispanic, one African-American and one French-Canadian. Psychology participants included 46 anglo-caucasians, one Native American Indian, one Native American Pacific Islander, two Hispanics, and one Armenian.

Counseling participants reported a mean of 40.45 months of computer experience, while psychology participants reported a mean of 63.38 months of computer experience. Site 1 participant classroom exposure to abnormal psychology with a DSM emphasis ranged from 1 to 15 weeks with a mean of 5.8 weeks. Ten Site 1 participants had experienced 5 weeks of Coun 580, one had experienced 2 weeks and had withdrawn from the course, and 2 had experienced not only 5 weeks of Coun 580 but had also completed an undergraduate course in abnormal psychology. Professional experience using DSM ranged from 0 to 10 months with a mean of 1.18 months. Site 2 participants reported a range of classroom exposure to DSM from 0 to 12 weeks with a mean of .76 weeks, with none having had any professional experience. Table 1 describes the gender and education level of counseling participants, and Table 2 describes the gender and education level of psychology participants.

Initially, 13 participants volunteered at the university counselor training site, but there was a loss of two participants from group 2, both males. Both participants completed the pretest and part of the experiment but failed to complete the posttest, reporting they were ill. One stated that looking at the computer screen made him nauseated. Both were dropped from the study. The total n after the loss of the two university participants was 11 from Site 1 and 49 from Site 2, with an overall n of 60 study participants.

TABLE 1. Gender and Education Level of Site 1, Counseling Sample

Ed Level	<u>Gender</u>		Totals
	Male	Female	
Grad. Yr. 1	4	5	9
Grad. Yr. 2	2	2	4

TABLE 2. Gender and Education Level of Site 2, Psychology Sample

Ed Level	<u>Gender</u>		Totals
	Male	Female	
Freshman	5	9	14
Sophomore	2	7	9
Junior	3	10	13
Senior	4	7	11
Grad. Yr. 1	0	2	2

Research Design

The study was a controlled experiment. Each session was completed in a four hour time period to control for time and participant interaction.

Research Procedure

Research assistants were trained and conducted the entire experiment. Research assistants assigned to read instructions and proctor sessions, were given four hours of training including participant management, script practice,

computer software orientation, and interview training. Research assistants assigned only to interviews were given one hour of training including use of recording equipment, general cues training to recognize problem-solving versus decision tree diagnostic decision making, and script familiarity. All verbal interaction between research assistants and participants was scripted (See Appendix C).

All potential volunteer participants were asked to read the letter of invitation and the informed consent form (See Appendix A). Individuals who volunteered were given instructions for dates and locations of the experiment sessions. One experiment was performed at Site 1 and one administered at Site 2 during 4 sessions. Random group assignment of participants, using a randomization table procedure, was completed prior to the experiment dates. Experimental treatments were conducted and data gathered during winter term, 1997. Groups 1 and 2 were assigned from the Site 1 volunteers and Groups 3 and 4 from the Site 2 volunteers. Participants were identified by the last four numbers of their Social Security numbers throughout the study so that proctors and scorers were blind to the identity of participants.

Participants were introduced to the structure of the DSM before the experiment began. At Site 1, this structural information was presented by the professor during the introductory phases of the Counseling 580 course. Site 2 students were required to read the introduction to the Diagnostic and Statistical Manual, Fourth Edition (American Psychiatric Association, 1994, pp. xxi-xxv, 1-9) during the two weeks between registration and participation and to sign an affidavit of completion at the beginning of the experiment (See appendix A). The Site 1 university professor agreed to forego presentation of Axis II instruction until after the experiment was completed. Instructors at both institutions agreed not to reveal the specific intent of the experiment, which was introduced simply as a study of psychodiagnostic learning.

Upon arrival, participants were given the instructions for and administered the pretest (See Appendix C for scripts). A demographic questionnaire was given at the time of the pretest and with the computer assisted instructional program to determine age by years, gender, computer

experience by months, classroom diagnostic experience by weeks, professional diagnostic experience by months, education level by years, undergraduate grade point average and cultural origin. Methodological hypotheses, described in Chapter IV, were developed to analyze the several potentially explanatory variables.

Upon completion of the pretest, participants were instructed to find their group assignments from a list of Social Security code numbers and to proceed to the assigned rooms. Groups 1 and 3 were then presented with a 35 minute video (See Appendix F for video script) and a follow-along handout orienting them to the Structured Clinical Interview model (Williams, et al., 1992, pp. A.1-A.7, D.1-D.5, G.1-G.7). Groups 2 and 4 were presented with a 35 minute video and follow-along handout written by the study author (See Appendix E for handout), orienting them to the problem-solving model of diagnostic decision making (Elstein, et al., 1978). Videotapes were produced for the experiment in lecture format (See Appendix F) and presented to three college instructors who verified that the video presentations were not biased toward one of the models.

Following the treatment session, each group was escorted individually to a computer laboratory equipped with Power Macintosh 7600 series computers, read instructions and orientation material for use of the Hypercard computer assisted learning program, Hyperaxis II (Patterson, 1990). Each was given up to two hours to complete study of DSM Axis II, determined ample time to review the entire program (Patterson and Yaffe, 1993).

The program, Hyperaxis II (Patterson, 1990), presents a "card" of user instructions, followed by a demographic questionnaire, and a "map" of the program. The nonsequential format provides the reader with: definitions of personality disorder and of each category of personality disorder, diagnostic criteria for each cluster and diagnosis, a clinical presentation of each diagnostic category along with the differential diagnosis, a case example of each diagnostic category, and finally, a practice skills test which gives immediate feedback and has access to previously presented information for review by the student. The program offers a straightforward presentation of each category without commentary regarding how to use diagnostic categories, controlling for

instructor bias. Each participant receives a diskette containing the program. This diskette records demographic information: the percentage of each diagnostic category reviewed, time spent on each section of the learning program, and scores from the rehearsal test. The program presents the learner with a comprehensive, interactive, and nonsequential format, easily learned within minutes.

Participants were instructed to return to the room used for the pretest whenever they thought they had achieved competence with Axis II diagnosis. Because of copyright restrictions (D. A. Patterson, personal communication, April 9, 1996) Hyperaxis II diskettes were collected by the proctor as subjects left the computer laboratory, at which time a posttest was administered. Pretests and posttests were scored by the study author.

Subsequent to the posttest, an interview was conducted by research assistants (See Appendix D). Interviews were planned to determine if participants had used the treatment orientations to which they were exposed as a treatment condition during posttest diagnostic decision making. The interviews were structured to assure consistency and decrease bias. Since no data exists to suggest how much exposure to the treatment conditions is required for orientation, this qualitative addition to the study was appropriate.

Test Instrument

Because the computer assisted learning program (Patterson, 1990) limited instruction to personality disorders, the test instrument was designed to assess participant proficiency after study of DSM Axis II diagnostic criteria. The traditional assessment instrument has been a pretest and posttest method utilizing randomly assigned case studies (Berven & Scofield, 1980; Chan, et al., 1993; Clay, et al.; Friedman, et al, 1995; Hayden, 1990; Janikowski, et al., 1989; Lambert & Meier, 1992; Patterson & Yaffe, 1993; Shamian, 1991). Consequently, the established assessment instrument was adopted for this study.

The test instrument was constructed from items drawn from the DSM-III-R Case Book (Spitzer, Gibbon, Skodol, Williams, & First, 1988, pp. 32-35, 37-40, 42-43, 53-54, 80-81, 107-110, 123-124, 163, 172, 175-176, 182-183, 197-198, 207-208, 233, 249-250, 424-425, 450-451). The third edition of the DSM (See appendix B) was chosen because the computer assisted learning program has not yet been upgraded to match the fourth edition (Spitzer, Gibbon, Skodol, Williams, & First, 1994); however, all cases chosen appear in the fourth edition. Eighteen cases were chosen for the testing instrument, 16 presenting the twelve personality disorders from the DSM-III-R, one presenting no diagnosis and one presenting an Axis I case study which has signs or symptoms in common with personality disorders.

Cases chosen for the study test instrument met the following criteria:

- (a) 14 cases that they represent each of the personality disorders, including Personality Disorders Not Otherwise Stated, that one case represent a diagnosis from Axis I and that one case contain no diagnosis, to control for the confounding factor of elimination,
- (b) that test items must not appear on the computer assisted learning program practice test,
- (c) that test items simulated real life scenarios: items varied from one to four diagnoses, some diagnoses being full disorders and others being features of but not fully meeting criteria of disorders.

Content validity of the cases was assumed, since the authors of the DSM-III-R Case Book are also the authors of the DSM-III-R (Borg & Gall, 1989, pp. 250-252). However, a committee of three mental health professionals -- chosen because an area of professional concentration on psychological diagnosis -- was also consulted to establish case study construct validity (Borg & Gall, 1989, pp. 255-556) during formation of the pretests and posttests (See Appendix D). The consultants included John Gram, Psy.D., Bernie Legner, Ph.D., and Neil Williamson, M.D..

Each item in the pretests and posttests was a verbatim presentation from the DSM-III-R Case Book with diagnosis and case discussion removed. Permission was requested of the copyright retainer to publish the case studies.

Items were randomly assigned to pretests and posttests with the two nonpersonality disorder cases randomly assigned to the two tests separately. To control for instrument error, using a randomization table procedure, the random assignment of cases was repeated three times and the resultant pairs of pretests and posttests were distributed to participants evenly during the study. Test version 1 was taken by 21 participants, test version 2 by 19 participants, and test version 3 by 20 participants. Table 3 describes the items in the three pretest and posttest versions.

TABLE 3. Items of the Three Test Versions

Item	Test Version 1		Test Version 2		Test Version 3	
	Pretest	Posttest	Pretest	Posttest	Pretest	Posttest
A	X			X		X
B	X			X	X	
C		X		X		X
D		X		X	X	
E	X			X		X
F		X	X		X	
G	X		X			X
H		X	X			X
I	X		X		X	
J		X		X		X
K		X	X		X	
L		X		X	X	
M	X		X			X
N	X		X			X
O		X	X		X	
P		X		X	X	
Q	X		X			X
R	X			X	X	

Scoring of pretests and posttests was based on correct identification of diagnosis in the following manner. Correct identification of Axis II disorders was considered to be use of the proper DSM terminology, i.e., "paranoid," "paranoid personality," or "paranoid personality disorder." An answer sheet was provided for each case study. The answer sheet was designed for recall responses to insure that participants had learned material rather than potentially guessed correct responses from a multiple choice list. Each answer sheet provided separate spaces for an Axis II full diagnosis, Axis II traits which did not meet adequate criteria for full diagnosis, Axis I full diagnosis, and Axis I traits. Participants were expected to identify Axis II diagnoses only.

Although the instrument used to assess diagnostic proficiency of participants has been accepted practice, the potential for test error has not been explored. Consequently, pretests and posttests were scored twice, using two separate scaling systems and each one was analyzed separately.

First, pretests and posttests were scored using a three point per item system. One point was awarded for identification of the primary diagnosis. A second point was awarded for identification of secondary diagnosis, secondary traits or recognition of no secondary traits. A third point was awarded for properly identifying whether the diagnoses met criteria for full diagnosis or were traits which did not meet criteria for full diagnosis. No score was given for recognition of Axis I diagnoses or features. An overall score of 54 points was available, 27 points for the pretest and 27 points for the posttest.

The second scoring scale was a one point per item. Each item was scored on the basis of correct primary diagnosis only. Maximum points on each of the six tests was 9.

Analysis

All analyses were carried out using the SPSS 4.0 software package. Several potentially explanatory variables were anticipated and included in the analysis -- site, test versions (instrument error), participant age, gender, computer experience, classroom exposure to psychodiagnosis, professional

experience with psychodiagnosis, education level, undergraduate GPA, cultural origin, apparent participant study preferences and participant attitudes toward CAI. The potential significance of these variables has been noted (de Mesquita, 1992; Patterson & Yaffe, 1993), but inclinations of the variables have been inconsistent.

Due to small sample size, $\alpha=.10$ would compensate for potential loss of power, particularly in light of the need to compare two samples with an unequal n . Since comparisons of the two groups would be most congruent using the same statistical analysis, and since precedent had been set during earlier research of psychodiagnostic learning (Elstein, et al., 1978), the choice of alpha was appropriate.

A research question was posed: "Did participants utilize the treatment condition (orientation to either problem-solving or binary decision tree decision-making model) during the posttest?" Research assistant posttest interviews were transcribed for analysis. Transcribed results were analyzed first by this author and then triangulated with a mental health professional.

Consensus was reached between raters on analysis procedures as follows. The mental health professional was briefed regarding the differences between problem-solving and decision tree models of decision-making. Discussions were held to attain as high a level of agreement as possible on what constituted multiple hypothesis versus decision tree decision-making methodology. The mental health professional was not briefed regarding which orientation each participant had received. Each rater independently read the transcripts from interviews performed at Site 1, attempting to determine which decision-making model each participant had utilized. After transcribed interviews were analyzed, results were compared and disagreements discussed to determine areas of confusion or disagreement, particularly regarding interviews which were too short to give clear indications or which gave only subtle clues. Each rater then reread Site 1 transcripts independently to reevaluate classifications. Disagreements were again discussed to reach as high a rate of agreement as possible regarding subtle cues of decision-making model use by participants. Then, Site 2 information was analyzed

independently by each rater. These results were then statistically analyzed for agreement using a signed test to determine if agreement reached statistical significance.

Summary

A pretest, posttest and experimental treatment were administered by trained personnel. A qualitative interview after the posttest was conducted to determine if participants had used the treatment condition, which was video generated to control for instructor bias. A computer assisted learning lab experience followed the treatment condition. Several potential explanatory variables were considered and demographic information gathered for analysis to consider their effects on the major hypotheses. These variable were tested using a Pearson product-moment procedure.

CHAPTER IV FINDINGS AND DISCUSSION

Introduction

The primary purpose of this study was to examine the effects of orientation to either multiple hypothesis or single hypothesis diagnostic decision-making on DSM diagnostic proficiency. Hypotheses were developed to direct systematic testing of the experimental treatment variables. Additionally, the study examined the possible relationship between site, age, gender, computer experience, education level, classroom exposure to DSM information and diagnostic decision making orientation, undergraduate grade point average, professional experience with psychodiagnostics, and cultural origin on development of diagnostic skill.

Secondarily, this study examined apparent study organization preferences of participants and attitudes toward computer assisted learning with the Hyperaxis II program. Because this software package records demographics including attitude, participant concentration within the program on location (including practice test and case studies), as well as time devoted to the learning package, ample data is available for a post hoc analysis of study organization preferences and attitudes. In this chapter, hypotheses will be presented, followed by results of analyses and appropriate statements of support or nonsupport of hypotheses.

Because of the variations of the generally accepted instrument for assessment of psychological diagnostic proficiency discovered in this study, the instrument were analyzed using two scoring scales to determine its internal consistency.

Preliminary Analysis of Instrument

A Student's t-test was performed to determine if there were differences between sites using both the three point scale and the one point scale. Each test version was analyzed separately to discover existing patterns.

Using the three point scale, pretest version 1 mean for Site 1 was 7.75 (sd=1.71) and for Site 2 was 3.88 (sd=2.69). Evidence existed of a difference between sites on pretest version 1 ($t_{19}=2.72$, $p=.014$). The version 1 posttest mean for Site 1 was 14.5 (sd=1.0) and for Site 2 was 7.82 (sd=4.22). Evidence existed of a difference between sites on posttest version 1 ($t_{19}=3.09$, $p<.00$).

The pretest version 2 mean for Site 1 was 3.0 (sd=2.0) and for Site 2 was 3.47 (sd=2.06). There was no evidence of a difference between sites on pretest version 2 ($t_{18}=-.37$, $p=.719$). The version 2 posttest mean for Site 1 was 6.0 (sd=1.0) and for Site 2 was 7.24 (sd=3.34). There was no evidence of a difference between sites on posttest version 2 ($t_{19}=-.62$, $p=.54$).

The pretest version 3 mean for Site 1 was 1.5 (sd=1.73) and for Site 2 was 2.33 (sd=2.22). There was no evidence of a difference between sites on pretest version 3 ($t_{17}=-.69$, $p=.5$). The version 3 posttest mean for Site 1 was 12.0 (sd=6.68) and for Site 2 was 13.67 (sd=5.38). There was no evidence of a difference between sites on posttest version 3 ($t_{17}=-.53$, $p=.61$).

An analysis of sites using the one point scale was then performed. The pretest version 1 mean for Site 1 was 5.5 (sd=1.0) and for Site 2 was 2.59 (sd=1.77). Evidence existed of a difference between sites on pretest version 1 ($t_{19}=3.13$, $p=.005$). The version 1 posttest mean for Site 1 was 5.25 (sd=.5) and for Site 2 was 3.71 (sd=1.72). Evidence existed of a difference between sites on posttest version 1 ($t_{19}=1.74$, $p=.097$).

The pretest version 2 mean for Site 1 was 1.67 (sd=1.16) and for Site 2 was 2.18 (sd=1.33). There was no evidence of a difference between sites on pretest version 2 ($t_{18}=-.62$, $p=.54$). The version 2 posttest mean for Site 1 was 4.0 (sd=0.0) and for Site 2 was 4.29 (sd=2.31). One sample had no variance so a t-test could not be performed for difference on the version 2 posttest.

The pretest version 3 mean for Site 1 was 1.75 (sd=1.71) and for Site 2 was 3.07 (sd=1.62). There was no evidence of a difference between sites on pretest version 3 ($t_{17}=-.34$, $p=.74$). The version 3 posttest mean for Site 1 was 4.5 (sd=2.38) and for Site 2 was 5.67 (sd=1.72). There was no evidence of a difference between sites on posttest version 3 ($t_{17}=-1.12$, $p=.28$).

Although test version 1 gave evidence of site difference, a pattern of nonsignificance occurred, so site samples were combined and a new t-test was performed. After combining test versions the overall pretest mean for Site 1 was 4.18 (sd=3.31) and for Site 2 was 3.26 (sd=2.39) using the three point scale. There was no evidence of difference between sites ($t_{58}=1.07$, $p=.29$). The posttest mean for Site 1 was 11.27 (sd=5.16) and for Site 2 was 9.41 (sd=5.13). There was no evidence of difference between sites ($t_{58}=1.09$, $p=.28$).

Using the one point scale the mean pretest score for Site 1 was 3.09 (sd=2.26) and for Site 2 was 2.29 (sd=1.57). There was no evidence of difference between sites ($t_{58}=1.41$, $p=.16$). The posttest mean for Site 1 was 4.64 (sd=1.43) and for Site 2 was 4.51 (sd=2.07). There was no evidence of difference between sites ($t_{58}=.19$, $p=.84$).

Tests for site differences suggested a pattern of no difference between sites. The test version 1 significance suggested that a problem existed with the instrument itself, so site samples were combined for further analyses. Further tests were performed to discover the instrument variations. Analysis of the instrument was performed to establish reliability before tests of hypotheses.

Reliability Analysis of Three Point Scale

Pearson Product-moment correlations among all items for the three point scale pretest items and posttest items are provided in Appendix G. Reporting the results of SPSS's reliability procedure with the Chronbach Alpha option, Table 4 describes the reliability of items in pretest version 1. Table 5 describes the reliability of items in posttest version 1. Table 6 describes the reliability of items in pretest version 2. Table 7 describes the reliability of items in posttest version 2. Table 8 describes the reliability of items in pretest version 3. Table 9 describes the reliability of items in posttest version 3.

TABLE 4. Reliability of Items in Pretest Version 1 Using 3 Point Scale

Item	Total Correlation	Alpha If Item Deleted
A	-.0048	.4606
B	.3529	.2979
E	.5043	.2135
G	.3698	.3501
I	-.0115	.4360
M	-.2818	.4615
N	.1485	.4046
Q	.1407	.4750
R	.2694	.3834

Alpha=.42, $F_{8,152}=10.17$, $p<.00$

TABLE 5. Reliability of Items in Posttest Version 1 Using 3 Point Scale

Item	Total Correlation	Alpha If Item Deleted
C	.3457	.5095
D	.3709	.5122
F	.0204	.5962
H	.1720	.5732
J	.3570	.5148
K	.3486	.5075
L	.4065	.5014
O	.3459	.5085
P	-.0550	.5856

Alpha=.57, $F_{8,152}=8.12$, $p<.00$

TABLE 6. Reliability of Items in Pretest Version 2 Using 3 Point Scale

Item	Total Correlation	Alpha If Item Deleted
F	.0630	-.3034
G	-.1966	-.1217
H	-.2439	.0265
I	-.0410	-.2239
K	.2135	-.3337
M	.0630	-.3034
N	-.0516	-.2143
O	.1338	-.4190
Q	-.2131	.1079

Alpha=.23, $F_{8,152}=5.84$, $p<.00$

TABLE 7. Reliability of Items in Posttest Version 2 Using 3 Point Scale

Item	Total Correlation	Alpha If Item Deleted
A	.1108	.3644
B	.1993	.3216
C	.1813	.3323
D	.1250	.3579
E	-.0461	.4906
J	.2510	.3018
L	.3424	.2877
P	.2481	.3409
R	.3450	.3313

Alpha=.38, $F_{8,152}=8.45$, $p<.00$

TABLE 8. Reliability of Items in Pretest Version 3 Using 3 Point Scale

Item	Total Correlation	Alpha If Item Deleted
B	.6260	.0591
D	-.2519	.5607
F	.4084	.3031
I	.0926	.4886
K	.6020	.2520
L	-.2519	.5607
O	.0000	.0000
P	.0000	.0000
R	.0000	.0000

Alpha=.48, $F_{5,90}=4.92$, $p=.001$

TABLE 9. Reliability of Items in Posttest Version 3 Using 3 Point Scale

Item	Total Correlation	Alpha If Item Deleted
A	.7031	.6850
C	.7572	.6666
E	.3874	.7273
G	.6583	.6779
H	.6531	.6818
J	.0527	.7797
M	.5912	.6972
N	.5912	.6972
Q	-.1147	.8206

Alpha=.74, $F_{8,144}=7.98$, $p<.00$

Reliability Analysis of One Point Scale

Pearson Product-moment correlations for pretest items and posttest items are provided in Appendix H. Table 10 describes the reliability of items in pretest version 1. Table 11 describes the reliability of items in posttest version 1. Table 12 describes the reliability of items in pretest version 2. Table 13 describes the reliability of items in posttest version 2. Table 14 describes the reliability of items in pretest version 3. Table 15 describes the reliability of items in posttest version 3.

TABLE 10. Reliability of Items in Pretest Version 1 Using 1 Point Scale

Item	Total Correlation	Alpha If Item Deleted
A	.0194	.4837
B	.2422	.3832
E	.6547	.1959
G	.4791	.2671
I	.0734	.4431
M	-.2327	.4962
N	.0369	.4665
Q	.0131	.4849
R	.3241	.3566

Alpha=.44, $F_{8,152}=3.86$, $p<.00$

TABLE 11. Reliability of Items in Posttest Version 1 Using 1 Point Scale

Item	Total Correlation	Alpha If Item Deleted
C	.3002	.4062
D	.1798	.4531
F	.0577	.4961
H	.0179	.5036
J	.2656	.4194
K	.2211	.4376
L	.3295	.3929
O	.4233	.3621
P	.0194	.4993

Alpha=.47, $F_{8,160}=3.72$, $p=.002$

TABLE 12. Reliability of Items in Pretest Version 2 Using 1 Point Scale

Item	Total Correlation	Alpha If Item Deleted
F	-.0614	.1534
G	.0187	.1008
H	.1043	.0233
I	.1811	.0355
K	.4564	-.1524
M	.2750	-.0960
N	.0424	.0822
O	-.2018	.2254
Q	-.2245	.2984

Alpha=.10, $F_{8,152}=2.39$, $p=.019$

TABLE 13. Reliability of Items in Posttest Version 2 Using 1 Point Scale

Item	Total Correlation	Alpha If Item Deleted
A	.1448	.6649
B	.3303	.6315
C	.4423	.6032
D	.1558	.6746
E	.2604	.6489
J	.3426	.6285
L	.4938	.5895
P	.5166	.5984
R	.4165	.6225

Alpha=.66, $F_{8,152}=6.89$, $p<.00$

TABLE 14. Reliability of Items in Pretest Version 3 Using 1 Point Scale

Item	Total Correlation	Alpha If Item Deleted
B	.4503	-.1118
D	-.2526	.4432
F	.3944	.0000
I	-.1140	.4318
K	.6005	-.1827
L	-.2526	.4432
O	.0000	.0000
P	.0000	.0000
R	.0000	.0000

Alpha=.29, $F_{5,90}=2.79$, $p=.02$

TABLE 15. Reliability of Items in Posttest Version 3 Using 1 Point Scale

Item	Total Correlation	Alpha If Item Deleted
A	.4239	.6578
C	.6463	.6281
E	.3150	.6811
G	.3467	.6731
H	.5426	.6576
J	.3487	.6728
M	.5633	.6194
N	.5862	.6133
Q	-.1232	.7655

Alpha=.69, $F_{8,144}=4.89$, $p<.00$

Inter-item correlations using both scoring scales indicate reliability varied no matter how tests were scored. Using the three point scale, correlations ranged from $-.28$ to $.76$ with only one test version (the version 3 posttest) giving substantial evidence of internal consistency. This test version had an alpha of $.74$, which increased to $.82$ if item Q was removed. Using the one point scale correlations ranged from $-.25$ to $.65$. Again, only the version 3 posttest manifested substantial internal consistency. This test version had an alpha of $.69$, which increased to $.77$ if item Q was removed. All reliability coefficients were significant. Other than posttest version 3 removal of one or more items would not appreciably affect the coefficient alpha.

Reliability variations suggested that validity problems might exist with the instrument. Consequently one way analyses of variance was used to analyze pretests and posttests for equivalence of form. Table 16 describes the 3 point scale pretest mean scores by test versions. Table 17 describes the 3 point scale posttest mean scores by test versions.

TABLE 16. Pretest 3 Point Scale Mean Scores by Test Versions

Test	<i>n</i>	Mean	S.D.	95% C.I. for Mean
1	21	4.619	2.94	3.28 to 5.96
2	20	3.40	2.01	2.46 to 4.34
3	19	2.16	2.16	1.14 to 3.18
Total	60	3.43	2.57	2.77 to 4.10

TABLE 17. Posttest 3 Point Scale Mean Scores by Test Versions

Test	<i>n</i>	Mean	S.D.	95% C.I. for Mean
1	21	9.10	4.65	6.98 to 11.21
2	20	7.05	3.12	5.59 to 8.51
3	19	13.32	5.52	10.66 to 15.98
Total	60	9.82		
	5.14	8.42 to 11.08		

Table 18 describes the 1 point scale pretest mean scores by test versions. Table 19 describes the 1 point scale posttest mean scores by test versions.

TABLE 18. Pretest Mean Scores By Test Version Using One Point Scale

Test version	Mean Score	S. D.	95% C.I.	Participant <i>n</i>
1	3.1429	2.0071	2.2292 to 4.0565	21
2	2.1000	1.2937	1.4591 to 2.7055	20
3	2.0000	1.5986	1.2295 to 2.7705	19
Total	2.4333	1.7209	1.9888 to 2.8779	60

TABLE 19 Posttest Mean Scores By Test Version Using One Point Scale

Test version	Mean Score	S. D.	C.I.	Participant <i>n</i>
1	4.0000	1.6733	3.2383 to 4.7617	21
2	4.2500	2.1244	3.2557 to 5.2443	20
3	5.4211	1.8654	4.5220 to 6.3201	19
Total	4.5333	1.9612	4.0267 to 5.0400	60

Using the three point scale, the analysis of variance for differences of pretest scores by test version supported the idea that test versions were not equivalent forms ($F_{2,57}=5.22$, $p=.008$). The difference occurred between test version 1 and test version 3. The analysis of variance for differences of posttest scores by test version also supported the idea that test versions were not equivalent forms ($F_{2,57}=9.7$, $p=.0002$). The differences occurred between test versions 1 and 2 and version 3, but not between versions 1 and 2.

Using the one point scale, the analysis of variance for differences of pretest scores by test version supported the idea that test versions were not equivalent forms ($F_{2,57}=2.94$, $p=.06$). Differences occurred between tests 2 and 3 and test 1, but not between tests 2 and 3. The analysis of variance for

differences in posttest scores by test version also supported the idea that test versions were not equivalent forms ($F_{2,57}=3.14$, $p=.05$).

Two instrument patterns had developed which suggested concern. The first was that there was little internal consistency in most of the test instruments no matter how they were scored. Only one of the six tests had evidence of reliability on both scalings. It appeared that outside of the potential for use of posttest version 3, combining results of the three instruments did not appear appropriate.

The second instrument pattern that developed was that there was little evidence of equivalence between forms. The assumption that random assignment of case studies would produce a valid instrument became suspect.

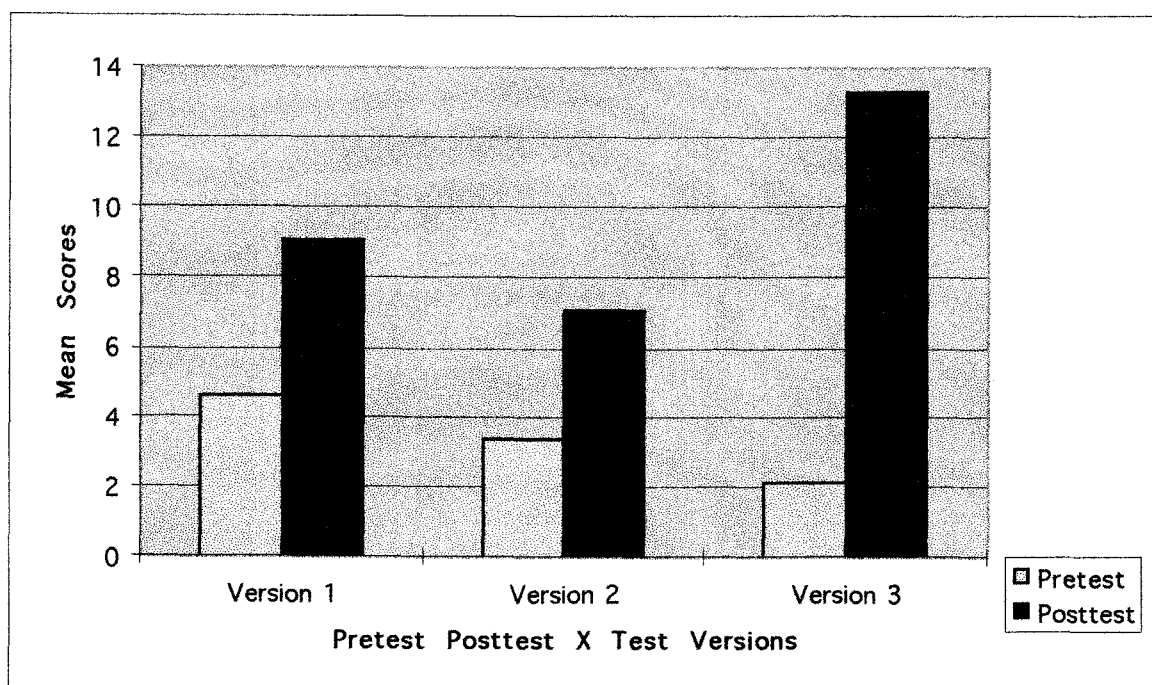
Although inference may not be assumed beyond the sample used in this study, the inconsistency of items within tests suggests that it is important to consider choice and placement of items when using case studies to assess diagnostic proficiency. Further, the patterns which developed between instruments suggested that further investigation was in order. It would be necessary to test hypotheses using separate analyses for each test version.

Hypothesis 1: Learning Effects

Hypothesis 1 stated that there is no significant difference between psychodiagnostic proficiency of students before and after receiving training in DSM taxonomy. Each test for the hypothesis was performed on both the three point scale and on the one point scale.

Using the three point scale, a one way analysis of variance indicated that there was a difference between test versions on both pretest ($F_{2,57}=5.22$, $p=.008$) and posttest ($F_{2,57}=9.7$, $p<.000$), so the tests for hypothesis were carried out by test version to determine existing patterns. Figure 1 illustrates the patterns of learning effects by test version.

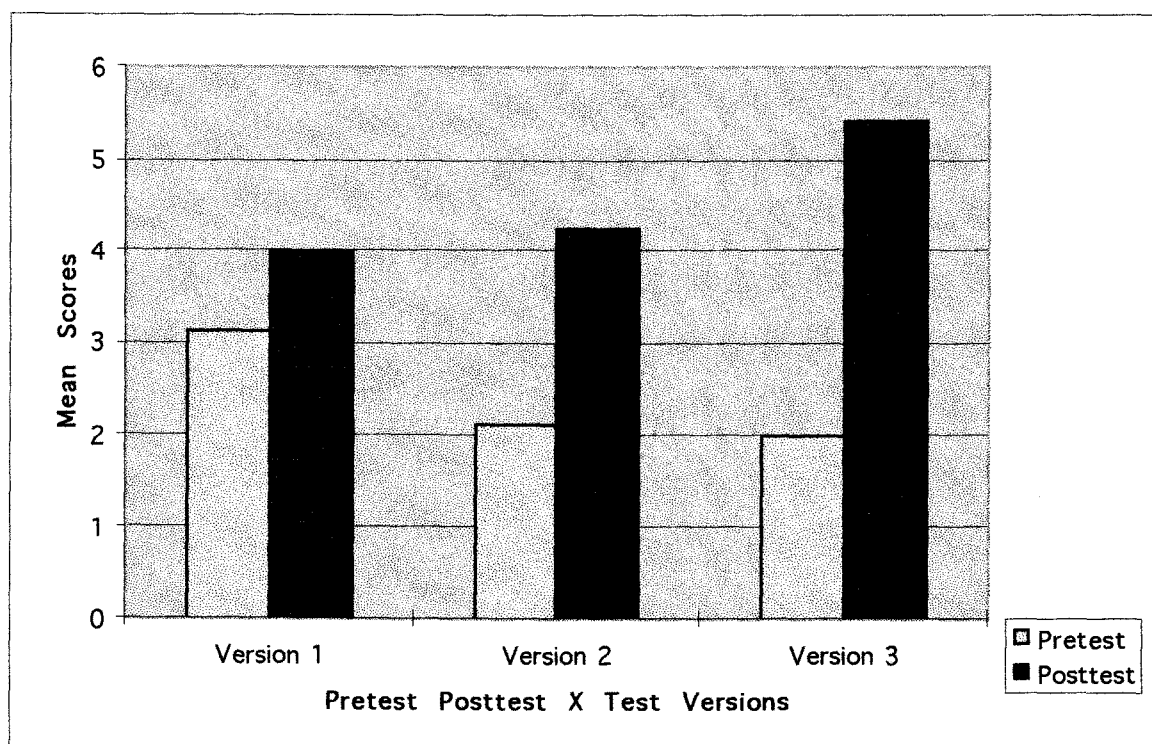
FIGURE 1. Mean 3 Point Scale Pretest and Posttest Scores by Test Versions



Paired sample t-tests were performed on each test version to determine learning effect. Analysis of test version 1 revealed convincing evidence that a learning effect did occur between pretest and posttest ($t_{20}=-4.94$, $p=.000$). Analysis of test version 2 revealed convincing evidence that a learning effect did occur between pretest and posttest ($t_{19}=-4.88$, $p=.000$). Analysis of test version 3 revealed convincing evidence that a learning effect did occur between pretest and posttest ($t_{18}=-10.39$, $p=.000$).

Using the one point scale, a one way analysis of variance determined that there was a difference between test versions on both pretest ($F_{2,57}=2.94$, $p=.06$) and posttest ($F_{2,57}=3.14$, $p=.05$), so the tests for hypothesis were carried out by test version to determine existing patterns. Figure 2 illustrates the patterns of learning effects by test version.

FIGURE 2. Mean 1 Point Scale Pretest and Posttest Scores by Test Versions



Paired sample t-tests were performed on each test version to determine learning effect. Analysis of test version 1 revealed a significant difference between pretest and posttest ($t_{20}=-1.81$, $p=.09$). Analysis of test version 2 revealed a significant difference between pretest and posttest ($t_{19}=-4.6$, $p=.000$). Analysis of test version 3 revealed a significant difference between pretest and posttest ($t_{18}=-8.56$, $p=.000$).

There was significant evidence that learning did occur during the study. Scoring both scales supports the alternative hypothesis, but also supports earlier findings, that the test forms are not equivalent. Again posttest version 3 appears to be the most reliable of the test forms.

Hypothesis 2: Main Effect of Treatment Conditions

Hypothesis 2 stated that there is no significant difference between psychodiagnostic proficiency of students who learn DSM taxonomy after receiving training in Structured Clinical Interview (single hypothesis decision tree) or problem-solving technique (multiple, competing hypotheses).

Both scoring scales were again analyzed to compare potential differences. A t-test was used to compare the effects of the the two treatment conditions.

Using the three point scoring scale the pooled t-test resulted in evidence for the idea that there was no association between treatment condition and posttest scores ($t_{17}=.51$, $p=.62$). Table 20 describes the means for the two treatment groups.

TABLE 20. 3 Point Scale Posttest Mean Scores by Treatment Condition

Treatment Group	Mean Posttest Score	SD
Decision Tree	9.10	3.542
Problem-Solving	8.1111	4.755

T-tests were performed on each test version to determine patterns. Test version 1 revealed no evidence of treatment effect (pooled $t_{19}=-.26$, $p=.8$). Test version 2 revealed no evidence of treatment effect (pooled $t_{18}=-1.24$, $p=.23$). Test version 3 revealed no evidence of treatment effect (pooled $t_{17}=.23$, $p=.82$). Table 22 describes the means for the two treatment groups by test version.

TABLE 21. 3 Point Scale Posttest Mean Scores by Treatment Condition and Test Version

Test	Group	Mean	SD	N
1	1	9.33	4.62	12
1	2	8.78	4.94	9
2	1	6.2	3.52	10
2	2	7.9	2.56	10
3	1	13.6	5.17	10
3	2	13.0	6.19	9

Using the one point scoring scale the pooled t-test resulted in evidence for the idea that there was no association between treatment condition and posttest scores ($t_{58}=-.53$, $p=.597$). Table 21 describes the means for the two treatment groups.

TABLE 22. 1 Point Scale Posttest Mean Scores by Treatment Condition

Treatment Group	Mean Posttest Score	SD
Decision Tree	4.4063	1.932
Problem-Solving	4.6786	2.019

Student's t-tests were performed on each test version to determine patterns. Test version 1 revealed no evidence of treatment effect (pooled $t_{19}=-.26$, $p=.8$). Test version 2 revealed no evidence of treatment effect (pooled $t_{18}=-1.17$, $p=.26$). Test version 3 revealed no evidence of treatment effect (pooled $t_{17}=.68$, $p=.51$). Table 22 describes the means for the two treatment groups by test version.

TABLE 23. 1 Point Scale Posttest Mean Scores by Treatment Condition and Test Version

Test	Group	Mean	SD	N
1	1	3.92	1.56	12
1	2	4.11	1.9	9
2	1	3.7	2.31	10
2	2	4.8	1.87	10
3	1	5.7	1.34	10
3	2	5.11	2.37	9

No evidence exists that treatment condition was associated with the effect on learning. Although learning occurred across both sites and treatment conditions, the cause for learning cannot be determined by the data.

Unhypothesized Findings

Interviews were held with each participant upon completion of the posttest to determine if the method of diagnostic decision making to which each had been exposed had been used during the posttest. Transcribed interviews resulted in the following results.

First, technical problems of various natures interfered with the audible recording of five interviews at Site 1, resulting in six full interviews, two partial interviews and three inaudible interviews. One interview began with the volume set at normal, but became inaudible when someone accidentally shut down the volume. The next interview, recorded on that tape was also inaudible until late in the interview, when the volume was turned up again. Two other interviews were held with the tape recorder on pause. The interviewer discovered the mistake after attempting to tape two interviews. The result was that neither interview was recorded. Finally, one interview was deleted when an interviewer rewound the tape recorder between interviews, recording the second interview over the first.

At Site 2, an attempt was made to prevent these technical problems by including training with the recorders before interviews began. However, eight of the 49 interviews were still inaudible. The first five, recorded inaudibly on day one of the experiment, were due to interviewers choosing a tape recorder which was conveniently available, but with which they were unfamiliar. Again, during retraining this problem was addressed and corrected, but on day two of the experiment three interviews were held in a hallway while other students moved between classes, resulting in overwhelming background noise.

Interviewers were briefed on the difference between multiple hypothesis and decision tree models of decision-making, so that they could distinguish between the two and were supplied with scripts containing structured interviews but were not informed of the hypotheses to be tested. Resultant quality of interviews was mixed. Some interviewers asked only the first of the four questions available to them, no matter how detailed or sketchy the participants' responses. Others faithfully carried out interviews until important data were gathered. Two other interviewers deviated from the script, both of these at Site 1, one so drastically that the participant stated, "I really don't know what you want here," and read the case study verbatim followed by the projected diagnosis. When the interviewer reiterated a question which deviated from the script, the participant stated, "I really don't understand what you mean. If you mean how did I come to that conclusion, I decided based on the symptoms."

There were 11 short interviews or interviews in which the participants continued to be vague about decision making process. Raters found it impossible to determine by which method these 11 participants decided diagnoses. Nevertheless, 31 interviews held adequate information to be analyzed, five from Site 1 and 26 from Site 2.

Of these, 16 were determined by both raters to have been probably using a binary decision-making process, two from Site 1 and 14 from Site 2. Of these 16, there was agreement between raters on 14 interviews.

A sign test was used to analyze interrater reliability. It was discovered that the raters did not tend to agree on the interpretation of the type of methodology used by participants ($p=.214$).

Typical responses among those interview responses on which raters agreed were as follows. "I saw evidence of paranoia in the article when he says, 'Refuses obviously sincere offers of help....' Throughout his life there have been numerous occasions in which he has displayed exaggerated suspiciousness....' Those two words led me to believe that he had a paranoid personality disorder."

"As I was reading Item A I remembered the criteria that described paranoia and the 85 year old man fits the criteria fairly well."

I remember this older gentleman who was depressed since the first grade so I felt that this was a really long-term depression he had going on; he had no interest or pleasure in anything. He was kind of like drifting away so to speak and in depression they kind of have self-pity. That's how I see it. He just sits alone; that was another thing. I chose depression probably from past experience with people I've known. Sitting home alone, doing nothing, having no interest in anything, just moping through life, that's what I have always associated depression with.

I remember that he can't ever feel comfortable socially, that he doesn't have very many friends, that he had no particular interest in dating. He didn't like having authority over people or giving instructions to them.... I remember from taking the computer tests that one of them, and I thought it was avoidance, was along the same line...

Twenty interviews were judged by at least one rater as being examples of multiple hypotheses decision-making. Of these, thirteen were agreed upon by both raters. Although most were more subtle than were many of the binary decision-making processes, a few appeared obvious.

For instance, one participant at Site 1 stated:

Well, I've talked already a lot about this case and tried to give you a diagnosis, but to be honest, it is hard for me to do that. It seems like there just isn't enough information. I mean, there are some clear symptoms, but I keep wanting to think, 'What if this is organic, or maybe there is something going on at school or in the family that I'm not aware of.' It's really hard to make a decision when there is so much to think about, you know?

Another Site 1 participant stated:

It's so confusing. We were taught in class that there's only one diagnosis, that it's best to only have one diagnosis, anyway; and then we were told in the film, you know, today, that we're supposed to think of a lot of things, which is how I naturally think anyway, which is why I don't really believe in the DSM anyway; but I couldn't get below several, well, 3 possibilities, and then I questioned them.

A Site 2 participant stated:

I learned the decision tree, but this makes more sense, to think the diagnosis over carefully. If you're going to write down a diagnosis in real life, it can be a dangerous thing, and you should be careful to think it over. It's hard, when you hear two ways to do things from the experts, to know which one is right, but this one seems right, to think things over carefully, consider a lot of possibilities.... I could have used a lot more time on the test to think about the case.

Another Site 2 participant told the interviewer:

I noticed that he was avoiding; that he does not like the responsibility; he feels a lot of anxiety at ...functions.... Another reason I put that is that he doesn't experience any anxiety or panic attacks; he basically takes the safe route in everything. He doesn't seem like a schizo; he doesn't seem dependent; he doesn't seem narcissistic. The other one that I thought of was maybe passive-aggressive because he did go to college and he did do some things... but avoiding was my main decision because he seemed to avoid any social situations.

Of group 1, binary decision tree oriented participants, whenever there was interrater agreement, nine participants were determined to be using decision tree methodology and seven were determined to be considering symptoms congruent with multiple hypotheses. There was disagreement on three interviews, and the rest were indistinguishable to one or both raters.

Of group 2 participants, those oriented to problem-solving diagnostic decision-making, both raters agree that six were using problem-solving techniques and that seven were using single hypothesis decision making on the case discussed during the interview. There was disagreement between raters regarding three group 2 participants. Eleven interviews were indistinguishable as one method or another according to at least one rater.

Although the interviews were not adequate to conclude that participants consistently used the orientations to which they were exposed in the treatment, enough evidence exists to suggest that some participants were not oriented by the treatment conditions, and more participants appeared not to be oriented to the treatment conditions than were. However, some participants appeared to be to some degree affected by the treatment conditions, but not enough to have any statistical relevance. This suggests that students are naturally oriented to one of the two diagnostic decision-making models as suggested by Elstein and associates (1978). It may also suggest that more time is needed to orient students to a preferred diagnostic decision-making model. Finally, it may suggest that skill levels are related to fund of information.

Research results have been mixed on this point. While de Mesquita (1992) and Elstein and associates (1978) suggested that diagnostic skills are not related to experience, Newell (1969) discovered that, in a closed information environment, deduction accuracy is directly related to fund of knowledge, and Feltovich (1981), studying medical student diagnostic abilities, found that logical competitor sets required an adequate fund of knowledge to overcome "dimensional restrictiveness" (Feltovich, 1981, p. 192). To some degree, both of these divergent research conclusions may be generally true. A fund of knowledge may be not only a necessity but quickly assimilated in a closed information environment. This argument may also be the most important when discussing the lack of difference between treatment condition effects. Although de Mesquita (1992) and Elstein and colleagues (1978) discovered that diagnostic ability was achieved quickly and that problem-solvers were consistently more accurate than were other types of diagnosticians, fund of knowledge may need a measurable threshold before this difference becomes apparent. Feltovich's dimensional restrictiveness began to disappear in the second year of medical training in sample populations who received training in logical competitor sets, but not in medical students who received training in traditional models of diagnosis. He found traditional students got "stuck in a chunk" and were not able to achieve "associative hypothesis spreading" when diagnostic error was discovered (Feltovich, 1981, p. 161). In fact, Katz (1990)

discovered that, when working with problem-solving skills, learning style and teaching method were overcome by self-study time until competence was achieved.

It is left to a longitudinal study to discover when during counselor diagnostic skills training problem-solving and decision tree orientations begin to make an effect on diagnostic proficiency, if preorientation affects outcome as suggested by previous studies (de Mesquita, 1992; Elstein, et al., 1978; Schmidt, et al., 1995; Goss, 1996). The evidence in this study does not support the Elstein and associates' (1978) study that early orientation to problem-solving during training improves diagnostic accuracy over other psychodiagnostic techniques, but Elstein and associates' sample population was drawn from advanced medical students and professional psychiatrists, which is congruent with the findings of Feltovich (1981).

Summary

Although inference may not be made beyond the study sample, reanalysis of previous studies which use case studies to determine diagnostic proficiency appear in order to determine if similar instrument variations exist. Traditional use of case studies to determine the process by which proficient diagnosticians make decisions would be seriously weakened by results comparable to those of this study. Even research which has used case studies to determine the effects of study of DSM on learning, the effects of CAI on the study of DSM, or the comparison of teaching methods on DSM learning would be weakened by the potential of test error.

The various results of the data analysis for this study suggests that researchers should not only take care in assignment of case studies to any assessment instrument and perform adequate correlation studies but that further work is in order to develop valid instrumentation for assessment of diagnostic proficiency.

There was support for the idea that a learning effect did exist in this study but not for a treatment effect on learning. Additional data were collected on

several characteristics of participants such as participant age, gender, computer experience, classroom exposure to psychodiagnosis, and education level. The potential relationship of these attributes will be examined in the next section.

Explanatory Variables

Data were gathered on several potential explanatory variables. Literature reviews and discussions with committee members suggested that attributes of participants had either previously been significant variables in other studies or could be in this study. These included age, education level, undergraduate grade point average, computer experience, gender, cultural origin, classroom experience with the DSM, and professional exposure to the DSM.

The explanatory variable of cultural origin could not be analyzed due to insufficient numbers of ethnic members in the sample. Although inference cannot be made beyond the sample itself, the lack of minorities represented at both sites reflects the reported state of minority representation in counselor related educational communities (Carey, Reinert, & Fontes, 1990). Because of the concerns of counselor educators about potential cultural bias in DSM diagnosis (Velasquez, John, & Brown-Cheatham, 1994; Ritchie, 1994) and the concern that cultural competency is not being achieved in counselor education (Carey, et al., 1990), an adequate representation of minority students to allow study of cross cultural training of psychodiagnostics would have added information to the literature. Particularly, the results of this study would have been interesting, if analyzable, because those few minority members who were represented presented interesting responses to the study's testing. Table 24 describes study participants by self-reported ethnic origin test version and treatment group.

TABLE 24. Ethnicity of Participants by Treatment Group and Test Version

Ethnicity	Group 1			Group 2			Total All Tests
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3	
Cauc	11	8	11	7	9	8	54
Afri/Am	0	0	0	0	0	1	1
Hisp	1	1	0	0	1	0	3
Nat/Am	0	1	0	1	0	0	2

The variable professional experience with the DSM could also not be analyzed. So few participants had experienced any professional experience with the DSM that analysis of professional DSM exposure could not have been analyzed with significance.

Data were gathered by the Hyperaxis II software on four research variables. These variables (percent of case studies reviewed, percent of practice test items reviewed, time spent studying the CAI, and percent of correct test items on the practice test included in Hyperaxis II) were analyzed to consider the question, "Were study preferences associated with test scores?" Additionally, the narrative recall answer sheet format for the study assessment instrument inadvertently collected data which were considered potentially important to the study and so analyzed post hoc with the planned explanatory variable data. Several participants wrote multiple hypotheses on the test form. In light of the interviews not providing adequate data to answer the question, "Did participants use the orientation to which the treatment condition exposed them," this inadvertent data -- although not considered to be reliable -- did provide another method to speculate about the question, "Were multiple diagnostic conjectures, written onto the posttest, associated with scores?" The data were recorded as the number of multiple hypotheses written on a participant's posttest and the number of items on which the participant wrote more than one diagnostic hypothesis.

The Instrument

A Pearson Product-moment correlation procedure was used to explain posttest scores; but, because the possibility of instrument variations existed, the variables were analyzed by test version. Correlation coefficient tables are provided in Appendix I.

The first question of interest was, "Are there explanatory variable patterns which explain the posttest scores?" Although some patterns do develop, the most important discovery is that explanatory variables are not consistent across either test versions or scoring scales. No single variable comprehensively explains posttest scores. It appears that not only are test versions not equivalent forms, but explanation of the scores is multifactor as well. Some attributes, study patterns and diagnostic decision-making orientations of participants are associated with posttest scores. However, the patterns that do develop are not consistent across all test versions. Still, some patterns do exist, and some variables are more associated with scores than others. Table 25 describes the association between significant variables and posttest scores.

The most consistent variable is multiple written hypotheses on the posttest answer sheets. Although the variable is positively correlated on only one test version using the 3 point scale, the correlation improves to two test version and from .36 to .49 on the same test version using the 1 point scale. This explains as much as 24 percent of the variance in test version 2. Education levels and age appear partially confounded with each other, but both appear to be to some extent associated with scores. Education level is generally positively correlated with scores, but in one case is negatively correlated with scores, although this may be an artifact. Computer experience appears to be negatively correlated with scores. Again, in one test version it appears to be positively correlated with scores, but this appears to be an artifact as well. Classroom exposure to DSM and percentage of correct items on the Hyperaxis II practice test were positively correlated with scores on one test version only. Problem-solving as rated by interview rater 1 and percentage of practice test items reviewed on Hyperaxis II were negatively correlated with

TABLE 25. Correlations And Significance of Variables to Posttest Scores By Test Version Using 3 Point And 1 Point Scales

MHNPO = Number of Multiple Hypotheses Written on Posttests
 MHCPO = Number of Multiple Hypothesis Items Written on Posttests
 AGE = Age of Participants
 EDATT = Attained Education Level of Participants
 GPA = Undergraduate Grade Point Average of Participants
 CXMO = Computer Experience of Participants by Month
 CDSM = Classroom Experience with DSM by Weeks
 PCTCREV = Percent of Case Studies Reviewed in Hyperaxis II
 PCTTREV = Percent of Practice Test Items Reviewed in Hyperaxis II
 PCTCTI = Percent Correct Practice Test Items in Hyperaxis II
 TIME = Time Spent in Study of Hyperaxis II by Minutes
 INTV 1 = Problem-Solving Decision Making as Determined by Interview Rater 1
 INTV 2 = Problem-Solving Decision Making as Determined by Interview Rater 2
 GENDER = Gender of Participants
 (Correlation Coefficient Stated First, P Value in Parentheses)

Variable	3 Point Scoring Scale			1 Point Scoring Scale		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
MHNPO		.36 (.06)		.39 (.04)	.49 (.01)	
MHCPO				.38 (.04)	.40 (.04)	
AGE	.64 (.07)			.62 (.00)		
EDATT	.78 (.00)			.46 (.02)		-.32 (.09)
GPA						
CXMO	.40 (.04)		-.69 (.00)			-.53 (.01)
CDSM	.51 (.01)			.46 (.12)		
PCTCREV						
PCTTREV	-.30 (.09)					
PCTCTI			.53 (.01)			.46 (.03)
TIME						
INTV 1						-.43 (.05)
INTV 2						
GENDER						

scores on one test version only. Grade point average, gender, time spent in study, and percentage of case studies reviewed on Hyperaxis II were not significantly correlated with posttest scores.

Because of varying patterns of explanatory variables all further analysis was performed using the three test versions. Also, potential explanatory variables which did not appear to be associated with posttest scores were reanalyzed to attempt to explain both study hypotheses.

Association of Explanatory Variables with Diagnostic Learning

The association of explanatory variables with the relationship associated with Hypothesis 1 was tested using a Pearson product-moment correlation procedure. The dependent variable was the arithmetic difference between pretest and posttest scores. A correlation coefficient matrix is provided in Appendix J. Table 26 describes the patterns of association between explanatory variables and the learning effect discovered in analysis of Hypothesis 1.

As in analysis of associations of explanatory variables with posttest scores, no comprehensive patterns developed to explain apparent differences between pretest and posttest scores. Nevertheless, the patterns that did develop allow speculation about the influences on learning effect in this study.

An interesting pattern developed between the scoring scales. The difference between pretest to posttest scores on the 3 point scale was better explained by participant attributes -- particularly test version 1 -- with variance accounted for ranging from 9 percent to 26 percent. The difference between pretest to posttest scores on the 1 point scale was best explained by the inclusion of multiple hypotheses on answer sheets -- particularly test version 3. This suggests that scoring on the test versions is not just related to the instrument itself but to multiple factors including the attributes and actions of participants. Although inference may not be made beyond the sample observed in this study, the study has evidence that case studies used as

TABLE 26. Correlations And Significance of Variables to Difference
between Pretest and Posttest Scores By Test Version
Using 3 Point And 1 Point Scales

MHNPO = Number of Multiple Hypotheses Written on Posttests
 MHCPO = Number of Multiple Hypothesis Items Written on Posttests
 AGE = Age of Participants
 EDATT = Attained Education Level of Participants
 GPA = Undergraduate Grade Point Average of Participants
 CXMO = Computer Experience of Participants by Month
 CDSM = Classroom Experience with DSM by Weeks
 PCTCREV = Percent of Case Studies Reviewed in Hyperaxis II
 PCTTREV = Percent of Practice Test Items Reviewed in Hyperaxis II
 PCTCTI = Percent Correct Practice Test Items in Hyperaxis II
 TIME = Time Spent in Study of Hyperaxis II by Minutes
 INTV 1 = Problem-Solving Decision Making as Determined by Interview Rater 1
 INTV 2 = Problem-Solving Decision Making as Determined by Interview Rater 2
 GENDER = Gender of Participants
 (Correlation Coefficient Stated First, P Value in Parentheses)

Variable	3 Point Scoring Scale			1 Point Scoring Scale		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
MHNPO		.48 (.02)			.59 (.00)	.37 (.06)
MHCPO		.43 (.03)			.53 (.01)	.38 (.05)
AGE	.51 (.01)					
EDATT	.47 (.01)					
GPA	.41 (.03)			.42 (.03)		
CXMO	.45 (.02)	-.30 (.10)	-.75 (.00)		-.30 (.10)	
CDSM	.30 (.09)					
PCTCREV						
PCTTREV						
PCTCTI			.49 (.02)		-.36 (.07)	.33 (.09)
TIME						
INTV 1			-.38 (.07)			-.55 (.01)
INTV 2						
GENDER						

diagnostic proficiency assessment instruments may need to be analyzed carefully not just for internal consistency but also for multiple factor influences on resultant scores, including both participant attributes and diagnostic decision-making styles.

The variable most consistently positively correlated with learning was multiple hypotheses written by participants on posttests. Although not significant across all test versions, evidence of problem-solving diagnostic decision-making was positively correlated on two test versions using the 1 point scale and on one test version using the 3 point scale. This discovery is congruent with the findings of Elstein and associates (1978), de Mesquita (1992), and Feltovich (1981) that diagnostic students who orient toward the problem-solving diagnostic decision-making model become more proficient diagnosticians faster than do those with a decision tree decision-making style. Although the treatment effect did not occur, participants who indicated they used problem-solving decision-making by writing multiple conjectures on tests scored higher. Interestingly, the more simplistic scoring scale appeared to be more sensitive to the influence of problem-solving on learning.

Data were collected by the software recording the apparent preferred study style of the participant. The variables were percentage of case studies reviewed, percentage of practice test items reviewed, percentage of correct test items, and time spent in study. Percentage of case studies reviewed ranged from 16 percent to 141 percent with a mean of 89 percent reviewed. Percentage of test items reviewed ranged from 50 percent to 233 percent with a mean of 123 percent reviewed. Percentage of correct practice test items ranged from 18 percent to 100 percent with a mean of 74 percent correct. Time spent in study of the CAI ranged from 26 minutes to 2 hours with a mean of 75 minutes.

These data were analyzed to determine their association with the learning effect. Percentage of correct practice test items was the only variable that had significant association with learning (3 pt. scale, test 3, $r=.49$, $p=.02$; 1 pt. scale, test 2, $r=-.36$, $p=.07$; test 3, $r=.33$, $p=.09$). The lack of correlation between preferred organization of study by participants and the difference between pretest and posttest scores suggests that the design of the computer

software was an effective instructional instrument which provided an equal opportunity to learn the material despite apparent study preferences.

The Hyperaxis II program not only supplies the student with learning strategies, such as immediate outcome feedback, which have been shown to be effective, but also nonsequential, student controlled learning environment, congruent with the various study organization preferences of individuals. The software may overcome some of the difficulties of classroom strategies which for some learners do not provide the needed environment for quick assimilation of information. The study results, which supported the idea that there was no difference between participant scores based on organization of material studied, may have some practical significance. Katz (1990) reported that learning problem-solving was hampered by mismatch of teaching and learning strategies. Katz discovered that teaching-learning style mismatched study participants required significantly more personal study time to improve skills than did style matched participants. An interesting possibility is that the nonsequential nature of Hyperaxis II provides the match to study strategies and negates differences.

Further support for suggestion that the CAI was an effective instructional instrument is the strong negative correlation between computer experience and learning. Apparently, the software design negated the effect of and need for computer experience. With only 2 to 3 minutes of instruction, all participants were able to use the learning program with a minimum of tutoring from research staff. This is a critical element in computer assisted learning. The alternative software investigated for use in this study, DTree (Spitzer, First, Williams & Gibbon, 1994), a DOS based program, required two hours of this author's time with technical assistance from two computer consultants to learn minimal use of the program.

Another intriguing result of this analysis is the negative correlation of problem-solving diagnostic decision-making as judged by interview rater 1 with learning scores on the same test version as a positive correlation of evidence on the written answer sheets that problem-solving was taking place. The weakness of interviews may raise questions, but many interviews had been

screened out before rating, so that the interviews which were rated did attempt to identify decision-making model of participants. Consequently, it appears that item choice for determination of participant decision-making has poor alternative forms reliability. This suggests that choice of items for think aloud protocols may an invalid method to determine participant diagnostic decision-making style as is multiple cases and written inventories. Inference may not be drawn beyond the sample in this study, but the finding suggests that alternative methods of investigation of diagnostic proficiency should be carefully analyzed before assumptions of validity are made. Table 27 describes the correlations between written multiple hypotheses on posttests and diagnostic decision styles of participants as determined by both interview raters.

TABLE 27. Correlations And P-Value of Multiple Written Hypotheses on Posttest to Interview Ratings of Multiple Hypothesis Use by Study Participants on Selected Posttest Items Using 3 Point and 1 Point Scoring Scales

MHNPO = Number of Multiple Hypotheses Written on Posttests
 MHCPO = Number of Multiple Hypothesis Items Written on Posttests
 (Correlation Coefficient Stated First, P Value in Parentheses)

Scale	Test Version	Interview Rater	MHNPO	MHCPO
3 Pt.	1	1	.09 (.37)	.20 (.24)
3 Pt.	1	2	.07 (.42)	.16 (.32)
3 Pt.	2	1	-.18 (.25)	-.23 (.32)
3 Pt.	2	2	.34 (.13)	.35 (.12)
3 Pt.	3	1	-.48 (.03)	-.49 (.03)
3 Pt.	3	2	-.23 (.23)	-.20 (.27)
1 Pt.	1	1	.09 (.37)	.20 (.24)
1 Pt.	1	2	.07 (.42)	.16 (.32)
1 Pt.	2	1	-.18 (.25)	-.13 (.32)
1 Pt.	2	2	.34 (.13)	.35 (.12)
1 Pt.	3	1	-.48 (.02)	-.49 (.03)
1 Pt.	3	2	-.24 (.23)	-.20(.27)

Association of Explanatory Variables to Treatment Effect

The association between the explanatory variables and the results of analysis of Hypothesis 2 was determined using a Pearson Product-moment procedure. Table 28 describes the association between the significant explanatory variables and the treatment effect by test version, using both the 3 point and 1 point scoring scales.

Significant correlates to group effect vary dramatically by test version but not by scoring scale. There are few explanatory variables which are associated with group effect as is expected, since no significant group effect existed. The few explanatory variables that did appear associated with any group effect which may have occurred were first of all age, education level and grade point average. There is some potential of a nonsignificant site effect on any difference between mean group scores. Correlations are mixed so that no trend can be described from the data. Some association exists in test version 2 between the learning effect and group scores. This might reflect the insignificant mean increase of problem-solving group scores over decision tree group scores. Finally, a negative association exists between gender and treatment effects on test version 3. Since gender is nonsignificant in all other analyses, this is assumed to be an artifact.

Further evidence is presented that explanatory variables for test scores are multifactor including not only nonequivalent forms but also participant attributes and actions which vary by test version. Otherwise, the treatment effect data reinforces the nonsignificance of treatment effect in this study.

TABLE 28. Significant Correlations Between Explanatory Variables
And The Treatment Effect Using 3 Point and 1 Point Scoring Scales

MHNPO = Number of Multiple Hypotheses Written on Posttests
 DIF = Difference Between Pretests and Posttests
 MHCPO = Number of Multiple Hypothesis Items Written on Posttests
 AGE = Age of Participants
 EDATT = Attained Education Level of Participants
 GPA = Undergraduate Grade Point Average of Participants
 CXMO = Computer Experience of Participants by Month
 CDSM = Classroom Experience with DSM by Weeks
 PCTCREV = Percent of Case Studies Reviewed in Hyperaxis II
 PCTTREV = Percent of Practice Test Items Reviewed in Hyperaxis II
 PCTCTI = Percent Correct Practice Test Items in Hyperaxis II
 TIME = Time Spent in Study of Hyperaxis II by Minutes
 INTV 1 = Problem-Solving Decision Making as Determined by Interview Rater 1
 INTV 2 = Problem-Solving Decision Making as Determined by Interview Rater 2
 GENDER = Gender of Participants
 (Correlation Coefficient Stated First, P Value in Parentheses)

Variable	3 Point Scoring Scale			1 Point Scoring Scale		
	Test 1	Test 2	Test 3	Test 1	Test 2	Test 3
MHNPO						
MHCPO						
DIF		.32 (.08)			.37 (.06)	
AGE		.32 (.08)		-.39 (.04)	.32 (.08)	
EDATT	-.39 (.04)					
GPA	.50 (.01)			.50 (.01)		
CXMO						
CDSM						
PCTCREV						
PCTTREV						
PCTCTI						
TIME						
INTV 1						
INTV 2						
GENDER			-.37 (.06)			-.37 (.06)

Summary

Analysis of potential explanatory variables resulted in inconsistent correlations with posttest scores, learning effect, and treatment conditions. No patterns emerged which could explain test version variations, which suggested that multiple factors beyond the scope of that presented in literature were associated with scores across test version. Since correlation studies had not produced explanatory patterns, a case study of item content was conducted.

Item Content Analysis

An analysis of items for frequency of correct answers was performed to determine if easy or difficult items existed. Because the total correct scores on the 3 point scale were similar to the total correct scores on the 1 point scale, only the 1 point scale was analyzed. A table of frequencies is provided in Appendix J. Items I, P, and R were consistently more difficult for participants while items A and H were consistently easier for participants to get correct scores. However, when items were removed from tests, reliability coefficients did not change dramatically. A table of Chronbach alpha coefficients without difficult and easy items is provided in Appendix J.

The factors related to test variations appeared to be related to the content of individual items. The DSM-III-R categorizes personality disorders into three clusters based on common signs and symptoms. Cluster A encompasses disorders which result in odd or eccentric behavior including Paranoid Personality Disorder, Schizoid Personality Disorder and Schizotypal Personality Disorder. Cluster B disorders appear dramatic, emotional, or erratic and includes Antisocial Personality Disorder, Borderline Personality Disorder, Histrionic Personality Disorder, and Narcissistic Personality Disorder. Cluster C is characterized by anxiety and fear and includes Avoidant Personality Disorder, Dependent Personality Disorder, Obsessive Compulsive Personality Disorder, and Passive Aggressive Disorder (American Psychiatric Association,

1987, p. 337). Several analyses were conducted to determine if the internal characteristics of these clusters were related to test variations.

An analysis of frequency of correct answers inside of test versions to determine if a learning effect occurred related to presence of other disorders within a cluster provided no evidence of learning effect. An analysis to determine if pairs of disorders within a cluster were associated with increase in score provided no evidence of pair influence. A table is provided in Appendix J that describes frequency of correct responses by cluster and test version.

An analysis that attempted to associate distribution of items with total score presented evidence of patterns of association between clusters and scores. Figure 3 illustrates the number of Cluster A items present and total mean scores on the three pretest versions. Figure 4 illustrates the number of Cluster A items present and total mean scores on the three posttest versions. Figure 5 illustrates the number of Cluster B items present and total mean scores on the three pretest versions. Figure 6 illustrates the number of Cluster B items present and total mean scores on the three posttest versions. Figure 7 illustrates the number of Cluster C items present and total mean scores on the three pretest versions. Figure 8 illustrates the number of Cluster C items present and total mean scores on the three posttest versions.

FIGURE 3. One-Point Scale Mean Pretest Scores on Three Test Versions and Corresponding Total Number of Items for Cluster A Personality Disorders

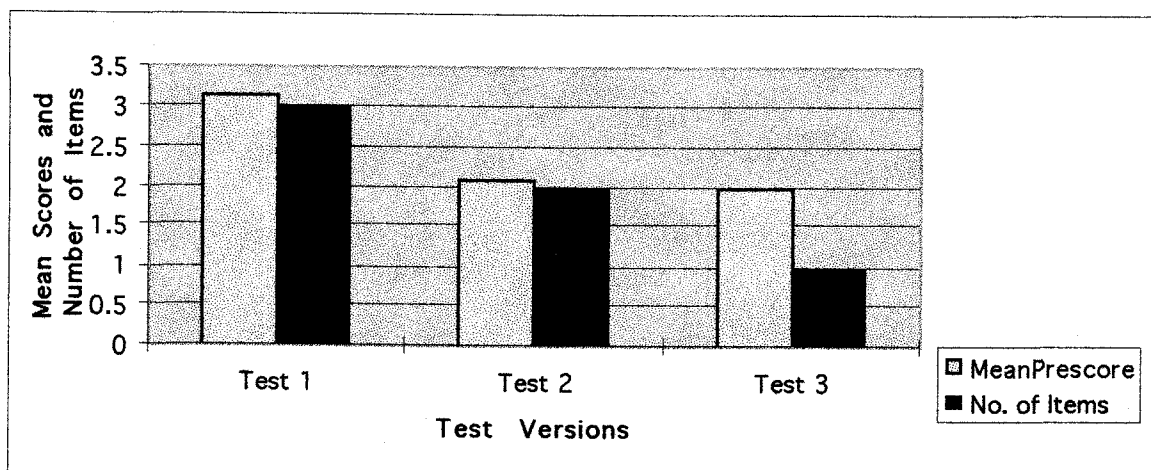


FIGURE 4. One-Point Scale Mean Posttest Scores on Three Test Versions and Corresponding Total Number of Items for Cluster A Personality Disorders

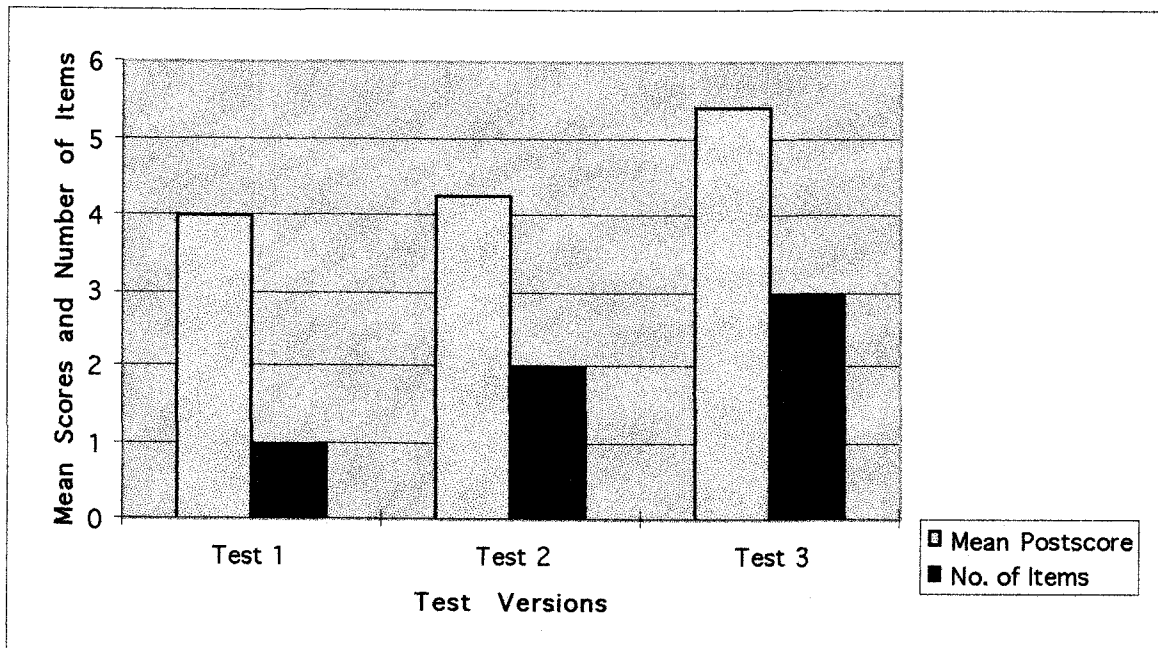


FIGURE 5. One-Point Scale Mean Pretest Scores on Three Test Versions and Corresponding Total Number of Items for Cluster B Personality Disorders

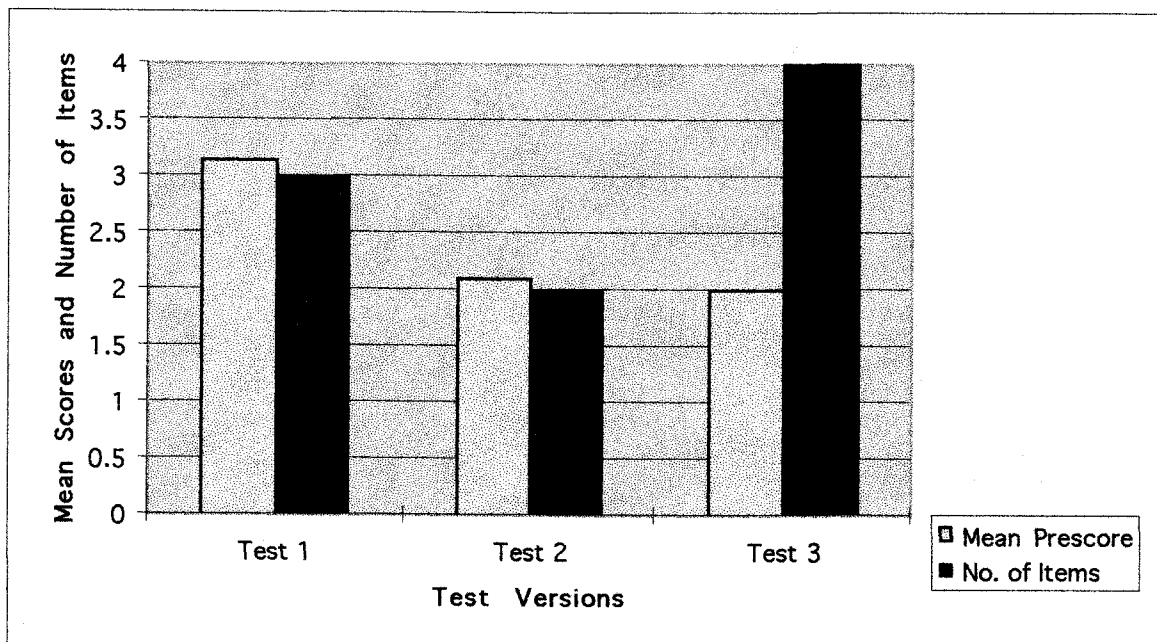


FIGURE 6. One-Point Scale Mean Posttest Scores on Three Test Versions and Corresponding Total Number of Items for Cluster B Personality Disorders

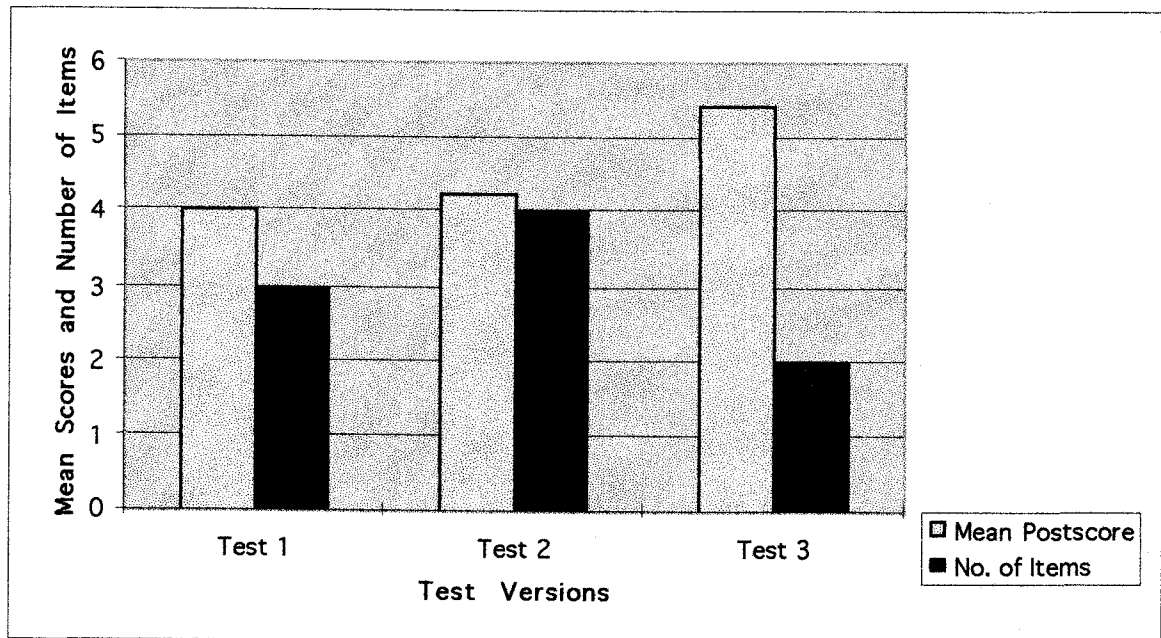


FIGURE 7. One-Point Scale Mean Pretest Scores on Three Test Versions and Corresponding Total Number of Items for Cluster C Personality Disorders

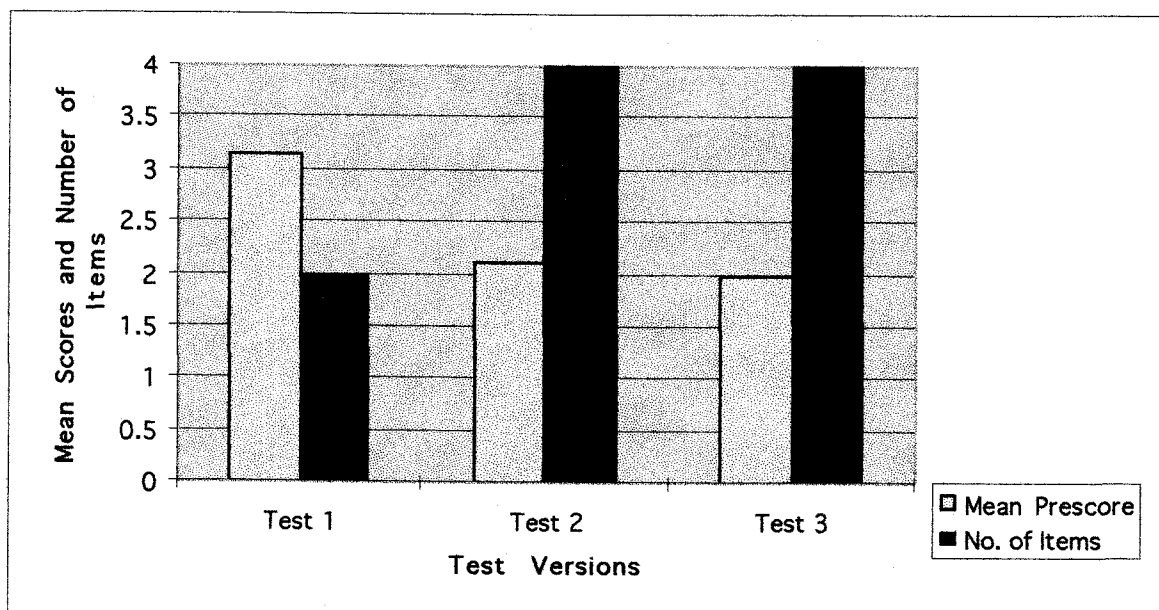
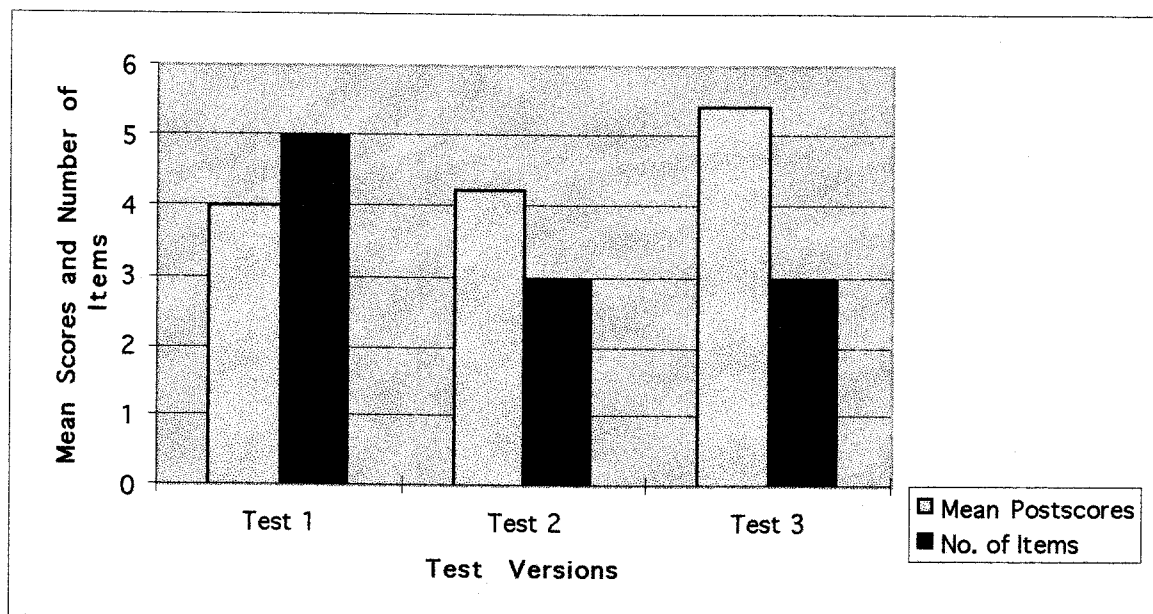


FIGURE 8. One-Point Scale Mean Posttest Scores on Three Test Versions and Corresponding Total Number of Items for Cluster C Personality Disorders



A clear pattern develops for Cluster A and somewhat of a pattern develops for Cluster C, the pattern is erratic for Cluster B, which also seems to reflect the relatively low reliability of this test version (3-point pretest=.23, 3-point posttest=.38, 1-point pretest=.1, 1-point posttest=.66).

Further evidence that personality disorder clusters were related to scores was presented by an analysis of the internal consistency of the diagnostic clusters by test version. Reliability coefficients were generally larger than were the reliability coefficients of test versions themselves (see Tables 4 to 16), except in clusters with too few items for an analysis. Additionally, cluster reliability coefficients varied by test version, suggesting that the content of individual items affected the overall reliability of any assignment of items inside a cluster. Moreover, reliability varied for many test version clusters according to grading scale used, which suggested that the separate scales were more sensitive to certain items than others. Table 29 describes the internal consistency of cluster assignments across test versions using the 3 point scale.

Table 30 describes the internal consistency of cluster assignments across test versions using the 1 point scale.

TABLE 29. Internal Consistency of Test Versions
by Personality Disorder Cluster Using 3 Point Scale

Test Version	Pre/Post	Diagnostic Cluster	N	Reliability Coefficient	F	P-Value
1	Pretest	A	3	.2358	5.20	.01
2	Pretest	A	2	.0000	.11	.75
3	Pretest	A	1	na ¹	na	na
1	Pretest	B	3	.4891	5.56	.01
2	Pretest	B	2	-.1569	.32	.58
3	Pretest	B	2	.3947	7.65	.00
1	Pretest	C	2	.5845	1.88	.19
2	Pretest	C	4	na ²	na	na
3	Pretest	C	4	na ²	na	na
1	Posttest	A	1	na ¹	na	na
2	Posttest	A	2	.4303	20.23	.00
3	Posttest	A	3	.7644	8.98	.00
1	Posttest	B	3	.5557	4.83	.10
2	Posttest	B	4	.2213	4.07	.01
3	Posttest	B	2	-.6811	.26	.61
1	Posttest	C	5	.2297	11.16	.00
2	Posttest	C	3	.2698	16.70	.00
3	Posttest	C	2	.8607	4.06	.03

1=Too few items

2=No participants diagnosed case correctly

TABLE 30. Internal Consistency of Test Versions
by Personality Disorder Cluster Using 1 Point Scale

Test Version	Pre/Post	Diagnostic Cluster	N	Reliability Coefficient	F	P-Value
1	Pretest	A	3	-.2931	4.27	.02
2	Pretest	A	2	.1176	.00	1.00
3	Pretest	A	1	na ¹	na	na
1	Pretest	B	1	.4313	5.38	.01
2	Pretest	B	3	-.1569	.32	.58
3	Pretest	B	2	.2044	3.49	.02
1	Pretest	C	2	.5845	1.88	.19
2	Pretest	C	4	na ²	na	na
3	Pretest	C	3	na ²	na	na
1	Posttest	A	1	na ¹	na	na
2	Posttest	A	2	.2339	7.11	.02
3	Posttest	A	3	.6231	5.26	.01
1	Posttest	B	3	.5104	1.36	.27
2	Posttest	B	4	.4429	2.11	.11
3	Posttest	B	2	-.0941	.39	.54
1	Posttest	C	5	.2783	6.73	.00
2	Posttest	C	3	.5567	13.40	.00
3	Posttest	C	3	.7989	.49	.62

1=Too few items
2=No participants diagnosed case correctly

Internal characteristics of clusters seem to affect scores, and internal characteristics of items seem to affect cluster scores and internal consistency. Since evidence that internal cluster and syndrome factors are the most consistent factor which explain test variations speculation about the nature of the clusters is in order.

Personality disorder Cluster A appears to be easy to discriminate. Frequency of correct answers ranged from .47 to .85 with a mean of .596 on posttests (See Appendix J). Each diagnosis inside the cluster has a single

distinct symptom or sign which appears to delineate it from other diagnoses. For instance, Schizoid individuals prefer social isolation but are not eccentric in behaviors like Schizotypal individuals.

Personality disorder Cluster B, although it shows a pattern of increased item difficulty from Cluster A, appears to have item difficulty which varies by item. Frequency of correct answers ranges from .35 to .667 with a mean of .498 on posttests. This suggests that some items are more difficult to conceptualize and learn than others. Also, no single item is easily discriminated. The highest frequency of correct posttest answers on Cluster A was .85 and on Cluster C was .947. In fact, it is not uncommon for experienced professionals to have difficulty with Histrionic Personality Disorder, which has features in common with dissociative disorders and with psychiatric disorders. Conversely, narcissism is a trait commonly referred to in society in general.

Personality disorder Cluster C shows a pattern of increased item difficulty, and the cluster appears to have an even wider range of item difficulty than Cluster B. Frequency of correct answers ranges from .10 to .947 with a mean of .564 on posttests. It appears that items may be distinguishable from other clusters but in the case of some paired syndromes not so much from each other. The signs and symptoms of some Cluster C disorders do not vary much from others. For instance, Passive Aggressive Personality Disorder varies from Avoidant Personality Disorder in the method by which individuals with the two diagnoses avoid responsibility.

Although inference may not be made beyond this study, it appears that the features of psychological diagnoses are multifactored and that these individual item factors are associated with the internal consistency and scores in diagnostic proficiency assessment instruments using case studies. It could be speculated that the features of individual diagnoses may be associated with the widely varied reliability studies of many psychological assessment instruments, since literature reports widely varied reliability of individual diagnoses for accepted instruments such as the MMPI and SCID.

It appears that any attempt at construction of a psychological disorders assessment instrument or psychological diagnostic proficiency instrument

should include an analysis of the content of diagnostic items and clusters, since reliability studies do not result in consistently high reliability coefficients. It also appears that the assumption that any single case study will be a valid assessment for measure of diagnostic proficiency may be inappropriate no matter what the testing protocol.

Finally, it appears that the internal characteristics of psychological disorder clusters affect how students learn them. A frequency of correct responses analysis of items and clusters by group revealed that clusters and items both were associated with differences in item difficulty between treatment conditions, which were masked by the overall insignificance of treatment effects. The treatment group which was exposed to the decision tree diagnostic model had a substantially higher frequency of correct answers on individual items and overall in personality disorder Cluster A. The treatment group which was exposed to the problem-solving diagnostic decision-making model had a substantially higher frequency of correct answers on individual items and overall in personality disorder Cluster B. Treatment condition appeared associated with higher frequency of correct answers on certain items for each treatment condition and none on others in personality disorder Cluster C. This suggested that early in training decision tree orientation may facilitate the learning of singular symptom syndromes, while problem-solving orientation may facilitate the learning of more complex syndromes. Table 31 describes the treatment effect by group, item and cluster.

The Elstein and colleagues (1978) discussion of associative memory of diagnostic learning and their discussion of remembering in chunks appears related to the internal factors of diagnoses as much as to the memory method discussed in the Elstein, et al. studies. The development of an instructional paradigm for DSM diagnostic skills for counselor education may be more related to the nature of the individual syndromes and their relation to category than to cognitive process. It appears that an instructor may not need to spend significant time on personality disorder Cluster A syndromes but should spend adequate time on Cluster B to point out the difficulty of these syndromes

on Cluster C to facilitate the ability to discriminate between alike syndromes. It also appears that instructors may be able to use both orientations to diagnostic decision-making to teach particular clusters, based on the complexity of the diagnosis.

TABLE 31. Frequency of Correct Responses to Assessment Instrument Items by Treatment Condition

DT = Decision Tree Diagnostic Decision-Making Model PS = Problem-Solving Diagnostic Decision-Making Model						
Cluster	Item	Pre/Posttest	Treatment Condition			
			DT % Correct	N	PS % Correct	N
A	A	Pretest	50	12	33	9
A	D	Pretest	10	10	11.1	9
A	M	Pretest	13.6	22	10.5	19
A	N	<u>Pretest</u>	<u>27.3</u>	<u>22</u>	<u>15.8</u>	<u>19</u>
		Mean	25.2	16.5	17.6	14
A	A	Posttest	90	20	78.9	19
A	D	Posttest	59.1	22	26.3	19
A	M	Posttest	60	10	44.4	9
A	N	<u>Posttest</u>	<u>60</u>	<u>10</u>	<u>33.3</u>	<u>9</u>
		Mean	67.3	15.5	47.7	14
B	B	Pretest	50	22	50	18
B	E	Pretest	33.3	12	11.1	9
B	I	Pretest	18.8	32	7.1	28
B	K	Pretest	30	20	0	19
B	L	<u>Pretest</u>	<u>0</u>	<u>10</u>	<u>22.2</u>	<u>9</u>
		Mean	26.4	19.2	18.1	16.6
B	B	Posttest	50	10	80	10
B	E	Posttest	50	20	57.9	19
B	J	Posttest	37.5	32	57.1	28
B	K	Posttest	66.7	12	66.7	9
B	L	<u>Posttest</u>	<u>34.4</u>	<u>22</u>	<u>52.6</u>	<u>19</u>
		Mean	47.7	19.2	62.8	17

C	F	Pretest	20	20	26.3	19
C	G	Pretest	40.9	22	21.1	19
C	H	Pretest	60	10	30	10
C	O	Pretest	0	20	15.8	19
C	P	Pretest	0	10	0	9
C	R	<u>Pretest</u>	<u>18.2</u>	<u>22</u>	<u>0</u>	<u>17</u>
		Mean	23.2	17.3	15.5	15.5
C	C	Posttest	75	32	64.3	28
C	F	Posttest	41.7	12	22.2	9
C	G	Posttest	100	10	77.8	9
C	H	Posttest	81.8	22	88.9	18
C	O	Posttest	66.7	12	77.8	9
C	P	Posttest	18.2	22	15.8	19
C	R	<u>Posttest</u>	<u>0</u>	<u>10</u>	<u>18.2</u>	<u>11</u>
		Mean	63.9	20	60.8	17.2

Participant Attitudes Toward Use of Hyperaxis II

Upon completion of the learning session Hyperaxis II presented a demographic questionnaire which asked each participant a series of questions. A post hoc analysis of data gathered from the Hyperaxis II program was conducted to determine the ease of use of the software and the preferences of participants toward use of the software.

Participants were asked, "How easy was this program to use?" An equal appearing interval scale was used to record responses, with the following scores: very easy = 1, easy = 2, difficult = 3, and very difficult = 4. A Kolmogorov-Smirnov procedure discovered strong evidence that the high rating could not have occurred by chance (MaxD=.42, $p < .10$). Table 32 describes participant responses.

Participants were then asked, "Overall, how well did you like this form of learning?" Again, a equal appearing interval scale was used to record responses, with the following scores: very much = 1, like it = 2, dislike it= 3, and strongly dislike it = 4. A Kolmogorov-Smirnov procedure discovered strong

evidence that the high rating could not have occurred by chance ($\text{MaxD}=.74$, $p<.10$). Table 33 describes participant responses.

TABLE 32. Hyperaxis II Ease of Use

	Very Easy (1)	Easy (2)	Difficult (3)	Very Difficult (4)	Total N	Mean
N	19	30	4	0	53	1.72
%	35.8%	56.6%	7.6%	0%	100%	

TABLE 33. Preference for Computer Assisted Learning

	Very Much (1)	Like It (2)	Dislike It (3)	Strongly Dislike It (4)	Total N	Mean
N	14	32	2	0	48	1.75
%	29.2%	66.7%	4%	0%	100%	

Three additional questions were asked, each with a space provided for comments. Comments were recorded and printed out for later analysis.

The first comment question was, "What is your attitude toward computers?" Few participants responded, but those who did expressed either strong dislike or highly positive comments. One participant stated, "Hate, rage, I loathe computers," but most stated such comments as, "confident," or, "I could spend all day on the computer...."

The second comment question was, "What did you like about this program?" Comments included sections of the program that participants

preferred and concentrated on, the ease of use, the nonsequential nature of the program, and the speed with which one could access and assimilate information.

The third comment question was, "What did you dislike about this program?" Participants who rated themselves highly on computer skills commented that the program had glitches, was too slow, needed a larger screen, or took too long to change stacks. Others commented that the scorekeeping bothered them, that there was not enough time to work with the program, that the program was not made available to them after the study, that the program was too long, that an answer was wrong on the test case, that Macintosh computers are not easy to use and that action buttons were unpredictable.

In light of research which suggests that study participant attitudes may not be a valid method for evaluation of the effectiveness of CAI software (Jones & McCormac, 1992) a conservative interpretation of the attitudes of participants in this study toward the software used is in order. Regardless, the trend toward participant enjoyment of learning on computer assisted program suggests that, since evidence has already been presented that the program was effective, it is an appropriate learning instrument and should be considered for more expansive use in counselor education.

Conversely, the development of computer assisted software which on one hand is user friendly and on the other hand is congruent with the various skills of computer users may be a difficult task. Further, the feedback of two participants in this study, that reading the computer monitor for over an hour resulted in nausea, could also be a drawback to use of the technology.

Limitations exist for use of computer assisted learning software that do severely limit widespread use of computer assisted diagnostic instructional software. First, development of programs is a time consuming and potentially difficult process. The author of Hyperaxis II reported that, because of both difficulties with revisions and with copyright limitations, he was unwilling to upgrade the software (D. A. Patterson, personal communication, April 9, 1996). Patterson reported that technical assistance or dual authorship with a computer

programmer is required for development of instructional software. Additionally, software copyright laws differ from textbook authorship. Although many texts exist providing instruction in psychodiagnostics, including use of the DSM such as Sullivan (1954), Seligman (1980), Anastasi (1988), Seligman (1994), Kaplan & Sadock (1990), Costa & Widiger (1994), and Halgin & Whitbourne (1993), the use of DSM diagnostics in software has been restricted to research purposes by the authors of both the DSM and of DTree (D. A. Patterson, personal communication, April 9, 1996), so that classroom use of diagnostic software must either be limited to research purposes only, use of expensive and cumbersome software developed for DOS systems (Spitzer, First, Williams & Gibbon, 1994), or alternative psychodiagnostic approaches than use of the DSM. The expense related to the presence of a computer technician in the diagnostic laboratory, purchase of restricted software or development of software is prohibitive to proliferation of software as a classroom tool, even though there is evidence that psychodiagnostic learning is an ideal use of computer assisted software.

Summary

Before testing of hypotheses could occur a test for potential site differences was in order. Analysis revealed that no significant site differences existed, so samples were combined. Because of the potential test variations, two scoring systems were devised and tested separately. It was discovered that no matter how the instrument was scored internal consistency of the six instruments varied and the three test versions were not equivalent forms. The only instrument which appeared to have internal consistency was posttest version 3. Correlation studies suggested that using internal reliability to construct an instrument was not promising.

Additionally, in a test of Hypothesis 1 it was discovered that there was strong evidence for the alternative hypothesis, that learning did take place across all three test versions. However, a test of Hypothesis 2 revealed that no

treatment effect occurred -- that participants, whether oriented to the treatment condition or not, did not score significantly differently by treatment group.

Interviews, performed to determine if participants used the treatment conditions to which they were exposed, were problematic. Several interviews were not recorded, not audible or were not conducted in such a way that substantial evidence existed on which decision-making model was used by participants. However, of those which could be analyzed there was little agreement between the interviews and actual evidence present on the test forms which did attest to the presence of multiple hypotheses on assessment instrument answer sheets.

Potential explanatory variables were inconsistently significant across test versions, scoring scales, pretests and posttests. This suggested that item difficulty and test variations were associated with factors internal to items. Factor analyses revealed that the personality disorder clusters were associated with test variations, internal consistency and frequency of correct item responses. Moreover, individual items within clusters appeared to affect test variations and the internal consistency of clusters according to distribution of items between pretest and posttest. This evidence suggested that a factor analysis of content is as important to construction of a diagnostic proficiency assessment instrument as are reliability studies. The evidence also suggested that clusters are unevenly learned because of the internal item factors. It appears that early orientations to diagnostic decision-making models may facilitate the learning of particularly diagnostic clusters based on item difficulty even though orientation does not have a significant affect on overall early diagnostic learning.

The findings of this study suggest that computer assisted diagnostic education software offers a potential contribution to the field of counselor education. CAI also offers a potential contribution as a research instrument which controls for instructor bias.

Evidence in this study and prior studies indicate that instruction of the problem-solving diagnostic decision-making model requires a more extensive period of time than provided in this study before the differences between it and

the decision tree model and their effects on diagnostic proficiency can be determined. The study suggests that a longitudinal study is in order.

Most importantly, this study reveals convincing evidence that case studies as diagnostic proficiency assessment instruments are suspect unless carefully analyzed for instrument error. Although inference may not be made beyond the study sample, past studies have not performed the advisable instrument analyses and so may not measure what is intended. A reanalysis of these studies in this light may weaken some of the studies' findings. The most appropriate conclusion to be drawn may be that the development of a valid and reliable instrument is the next logical step before further research in diagnostic proficiency and training can proceed.

CHAPTER 5 CONCLUSIONS, IMPLICATIONS, AND LIMITATIONS

Introduction

This final chapter presents a summary of the research and interpretation of the findings. Along with a discussion of the limitations inherent in this study, conclusions will be presented with considerations for counselor education programs and implications for future research.

Summary of Study

The purpose of the present study was to examine the effects of two methods of psychodiagnostic decision-making on, first, the integration of the model and, second, the ability to make proficient diagnostic decisions while in training.

In order to examine the questions, 60 participants from two higher education sites were randomly assigned to two treatment groups. Participants in each group were presented with a 35 minute orientation to one of two treatment conditions -- either an orientation to a binary decision tree model, based on the Structured Clinical Interview for the DSM (First, et al., 1996), or to a problem-solving model, multiple competing hypotheses (Elstein, et al., 1978) -- and then given up to two hours of training in a computer assisted learning laboratory, utilizing the software, Hyperaxis II (Patterson, 1990), employing a pretest and posttest design. The pretest and posttest were made up of 18 cases drawn from the DSM-III-R Casebook (Spitzer, et al., 1989), which were randomly assigned to pretests and posttests. This procedure was repeated three times and three different test versions were prepared for equal dissemination to participants to control for instrument error. Tests were scored using two scales, a 3 point scale and a 1 point scale. Participants were interviewed upon completion of the posttest by trained research assistants to determine if they predominately used the treatment conditions to which they were exposed for making diagnostic decisions on the posttest.

Several explanatory variables were introduced into the study based on a review of the literature which gave evidence that in past studies all variables had significant effects on data in at least one of the studies. These variables included site related variables as well as age, gender, education level, computer experience, classroom preexposure to DSM diagnosis, undergraduate grade point average, cultural origin, and professional experience with psychodiagnosis.

Nonhypothesis items were analyzed based on data gathered by the Hyperaxis II software. These data included attitudes of participants toward the Hyperaxis II program, percentage of case studies and practice test items examined by participants, and time investments of participants using Hyperaxis II. These data allowed an examination of motivation, attitude and study organization preferences of participants and in some cases the effects of these variables on scores on the posttest.

Further analyses were done: first -- of the interviews to attempt to determine the methodology of participants for decision making during the posttest. Second -- the discovery that the narrative format of the test instrument resulted in written, speculated multiple diagnoses and symptoms by many participants, provided data for a post hoc quantitative analysis of the potential association between written multiple conjectures by participants and the main study hypotheses.

Analyses presented evidence that content factors were most influential on scores, test variations, and internal consistency. This evidence suggested that the most important factors for construction of a diagnostic proficiency assessment instrument as well as for teaching diagnostic skills may be the characteristics of diagnostic clusters and individual syndromes inside each cluster.

Summary of Discussion

Analysis of data revealed no significant differences between sites, so samples were combined for all further analyses. A one way analysis on test

scores variance resulted in evidence of variations between test versions using both the 3 point scoring scale and the 1 point scoring scale.

Random assignment of case studies from the DSM-III-R Casebook (Spitzer, et al., 1989) was shown to be unreliable, even though it is a traditional assessment instrument for studies of diagnostic decision-making and of teaching diagnostic skills (Lambert & Meier, 1992; Janikowski, et al., 1989; Berven & Scofield, 1980; Chan, et al., 1993; Hayden, 1990; Patterson & Yaffe, 1993; Shamian, 1991; Friedman, et al., 1995; Clay, et al., 1995). Further, it was also clear that even among professionals the major binary inventory, SCID, has low reliability (Williams, et al., 1992; First, et al., 1995; Kendler & Roy, 1995; Hillis, 1995; Kennedy, et al., 1995; Steiner, et al., 1995; Jacobsen, et al., 1995; Schotten, et al., 1993). Also, SCID has not yet been studied as a teaching instrument (even though it has been introduced as such). It appears that the two major instruments used to measure change in psychodiagnostic skills in counseling related students present internal validity problems.

Correlation studies discovered that no matter how the instrument was scored, internal consistency of the six instruments varied and the three test versions were not equivalent forms. Only posttest version 3 appeared to have relatively high internal consistency. Correlation studies suggested that using internal reliability to construct an instrument was not promising.

A test of Hypothesis 1 discovered strong evidence for the alternative hypothesis: learning did take place across all three test versions. A test of Hypothesis 2 revealed that no treatment effect occurred: participants did not score significantly differently based on treatment condition.

Evidence existed that multiple factors affected learning. Factor analyses appear to be appropriate when constructing a case study based diagnostic proficiency assessment instrument.

Qualitative interviews of participants to determine the diagnostic decision-making model used by a think aloud protocol using one posttest item were not generally well performed. However, those that were successful suggested that some participants were affected by orientation to one of the treatment conditions, but more participants were not than were affected. Results

of analysis of the interviews also suggested that the use of a single item from the posttest may not actually identify the decision-making model used by participants. Inadvertent data collected from the posttest answer sheets appeared more reliable for this purpose and suggested that use of multiple hypotheses was to some degree positively correlated with increased scores.

Results of the study suggested that computer assisted diagnostic education software may contribute substantially to the field of counselor education. CAI also may contribute as a research instrument.

A more extensive time period than provided in this study appears to be necessary to orient students to a diagnostic decision-making model and to study the differences between resultant diagnostic proficiency.

Limitations

The sample for this study was drawn from two select, criterion populations, so no inferences may be made beyond these populations. However, random assignment to treatment groups allows limited inference to be made regarding cause.

Another limitation of this study is a small sample population both overall and per site, particularly at Site 1. This limitation is endemic to study of graduate students in an experimental setting, so that replication of this study with graduate students on a larger scale would require a funding source adequate to several sites. In fact, the small number of counseling students limits inference to counselor education.

Further, use of undergraduates, particularly from a psychology program, limits direct application to a counselor education program per se. On the other hand, since the objective of this study was to observe and measure the influence of early orientation to diagnostic decision-making models on psychodiagnostic education, the study accomplishes its purpose.

Additionally, events at Site 1 placed limitations on interpretation of data gathered. Although reservations for rooms, media equipment, and the computer laboratory had been made significantly in advance and confirmed

twice before the experiment, research staff was confronted on the day of the experiment with locked doors, canceled reservations for video equipment, and computer laboratory technicians unaware of reservations. Although the study took place almost as scheduled, the delays caused an increase in the planned time of the study from 4 hours to 4:45, with the conclusion of the study at almost 12:45 p.m.. Impatience was evident in some participants, and one participant, who reported himself to be diabetic, appeared confused during the last hour of the study time and did not take the posttest. It is difficult to define the effect these events had on resultant quantitative data, but three of the four interviewers reported that participants were impatient during the posttest interviews, with the reported results that interviewers were somewhat distracted and uncomfortable with elongating the interviews. Although this may also have affected the posttest as well, Site 1 participants spent as much time answering questions on the posttest as did Site 2 participants, averaging 32 minutes with a maximum time of one hour.

Another limitation was reliance on outside sources for provision of background knowledge for participants. Two Site 1 participants, who had spent 5 weeks in a classroom concentrating study on the DSM, reported difficulty recognizing the difference between traits and full diagnosis of personality disorders. Furthermore, for an unknown reason, Site 1 participants had as much difficulty identifying Axis I diagnoses (which they had studied for 3 weeks) as did participants at Site 2 who had minimal exposure to specific disorders before the study. Four Site 1 participants reported having been oriented in class to single diagnosis, decision tree based psychodiagnostic decision-making. However, since only a few participants discussed the effect of the orientation on decision-making during the posttest, discussion is speculative beyond the statements of these individual participants.

Because Site 2 participants were drawn from 9 separate courses, scheduling of discussion and lecture time was so difficult that it was abandoned in favor of reading material. Reliance on written material explaining the structure of the DSM to Site 2 participants also limits discussion of its effects. Two participants stated that they did not understand the difference between

Axes I and II, while three participants reported a lack of understanding of traits versus full diagnosis, even though all participants signed statements that they had read the prestudy material.

The most severe limitation of this study is the qualitative portion, interviews of participants after completion of the posttest, and interpretation of those interviews. Although exact scripts were supplied and training provided to each interviewer, few of the interviewers followed the instructions during the interviews with the result that less than adequate information was obtained from the majority of participants and none from several.

The major inhibition imposed by a lack of comprehensive interview is that there can be no accurate determination as to whether participants consistently used the orientation to which they were exposed during the posttest. Indications from the post hoc analysis of multiple postulates written on the posttest would indicate that the trend did develop, but this post hoc data was neither planned nor consistently reliable.

The dearth of indepth interviews contributed heavily to interrater unreliability during interpretation of data. Most of the disagreements between raters were related to short interviews, terminated by interviewers without having asked even the second of the four questions. Although interviews improved with retraining between Site 2 session days, interviewers still tended to hold short interview sessions and seldom completed the four interview questions provided on the script. During the retraining session at Site 2, interviewers described their performance during the first experiments as adequate and productive and had difficulty understanding the need for more lengthy interviews.

It is difficult to speculate how much improvement in interrater reliability would have accompanied more comprehensive interviews, but studies which have used interview raters to determine methodology of diagnosticians has been significantly more successful (de Mesquita, 1992; Elstein, et al., 1978). On the other hand, in this study the lack of agreement between interviews which were successful and the inadvertent evidence of written multiple hypotheses on

instrument answer sheets also suggests a limitation on the idea that interviews can determine the actual decision-making model used by participants.

Recommendations for Further Research

The most important consideration that results from this study is that the development of instrumentation for diagnostic skills assessment appears to be an important task. In a field founded in development of assessment instruments, adoption of less than valid or reliable assessment instruments in counselor diagnostic education does not achieve the ends of adequately trained professionals nor of consistency within the field.

Pilot studies should be undertaken to determine internal consistency before instruments are used for study or assessment. Additionally, the construction of diagnostic assessment instruments should include factor analysis of item difficulty.

The response of the counselor community toward development of fresh perspectives, effective and ethical training of diagnosis and development of valid and reliable instrumentation appears an appropriate goal. Further, there is opportunity to not only increase credibility alongside other mental health professionals but to provide modeling for effective and ethical use of diagnosis.

This study did not provide evidence of early orientation to a diagnostic decision-making model nor effect of that orientation on proficiency. However, the development and teaching of a model compatible with counseling theory and dedicated to increase in diagnostic proficiency is appropriate. Study of an instructional paradigm is important, particularly since studies in the counseling field suggest that the DSM is being adopted. It is also important in light of evidence that counseling is expanding into mental health fields rapidly and that mental health professionals view psychological diagnosis as a skill deficiency among counselors.

Factor analyses suggested that students may learn individual diagnostic clusters more proficiently using an orientation which matches the complexity of the diagnoses within individual clusters. Exploration of this potential paradigm

for diagnostic instruction is appropriate. If results of further studies add evidence that content not only affects diagnostic proficiency assessment but also diagnostic learning, the adoption of decision tree and problem-solving instruction would be in order.

Moreover, the positive results of this study strongly suggests that computer assisted technology limitations need be overlooked and that computer assisted instruction be developed. Not only was there evidence of a learning effect, but evidence existed that preferred student study organization preferences were easily and consistently matched by the software package and so negated problems introduced by teaching styles. Software development would require working relationships between traditionally separate departments in the university community and better communication between institutions to share technologies already developed.

Development and testing of valid and reliable instruments for assessment of the development of diagnostic skills is in order. The lack of validity and reliability in this study suggests that replication is not appropriate until a valid and reliable instrument is developed. The doubtful validity and reliability of other binary or case study based inventories reviewed for this study appears to also weaken other research results. If case studies are to be used, then extensive testing to develop an instrument which measures change in skills and does so consistently is important. Pilot studies should precede use of assessment instruments to determine internal consistency. Moreover, factor analysis of item difficulty should be undertaken during construction of diagnostic assessment instruments.

Other existing instruments, the Modified Essay Examination (Brown, 1987), the Diagnostic Inventory Rating Scale (Bögels, 1994), and the psychodiagnostic learning inventory developed by Boshuizen and associates (1995) have been studied little and only with medical students. Replication of those studies would not only reinforce findings of the authors but also potentially provide improved alternatives over traditional testing methods. Adaptation of these instruments for use with non psychiatrically trained counselors would also allow study of the instrument for validity and reliability

with counseling students. If replication either gives evidence of unreliability or invalidity of the above instruments or investigation suggests that the instruments are inappropriate for other reasons, then development of an alternative instrument is merited. A distinct possibility is continued work with a case study based inventory until validity and reliability are achieved.

The development of instrumentation is particularly important in light of the difficulty of this study establishing and maintaining a group of interviewers capable of and interested in data collection. If replication of qualitative methods of data collection are undertaken, lengthy and thorough training is warranted. This would require either the budget or other incentives that would guarantee the time involvement of research assistants. Training sessions should include training films, role plays until measured competence is achieved, practice with and consistent supply of electronic recording equipment, and facilities which would guarantee comprehensive data gathering without added bias on the part of research assistants.

Think-aloud protocols based on case studies may prove to be just as invalid and unreliable as written inventories. The evidence presented in this study that interviews did not have alternate forms reliability with obvious evidence of multiple hypotheses written on assessment instrument answer sheets suggests that this alternative method may be just as flawed as an undeveloped written instrument. The choice and testing of case studies for any form of inventory appears to be a more critical element than the instrument protocol.

It is also recommended that development and use of computer assisted instruction continue. The evidence presented by this study that preferred student study style and teaching style mismatch may be easily overcome by user friendly software developed according to the standards set out in the literature for effective use of technology, controls for instructor bias and learning styles so that confounding variables are substantially decreased.

A longitudinal study of psychodiagnostic education of counselor trainees would provide a valuable addition to the literature and important insights into effective teaching methodologies and skills development chronology. Also,

studies across several institutions would permit indepth reports of the state of the art of psychodiagnostic education, beyond the survey material presently available and would increase the ability of research to infer beyond specific sample populations. These studies would also provide research data specific to the counseling field, presently reliant on a few studies of counseling diagnostic education and more heavily reliant on medical research into diagnostic learning. They would also begin to provide a base of information and the beginning of a dialogue within counselor education about the appropriate and ethical base for counselor psychodiagnostic philosophy and development rather than a dialogue of whether to or not to diagnose. Finally, these studies could provide leadership across other professional counseling related fields toward an increased ethical and caring diagnostic standard.

Bibliography

Abouserie, R., Moss, D., & Barasi, S. (1992). Cognitive style, gender, attitude toward computer-assisted learning and academic achievement. Educational Studies, 18 (2), 151-160.

Albanese, M. J., Bartel, R. L., Bruno, R. F., Morgenbesser, M. W., Schatzberg, A. F. (1994). Comparison of measures used to determine substance abuse in an inpatient psychiatric population. American Journal of Psychiatry, 151 (7), 1077-1078.

Alpert, D. (1986). A preliminary investigation of computer-enhanced counselor training. Computers in Human Behavior, 2, 63-70.

American Counseling Association. (1995, April). Code of Ethics and Standards of Practice. Alexandria, VA: Authors.

American Psychiatric Association. (1952). Diagnostic and statistical manual of mental disorders. Washington DC: Authors.

American Psychiatric Association. (1994). Diagnostic and statistical manual of mental disorders (4th ed.). Washington DC: Authors.

Anastasi, A. (1988). Psychological Testing. New York: Macmillan Publishing Company.

Anderson M., & Hornby, P. (1990). Computer use in psychology instruction. Behavior Research Methods, Instruments, & Computers, 22 (2), 194-199.

Atkinson, D. R., Worthington, R. L., Dana, D. M., & Good, G. E. (1991). Etiology beliefs, preferences for counseling orientations, and counseling effectiveness. Journal of Counseling Psychology, 38 (3), 258-264.

Banyan, C. D., & Stein, D. M. (1990). Voice synthesis supplement to a computerized interviewing training program: Retention effects. Teaching in Psychology, 17 (4), 260-263.

Barlow, D. H. (1991). Introduction to the special issues on diagnoses, dimensions, and DSM-IV: The science of classification. Journal of Abnormal Psychology, 100 (3), 243-244.

Barrett-Lennard, G. T. (1962). Dimensions of therapist response as causal factors in therapeutic change. Psychological Monographs: General and Applied, 76 (43), 1-36.

Benner, D. G. (1985). Baker encyclopedia of psychology. Grand Rapids, MI: Baker Book House.

Berven, N. L. (1985). Reliability and validity of standardized case management simulations. Journal of Counseling Psychology, 32 (3), 397-409.

Berven, N. L., & Scofield, M. E. (1980). Evaluation of clinical problem-solving skills through standardized case-management simulations. Journal of Counseling Psychology, 27 (2), 199-208.

Bögels, S. M. (1994). A structured-training approach to teaching diagnostic interviewing. Teaching of Psychology, 21 (3), 144-149.

Borg, W. R., & Gall, M. D. (1989). Educational research: An introduction (5th ed.). New York: Longman.

Boshuizen, H. P. A., Machiels-Bongaerts, M., Schmidt, H. G., & Hermans, H. (1995, April). Monitoring the development of expertise in a problem-based curriculum. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Brenneis, C. B. (1994). The skewing of psychiatry. American Psychiatry, 18 (2), 71-80.

Brothen, T. (1995). A computer-assisted difference threshold exercise for introductory psychology. Teaching of Psychology, 22 (2), 136-138.

Brown, S. (1985). Treating the alcoholic: A developmental model of recovery. New York: John Wiley and Sons.

Brown, S. W. (1987). Teaching and assessing clinical skills using a modified essay examination. Poster session presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Cade, B., & O'Hanlon, W. H. (1993). A brief guide to brief therapy. New York: W. W. Norton & Company.

Campbell, C. C., Dahir, C., Johnson, L., & Sheldon, C. (1996). School counseling program standards. The National Standards for School Counseling Programs. Miami: The American School Counselor Association.

Carey, J. C., Reinat, M., & Fontes, L. (1990). School counselors' perceptions of training needs in multicultural counseling. Counselor Education and Supervision, 29 (2), 155-169.

Chan, F., Rosen, A. J., Wong, D. W., & Kaplan, S. (1993). Evaluating rehabilitation caseload management skills through computer simulations. Journal of Counseling and Development, 71, 493-498.

Clarkin, J. F., Hull, J. W., Cantor, J., & Sanderson, C. (1993). Borderline personality disorder and personality traits: A comparison of SCID-II BPD and NEO-PI. Psychological Assessment, 5 (4), 472-476.

Clay, D. L., Hagglund, K., Frank, R. G., & Elliot, T. R. (1995). Enhancing the accuracy of depression diagnosis in patients with spinal cord injury using Bayesian analysis. Rehabilitation Psychology, 40 (3), 171-180.

Cleghorn, J. M. (1985). Formulation: A pedagogic antidote to DSM-III. Comprehensive Psychiatry, 26 (6), 504-512.

Cook, E. P., Berman, E., Genco, K., Repka, F., Shrider, J. (1986). Essential characteristics of master's level counselors: Perceptions of agency administrators. Counselor Education and Supervision, 27, 146-152.

Cook, E. P., Warnke, M., & Dupuy, P. (1993). Gender bias and the DSM-III-R. Counselor Education and Supervision, 32 (3), 311-321.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. Cognition, 58, 1-73.

Costa, P. T., & Widiger, T. A. (1994). Personality disorders and the five-factor model of personality. Washington, DC: American Psychological Association.

Council for Accreditation of Counseling and Related Educational Programs. (1994). CACREP accreditation standards and procedures manual. Alexandria, VA: Authors.

Council for Accreditation of Counseling and Related Educational Programs. (1995). The future of CACREP (Report to ACES regional meetings). Alexandria, VA: Authors.

Cummings, A. L., Martin, J., Hallberg, E., & Slemon, A. (1992). Memory for therapeutic events, session effectiveness, and working alliance in short-term counseling. Journal of Counseling Psychology, 39 (3), 306-312.

de Mesquita, P. B. (1992). Diagnostic problem solving of microcomputer use by academic psychologists. Teaching of Psychology, 16 (3), 145-147.

Dennison, R. (1993). Documenting for success in psychiatric services. Rockville, MD: National Community Mental Healthcare Council.

Downing, H. D., & Paradise, L. V. (1989). Using the DSM-III-R in counseling. Journal of Counseling and Development, 68, 226-227.

Einhorn, H. J. (1972). Expert measurement and mechanical combination. Organics, Behavior, and Human Performance, 7, 86-106.

Elmore, P. B., Ekstrom, R. B., Diamond, E. E., & Whittaker, S. (1993). School counselors' test use patterns and practices. The School Counselor, 41, 73-80.

Elstein, A. S., Shulman, L. S., Sprafka, S. A., Allal, L., Gordon, M., Jason, H., Kagan, N., Loupe, M. J., & Jordan, R. D. (1978). Medical problem solving: An analysis of clinical reasoning. Cambridge, MA: Harvard University Press.

Elwood, R. W. (1993). The clinical utility of the MMPI-2 in diagnosing unipolar depression among male alcoholics. Journal of Personality Assessment, 60 (3), 511-521.

Eysenck, H. J. (1952). The effects of psychotherapy: An evaluation. Journal of Consulting Psychology, 16 (5), 319-324.

Feletti, G. I., & Smith, E. K. M. (1986). Modified essay questions: Are they worth the effort? Medical Education, 20, 126-132.

Feltovich, P. J. (1981). Knowledge based components of expertise in medical diagnosis. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

First, M. B., Spitzer, R. L., Gibbon, M., & Williams, J. B. W. (1996). Structured clinical interview for DSM-IV. New York: Biometric Research Department, New York State Psychiatric Institute.

First, M. B., Williams, J. B. W., Spitzer, R. L. (1988). The concise guide to using DTree: The electronic DSM-III-R. Toronto, ONT: Multi-Health Systems.

Fisher, S. R., & Martin, C. J. (1981). Standardized interviews in psychiatry: Issues in reliability. British Journal of Psychiatry, 139, 139-143.

Fitzgerald, L. F., Weitzman, L. M., Gold, Y., & Unerod, M. (1988). Academic harassment: Sex and denial in scholarly garb. Psychology of Women Quarterly, 12 (3), 329-340.

Fong, M. L. (1993). Teaching assessment and diagnosis within a DSM-III-R framework. Counselor Education and Supervision, 32 (3), 276-285.

Frances, A. J., First, M. B., Widiger, T. A., Miele, G. M., Tilly, S. M., Davis, W. W., & Pincus, H. A. (1991). An a to z guide to DSM-IV conundrums. Journal of Abnormal Psychology, 100, 407-412.

Friedman, D., Massaro, D. W., Kitzis, S. N., & Cohen, M. M. (1995). A comparison of learning models. Journal of Mathematical Psychology, 39 (2), 164-178.

Furlong, M. J., & Hayden, D. C. (1993). Computer-assisted training of the DSM-III-R in counselor education. Counselor Education and Supervision, 32 (4), 298-310.

Gentry, D. B. (1992). Using computer aided interactive video technology to provide experiential learning for mediation trainees. Journal of Divorce & Remarriage, 17 (3), 57-74.

Ginter, E. J. (1991). Mental health counselor preparation: Experts' opinions. Journal of Mental Health Counseling, 13 (2), 187-203.

Goldberg-Wood, G., & Middleman, R. (1987). Interviewing skills for the human services: Learner's workbook. Park Forest, Ill: OUT ST Software.

Goldman, L. (1994). The marriage between tests and counseling redux: Summary of the 1972 article. Measurement and Evaluation in Counseling and Development, 26, 214-216.

Goodyear, R. K. (1990). Research on the effects of test interpretation: A review. The Counseling Psychologist, 18 (2), 240-253.

Goss, J. R. (1996). Teaching clinical reasoning to second-year medical students. Academic Medicine: Journal of the Association of American Medical Colleges, 71 (4), 349-352.

Grob, G. N. (1991). Origins of DSM-I: A study in appearance and reality. American Journal of Psychiatry, 148 (4), 421-431.

Guenther, W. C. (1964). Analysis of Variance. Englewood Cliffs, NJ: Prentice-Hall.

Halgin, R. P., & Whitbourne, S. K. (1993). Abnormal psychology: The human experience of psychological disorders. Fort Worth: Harcourt Brace Javanovich College Publishers.

Hand, D. J., & Taylor, C. C. (1987). Multivariate analysis of variance and repeated measures. London: Chapman and Hall.

Hayden, D. C. (1990). A DSM-III-R computer tutorial for abnormal psychology. Teaching of Psychology, 17 (3), 203-205.

Hersen, M., & Bellack, A. S. (1981). Behavioral assessment: A practical handbook (2nd ed.). New York: Pergamon Press.

Hersen, M., & Turner, S. M. (1985). Diagnostic Interviewing. New York: Plenum Press.

Hills, H. A. (1995). Diagnosing personality disorders: An examination of the MMPI-2 and MCMI-II. Journal of Personality Assessment, 65, 21-34.

Hogarth, R. M. (1974). Process tracing in clinical judgment. Behavioral Science, 19, 298-313.

Hohenshil, T. H. (1993). Teaching the DSM-III-R in counselor education. Counselor Education and Supervision, 32, 267-275.

Holland, J. L. (1985). Making vocational choices: A theory of careers (2nd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Hornby, P. A., & Anderson, M. D. (1994). Computer use in psychology instruction: A survey of individual and institutional characteristics. Behavior Research Methods, Instruments, & Computers, 26 (2), 250-254.

Jacobsen, L., Perry, S., & Frances, A. (1995). Diagnostic agreement between the SCID-II screening questionnaire and the Personality Disorder Examination. Journal of Personality Assessment, 65 (3), 428-433.

Jacquez, J. A. (Ed.). (1964). The diagnostic process. Ann Arbor, MI: Mallooly Lithographing.

Janikowski, T. P., Berven, N. L., Meixelsperger, M. K., & Roedl, K. E. (1989). A computer-based case simulation to assess skill in predicting client behavior. Rehabilitation Counseling Bulletin, 33 (2), 127-139.

Jenkins, T. M., & Dankert, E. J. (1981). Results of a three-month PLATO trial in terms of student attitudes. Educational Technology, 21 (3), 44-47.

Johnson, J., Horwath, E., & Weissman, M. M. (1991). The validity of major depression with psychotic features based on a community study. Archives of General Psychiatry, 48 (12), 1075-1081.

Jones, B. T., & McCormac, K. (1992). Empirical evidence shows that measuring users' opinions is not a satisfactory way of evaluating computer-assisted learning in nurse education. International Journal of Nursing Studies, 19 (4), 411-425.

Kaplan, H. I., & Sadock, B. J. (1990). Pocket handbook of clinical psychiatry. Baltimore: Williams & Wilkins.

Katz, N. (1990). Problem solving and time: Functions of learning style and teaching methods. The Occupational Therapy Journal of Research, 10 (4), 221-236.

Keller, M. B., Klein, D. N., Robert, M. A., Hirschfeld, R. M. A., Kocsis, J. H., McCullough, J. P., Miller, I., First, M. B., Holzer, C. P., Keitner, G. I., Marin, D. B., & Shea, T. (1995). Results of the DSM-IV mood disorder field trial. American Journal of Psychiatry, 152 (6), 843-849.

Kendler, K. S., & Roy, M. A. (1995). Validity of a diagnosis of lifetime major depression obtained by personal interview versus family history. American Journal of Psychiatry, 152 (11), 1608-1614.

Kennedy, S. H., Katz, R., Rockert, W., Mendlowitz, S., (1995). Assessment of personality disorders in anorexia nervosa and bulimia nervosa: A comparison of self-report and structured interview methods. Journal of Nervous and Mental Disease, 183 (6), 358-364.

Kronk, A. P. (1985). Computer assisted counselor training: A client simulation (Doctoral dissertation, University of Michigan, 1985) Dissertation Abstracts International, 44, 8512447.

Krueger, D. W. (1982). Neurotic behavior and the alcoholic. In E. M. Pattison and E. Kaufman (Eds.), Encyclopedic handbook of alcoholism. New York: Gardner Press, pp. 607-618.

Kutchins, H., & Kirk, S. A. (1988). The business of diagnosis: DSM-III and clinical social work. Social Work, 8 (2), 215-220.

Lahey, B. B., Applegate, B., Barkley, R. A., Barfinkel, B., McBurnett, K., Kerdyk, L., Greenhill, L., Hynd, G. W., Frick, P. J., Newcorn, J., Biederman, J., Ollendick, T., Hart, E. L., Perez, D., Waldman, I., & Shaffer, D. (1994). DSM-IV field trials for oppositional defiant disorder and conduct disorder in children and adolescents. American Journal of Psychiatry, 1512 (8), 1163-1173.

Lambert, M. E. (1989). Using computer simulations in behavior therapy training. Computers in Humans Services, 5 (3-4), 1-12.

Lambert, M. E., & Lenthall, G. (1989). Effects of psychology courseware use on computer anxiety in students. Computers in Human Behavior, 5, 207-214.

Lambert, M. E., & Meier, S. T. (1992). Utility of computerized case simulations in therapist training and evaluation. Journal of Behavioral Education, 2, 73-84.

Luborsky, , Chandler, M., Auerback, A. H., Cohen, J., & Bachrach, H. M. (1971). Factors influencing the outcome of psychotherapy: A review of quantitative research. Psychological Bulletin, 75 (3), 145-185.

Lusted, L. B. (1968). Introduction to medical decision making. Springfield, Ill: Thomas.

Macaskill, N., Geddes, J., & Macaskill, A. (1991). DSM-III in the training of British psychiatrists: A national survey. The International Journal of Social Psychiatry, 37 (3), 182-186.

MacFadden, R. J. (1990). Sexual abuse assessment training: Developing CAI in child welfare. Computers in Human Services, 5 (3-4), 29-41.

MacKinnon, R. A., & Yudofsky, S. C. (1986). The psychiatric evaluation in clinical practice. Philadelphia: J. B. Lippincott Company.

Maple, F. F. (1994). The development of goal-focused interactive videodiscs to enhance student learning in interpersonal practice methods class. Computers in Human Services, 11 (3-4), 343-357.

Maser, J. D., Kaelber, C., & Weise, R. E. (1991). International use and attitudes toward DSM-III and DSM-III-R: Growing consensus in psychiatric classification. Journal of Abnormal Psychology, 100 (3), 271-279.

McCabe, G. E. (1956). How substantial is a substantial validity coefficient? The Personnel and Guidance Journal, 34, 340-344.

McGuire, C. H. (1985). Medical problem-solving: A critique of the literature. Journal of Medical Education, 60 (8), 587-595.

Miller, G. M., & Larabee, M. J. (1995). Sexual intimacy in counselor education and supervision: A national survey. Counselor Education and Supervision, 34 (4), 332-343.

Millon, T. (1983). The DSM-III: An insider's perspective. American Psychologist, 38 (7), 804-814.

Moore, D. S., & McCabe, G. P. (1993). Introduction to the practice of statistics (2nd ed.). New York: W. H. Freeman and Company.

Murphy, H. B. M. (1980). Hidden barriers to the diagnosis and treatment of alcoholism. Journal of Studies on Alcohol, 41 (5), 417-428.

Nelson, M. L. (1989). The relation of gender to power and involvement in supervision. Unpublished doctoral dissertation, University of Oregon, Eugene.

Nelson, M. L., & Holloway, E. L. (1990). Relation of gender to power and involvement in supervision. Journal of Counseling Psychology, 37(4), 473-481

Nelson-Gray, R. O. (1991). DSM-IV: Empirical guidelines from psychometrics. Journal of Abnormal Psychiatry, 100 (3), 308--315.

Newell, A. (1969). Heuristic programming: Ill structured problems. In J. Aronofsky (Ed.), Progress in operations research (Vol. 3). New York: Wiley.

Niles, S. G., & Pate, R. H. (1989). Competency and training issues related to the integration of career counseling and mental health counseling. Journal of Career Development, 16, 63-71.

Oregon Progress Board. (1994). Oregon benchmarks: Standards for measuring statewide progress and institutional performance. Salem, OR: The Oregon Economic Development Department.

Oregon Reg. Stat. 309-51-035 (7 [a]-[b])

Oregon Reg. Stat. 336.650 (1992).

Oregon Senate Bill 1099 (1991).

Pask, G. (1976). Styles and strategies of learning. British Journal of Educational Psychology, 46, 128-148.

Patterson, D. A. (1990). Hyperaxis II. Louisville, TN: Random Order Software.

Patterson, D. A., Lee, M., & Evers, E. (1992). Mapping hypermedia informational space utilization. Unpublished manuscript, University of Tennessee at Knoxville.

Patterson, D. A., & Yaffe, J. (1993). Using computer-assisted instruction to teach Axis-II of the DSM-III-R to social work students. Research on Social Work Practice, 3 (3), 343-357.

Polson, D. A. D. (1995). Fostering multiple repertoires in undergraduate behavior analysis students. The Behavior Analyst, 18 (2), 293-299.

Pope, K. S., Levenson, H. S., & Schoner, L. R. (1979). Sexual intimacy in psychological training. American Journal of Psychotherapy, 34 (8), 682-689.

Poulin, J. E., & Walter, C. A. (1990). Interviewing skills and computer assisted instruction: BSW student perceptions. Computers in Human Behavior, 7 (3-4), 179-197.

Prentky, R. A. (1994). In R. J. Corsini (Ed.), Encyclopedia of psychology, Vol 3 (2nd ed.). New York: John Wiley & Sons.

Putney, M. W., Worthington, E. L., & McCullough, M. E. (1992). Effects of supervisor and supervisee theoretical orientation and supervisor-supervisee matching on interns' perceptions of supervision. Journal of Counseling Psychology, 39 (2), 258-265.

Raffoul, P. R., & Holmes, K. A. (1986). DSM-III content in social work curricula: Results of a national survey. Journal of Social Work Education, 10, 24-31.

Ramsey, F., & Schafer, D. (1995). The statistical sleuth: A second course in statistical data analysis. Corvallis, OR: Oregon State University Book Store.

Rappaport, E. (1975). Effects of dogmatism and anxiety during computer-assisted learning. Psychological Reports, 37, 1055-1065.

Reid, W. H., & Wise, M. G. (1989). DSM-III-R training guide: For use with the American Psychiatric Association's diagnostic and statistical manual of mental disorders (third edition-revised). New York: Brunner/Mazel.

Ritchie, M. H. (1990). Counseling is not a profession - yet. Counselor Education and Supervision, 29, 220-227.

Ritchie, M. H. (1994). Cultural and gender biases in definitions of mental and emotional health and illness. Special section: Critical issues in research in counselor education and supervision. Counselor Education and Supervision, 33 (4), 344-348.

Ritchie, M. H., Piazza, N. J., & Lewton, J. C. (1991). Current use of the DSM-III-R in counselor education. Counselor Education and Supervision, 30, 205-211.

Roe, A. (1956). The psychology of occupations. New York: Wiley.

Rosenberg, C. E. (1979). The therapeutic revolution: Medicine, meaning, and social change in nineteenth-century America. In M. J. Vogel (Ed.), The therapeutic revolution: Essays in the social history of American medicine. Philadelphia: University of Pennsylvania Press.

Saigh, P. A. (1992). Structured clinical interviews and the inferential process. Journal of School Psychology, 30, 141-149.

Schachter, S., & Singer, J. E. (1962). Cognitive, social, and physiological determinants of emotional state. Psychological Review, 69, 379-399.

Schmidt, H. G., Machiels-Bongaerts, M., Hermans, H., Cate, O., Venekamp, R., & Boshuizen, P. A. (1995, April). The development of diagnostic competence: A comparison between a problem-based, an integrated, and a conventional medical curriculum. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Schotte, C., de Doncker, D., Maes, M., Cluydts, R., (1991). Low MMPI diagnostic performance for the DSM-III-R obsessive-compulsive personality disorder. Psychological Reports, 69 (3), 795-800.

Schotte, C., de Doncker, D., Maes, M., Cluydts, R. (1993). MMPI assessment of the DSM-III-R histrionic personality disorder. Journal of Personality Assessment, 60 (3), 500-510.

Seligman, L. (1980). Assessment in developmental career counseling. Cranston, RI: The Carroll Press.

Seligman, L. (1986). Diagnosis and treatment planning in counseling. New York: Human Sciences Press.

Seligman, L. (1993). Teaching treatment planning. Counselor Education and Supervision, 32, 287-297.

Sexton-Radek, K. (1993). Using computers to teach the role of professional psychologists. Teaching in Psychology, 20 (4), 248-249.

Shamian, J. (1991). Effects of teaching decision analysis on student nurses' clinical intervention decision making. Research in Nursing and Health, 14, 59-66.

Sharf, R. S., & Lucas, M. (1993). An assessment of a computerized simulation of counseling skills. Counselor Education and Supervision, 32, 254-266.

Shaw, L. W. (1979). A study of empathy training effectiveness: Comparing computer assisted instruction, structured learning training and encounter training exercises (Doctoral dissertation, Syracuse University, 1978) Dissertation Abstracts International, 38, 7908563.

Shore, K. (1992). Managed care as a totalitarian regime. Psychotherapy in Private Practice 14 (2), 67-75.

Siegel, S. (1956). Nonparametric statistic for the behavioral sciences. New York: McGraw-Hill Book Company.

Sinclair E., Forness, S. R., & Alexson, J. (1985). Psychiatric diagnosis: A study of its relationship to school needs. The Journal of Special Education, 19 (3) 333-344.

Smith, M. L., & Glass, G. V. (1977). Meta-analysis of psychotherapy outcome studies. American Psychologist, 32, 752-761.

Smith, P. J., Galdes, D., Fraser, J., Miller, T., Smith, J. W., Svirbely, J. R., Blazina, J., Kennedy, M., Rudmann, S., & Thomas, D. L. (1991). Coping with the complexities of multiple-solution problems: A case study. International Journal of Man Machine Studies, 35 (4), 429-453.

Smith, W. G. (1966). A model for psychiatric diagnosis. Archives of General Psychiatry, 14, 521-529.

Spitzer, R. L., & Endicott, J. (1968). DIAGNO: A computer program for psychiatric diagnosis utilizing the differential diagnostic procedure. Archives of General Psychiatry, 18, 746-756.

Spitzer, R. L., & Fleiss, J. (1974). A re-analysis of the reliability of psychiatric diagnosis. British Journal of Psychiatry, 125, 341-347.

Spitzer, R. L., Gibbon, M., Skodol, A. E., Williams, J. B. W., & First, M. B. (1989). The DSM-III-R Case Book. Washington DC: American Psychiatric Press.

Spitzer, R. L., Gibbon, M., Skodol, A. E., Williams, J. B. W., & First, M. B. (1994). The DSM-IV Case Book. Washington DC: American Psychiatric Press.

Steinberg, M., Cicchetti, D., Buchanan, J., Rakfeldt, J., & Rounsaville, B. (1994). Distinguishing between multiple personality disorder (dissociative identity disorder) and schizophrenia using the structured clinical interview for DSM-IV dissociative disorders. Journal of Nervous and Mental Disease, 182 (9), 495-502.

Steiner, J. L., Tebes, J. K., Sledge, W. H., & Walker, M. L. (1995). A comparison of the Structured Clinical Interview for DSM-III-R and clinical diagnoses. Journal of Nervous and Mental Disease, 183 (6), 365-369.

Strong, E. K., & Campbell, D. P. (1974). Strong-Campbell Interest Inventory. Stanford, CA: Stanford University Press.

Sullivan, H. S. (1954). The psychiatric interview. New York: W. W. Norton & Company.

Super, D. E. (1990). A life-span, life-space approach to career development. In D. Brown, & L. Brooks (Eds.). Career choice and development: Applying contemporary theories to practice (pp. 197-261). San Francisco: Jossey-Bass.

Thelen, M. H., Mintz, L. B., & Vander Wal, J. S. (1996). Psychological Assessment, 8 (2), 219-221.

Thomas, D. L., & Bostow, D. E. (1991). Evaluation of pre-therapy computer-interactive instruction. Journal of Computer-Based Instruction, 18 (2), 66-70.

Vaillant, G. (1981). Dangers of psychotherapy in the treatment of alcoholism. In M. Bean and N. Zinberg (Eds.), Dynamic approaches to the understanding and treatment of alcoholism. New York: Free Press, 55-96.

Velasquez, R. J., Johnson, R., & Brown-Cheatham, M. (1993). Teaching counselors to use the DSM-III-R with ethnic minority clients: A paradigm. Counselor Education and Supervision, 342, 323-331.

Waldo, M., Brotherton, W. D., & Horswill, R. (1993). Integrating DSM-III-R training into school, marriage and family, and mental health counselor preparation. Counselor Education and Supervision, 32, 332-342.

Warnath, C. F. (1975). Vocational theories: Direction to nowhere. Personnel and Guidance Journal, 52 (6), 422-428.

Warner, H. R., Toronto, A. F., Veasey, G., & Stephenson, R. (1961). A mathematical approach to medical diagnosis: application to congenital heart disease. Journal of the American Medical Association, 177, 177-183.

West, J. D., Hosie, T. W., & Mackey, J. A. (1988). The counselor's role in mental health: An evaluation. Counselor Education and Supervision, 28, 233-239.

Widiger, T. A., Frances, A. J., Pincus, H. A., Davis, W. W., & First, M. B. (1991). Toward an empirical classification of the DSM-IV. Journal of Abnormal Psychology, 100 (3), 280-288.

Wilcoxon, S. A. (1990). Community mental health counseling: An option for the CACREP dichotomy. Counselor Education and Supervision, 29 (3), 401-410.

Williams, J. B. W., Gibbon, M., First, M. B., Spitzer, R. L., Davies, M., Borus, J., Howes, M. J., Kane, J., Pope, H. G., Rounsaville, B., & Wittchen, H. (1992). The structured clinical interview for DSM-III-R (SCID). Archives of General Psychiatry, 49, 630-636.

World Health Organization, International Classification of Diseases (9th ed.). New York: United Nations Press.

Zimpfer, D. G. (1996). Professional Issues: Five-year follow-up of doctoral graduates in counseling. Counselor Education and Supervision, 35, 218-228.

Zunker, V. G. (1994). Career counseling: Applied concepts of life planning (4th ed.). Pacific Grove, CA: Brooks/Cole Publishing.

APPENDICES

Appendix A

Letter to Subjects

Dear Southern Oregon State College student:

As a doctoral student in counseling at Oregon State University, involved in dissertation research, I am contacting you to solicit your *Voluntary* participation in an experimental study. Accompanying this letter you will find an informed consent document. If you wish to participate in the study, please sign, date, and return it to me either personally or through your professor.

This study is a comparison of methods of teaching psychodiagnosis, using the DSM-III-R and computer-assisted learning. The object of the experiment is to discover which of two primary methods, presently used in psychology education, results in the most proficient diagnosticians.

Participants will be required to read a preparatory handout entitled Introduction to the Structure of the Diagnostic and Statistical Manual of Mental Disorders as a prerequisite for participation in the experiment, time to be set with the subject's agreement. On Friday, February 21, 1997, or on February 28, 1997, at 8:30 a.m. the actual study will begin. Students will take a brief pretest, be divided into two groups, each group being presented with a short video - describing an approach to psychodiagnosis which is currently used in psychology - and then will spend two hours in a diagnostic learning laboratory. A brief posttest, followed by a short interview, will finish the experiment. Four hours will be spent in the complete study.

Subjects will receive a coded identification number upon registration, and that code will be used throughout the study. The code will be needed to tabulate results and match pretest to posttest. All information received will

remain confidential. After results are matched, identifying information, including codes, will be obliterated.

In return for participation in this study, students at Southern Oregon State College who complete the study will learn diagnostic skills and will have their names entered into a drawing for a \$25 gift certificate at the SOSC Bookstore. Also, students who complete the study will receive extra credit in a select class in the Psychology Department. Extra credit varies in each class in which it is offered and the amount of extra credit is at the discretion of the Professor. Professors have informed classes of the amount available in each class. Extra credit may be taken in only one class for participation in the study. Results of the study will be made available to interested students at SOSC, and students interested in receiving training in the alternative diagnostic method to that which she/he received will be able to do so during Spring or Summer Quarters, 1997.

The deadline for application to the study will be February 14, 1997. Because space in the computer laboratory is limited, once an adequate number of participants are registered, no further applications will be accepted. Your participation will be greatly appreciated, will contribute to psychological diagnosis literature and will be much appreciated.

Sincerely yours,

Louis Downs

Letter to Subjects

Dear Oregon State University student:

As a doctoral student in counseling at Oregon State University, involved in dissertation research, I am contacting you to solicit your *Voluntary* participation in an experimental study. Accompanying this letter you will find an informed consent document, which, if you wish to participate in the study, will need to be signed, dated, and returned to me either personally or through your professor.

This study is a comparison of methods of teaching psychodiagnosis, using the DSM-III-R and computer-assisted learning. The object of the experiment is to discover which of two primary methods, presently used in psychology education, results in the most proficient diagnosticians.

Participants are being solicited only from DSM-IV/Abnormal Behavior, Coun 580, because you will have received basic training in the structure of the DSM and the nature of psychodiagnosis, needed preparation for the study. On Friday, February 7, 1997, at 8:30 a.m. the actual study will begin. Students will take a brief pretest, be divided into two groups, each group being presented with a short video - describing an approach to psychodiagnosis which is currently used in psychology - and then will spend two hours in a diagnostic learning laboratory. A brief posttest will finish the experiment. Altogether, four hours will be spent in the study.

Subjects will receive a coded identification number upon registration, and that code will be used throughout the study. The code will be needed to tabulate results and match pretest to posttest. All information received will remain confidential. After results are matched, identifying information, including codes, will be obliterated.

In return for participation in this study, students at Oregon State University who complete the study will learn diagnostic skills and will have their names entered into a drawing for a \$25 gift certificate at the OSU Bookstore. Also, students who complete the study will receive 10% extra credit in Coun 580, a bonus offered by the instructor. Results of the study will be made available to interested students at OSU, and students interested in receiving training in the alternative diagnostic method to that which she/he received will be able to do so during Spring or Summer Quarters, 1997.

The deadline for application to the study will be February 1, 1997. Your participation will be greatly appreciated, will contribute to psychological diagnosis literature and will be much appreciated.

Sincerely yours,

Louis Downs

Informed Consent Document, OSU subjects

Research project: Comparison of DSM diagnostic teaching techniques and resultant diagnostic proficiency.

Investigators: Louis Downs, doctoral student in Counseling, OSU
Dr. James Firth, Professor, Counseling, OSU

Purpose of the research project:

The purpose of the research is to determine if the instruction in the two major paradigms of psychodiagnosis affect the ability of students, studying DSM diagnosis, to made accurate diagnoses of case studies.

Procedures: I have received and oral and a written explanation of this study and I understand as a participant in this study that the following things will happen:

(a) My participation in this research is voluntary and that I am being offered an incentive of a drawing to be held at the end of the study in which I will have an equal chance with all other participants to win a gift certificate at the local college bookstore. Also, upon completion of the study, I will receive 10% extra credit toward my final grade in Coun 580, DSM/-IV/Abnormal Behavior.

(b) I will be asked to respond to a pre and post test and a demographic questionnaire to accumulate research data.

(c) The duration of my participation in the experiment is 4 hours over the course of one day.

(d) I will not be identified by name in any thesis, publication or presentation prepared by the researcher.

(e) All research records will be kept in a private locked location, with only the investigators allowed access to the information.

(f) I have a right to discontinue participation at any time, with no obligation.

(g) Results are available to me and may be obtained by contacting the investigators.

(h) I am aware that additional help will be offered to me by the investigator if there is any problem due to my participation in the study or if I wish alternative training.

This project has been reviewed and approved by the Oregon State University committee for Research Involving Human Subjects. The committee believes that the research procedures adequately safeguard the subject's privacy, welfare, civil liberties, and rights. My signature below indicates that I have read and understand the procedures described above and give my voluntary consent.

Name of subject

Signature of subject

Subject's phone number

Date signed

Social Security Number

SOSC Informed Consent Document

Research project: Comparison of DSM diagnostic teaching techniques and resultant diagnostic proficiency.

Investigators: Louis Downs, doctoral student in Counseling, OSU
Dr. James Firth, Professor, Counseling, OSU

Purpose of the research project:

The purpose of the research is to determine if the instruction in the two major paradigms of psychodiagnosis affect the ability of students, studying DSM diagnosis, to make accurate diagnoses of case studies.

Procedures: I have received and oral and a written explanation of this study and I understand as a participant in this study that the following things will happen:

(a) My participation in this research is voluntary and that I am being offered an incentive of a drawing to be held at the end of the study in which I will have an equal chance with all other participants to win a gift certificate at the local college bookstore. Also, I will receive extra credit in a select class in the Psychology Department. I understand that extra credit varies in each class in which it is offered and the amount of extra credit is at the discretion of the Professor. Professors have informed classes of the amount of extra credit available in each class. Extra credit may be taken in only one class for participation in the study.

(b) I will be asked to respond to a pre and post test and a demographic questionnaire to accumulate research data.

(c) The duration of my participation in the experiment is 4 hours over the course of one day for the experiment. Duration of my participation includes reading a preparatory handout entitled Introduction to the Structure of the Diagnostic and Statistical Manual of Mental Disorders. I will sign a statement on the day of the experiment to verify that I have read the introductory material.

(d) I will not be identified by name in any thesis, publication or presentation prepared by the researcher.

(e) All research records will be kept in a private locked location, with only the investigators allowed access to the information.

(f) I have a right to discontinue participation at any time, with no obligation.

(g) Results are available to me and may be obtained by contacting the investigators.

(h) I am aware that additional help will be offered to me by the investigator if there is any problem due to my participation in the study or if I wish alternative training.

This project has been reviewed and approved by the Oregon State University committee for Research Involving Human Subjects. The committee believes that the research procedures adequately safeguard the subject's privacy, welfare, civil liberties, and rights. My signature below indicates that I have read and understand the procedures described above and give my voluntary consent.

Name of subject

Signature of subject

Subject's phone number

Date signed

Social Security Number

Questions about this research should be directed to Dr. James Firth, Associate Professor of Counseling, Oregon State University, (541) 737-5973 or Louis Downs, research investigator, (541) 757-7440. Any other questions should be directed to Mary Nunn, OSU Research Office, (541) 737-0670.

I _____ have read the 6 page introductory material the "Structure of the Diagnostic and Statistical Manual of Mental Disorders" in preparation for the Diagnostic Experiment I am about to participate in.

Signature

Date

Appendix B**Pretest and Posttest Instrument From the DSM-III-R Casebook****Pretest Questionnaire**

Social Security Number_____

Age_____ Gender ___Male ___Female

Education Level_____ Undergraduate GPA_____

Cultural Origin (Race)_____

Months of computer experience_____

Weeks of classroom exposure to DSM diagnosis_____

Professional experience with DSM diagnosis (including internship)

Years_____ Months_____

Test Items (Case Studies)

Item A

An 85-year-old man is seen by a social worker at a senior citizen's center for evaluation of health-care needs for himself and his bedridden wife. He is apparently healthy, with no evidence of impairment in thinking or memory. He has been caring for his wife, but has been reluctantly persuaded to seek help because her condition has deteriorated, and his strength and energy has decreased with age.

A history is obtained from the subject and his daughter. He has never been treated for mental illness, and in fact has always claimed to be "immune to psychological problems" and to act only on the basis of "rational" thought. He had a moderately successful career as a lawyer and businessman. He has been married for 60 years, and his wife is the only person for whom he has ever expressed tender feelings, and is probably the only person he has ever trusted. He has always been extremely careful about revealing anything of himself to others, assuming that they are out to take something away from him. He refuses obviously sincere offers of help from acquaintances because he suspects their motives. He never reveals his identity to a caller without first questioning him as to the nature of his business. Throughout his life there have been numerous occasions on which he has displayed exaggerated suspiciousness, sometimes of almost delusional proportions (e.g., storing letters from a client in a secret safe deposit box so that he could use them as evidence in the event that the client attempted to sue him for mismanagement of an estate).

He has always involved himself in "useful work" during his waking hours, and claims never to have time for play, even during the 20 years he has been retired. He spends many hours monitoring his stock-market investments, and has had altercations with his broker when he suspected that an error on a monthly statement was evidence of the broker's attempt to cover up some fraudulent deal (Spitzer, et. al., 1989, pp. 163).

Item B

Hank Allen was charged with the murder of ten women. His wife, Jody, who eventually testified against him, had worked as his partner, luring victims to their deaths.

Wanting to further her husband's fantasy of finding the "perfect lover," Jody had accompanied him to shopping centers or county fairs and talked young girls into climbing into their customized van. Once inside, the victims were confronted by her husband, who held a handgun and bound them with adhesive tape. Most were teen-agers, though two of the final victims were adults; the youngest was 13. The oldest victim, 34, was a bartender who closed up late one night, went out to her car, then rolled down her window to talk to the couple, who had been inside drinking and who now approached her. The Allens kidnapped her and drove her back to their residence. While Jody sat inside watching an old movie on television, Hank assaulted his victim in the back of the van, scripting her to play the role of his teen-age daughter. When he was through, Jody rejoined him, and they drove away in the early morning hours, the radio blaring to drown out the sounds of Hank in the back of the van, strangling his victim to death. That evening they celebrated his birthday at a restaurant.

Most of Hank's victims were petite blonds like Jody and Hank's own daughter. All were sexually abused, then shot or strangled to death; several were buried in shallow graves. One, a pregnant 21-year-old hitchhiker (Jody was also pregnant at the time), was raped, strangled, and buried alive in sand.

Hank rated the sexual performance of each of his victims, and always made sure that Jody knew she was never number-one. Jody tried to redeem herself in the eyes of her difficult husband by submitting to his every demand. Even when she finally separated from him, she was unable to say "no." They had been apart for several months when Hank called her, asking that they get together one more time. She agreed, and that day they claimed their ninth and tenth victims.

Hank's violence was a legacy from his father. When he was born, his father, 19, was serving a prison sentence for auto theft and passing bad checks. A later conviction earned him a term for second-degree robbery, but he escaped. In an ensuing saga of recapture, escape, recapture and escape, he killed a police officer and a prison guard, blinding the latter by tossing acid into his face before beating him to death. A short time before he was executed, his father wrote: "When I killed this cop, it made me feel good inside. I can't get over how good it did make me feel, for the sensation was something that made me feel elated to the point of happiness...."

Often told that he was going to be just like his father when he grew up, Hank was 16 when he learned that his father had been captured and executed in a gas chamber after his mother betrayed his hiding place. Hank later confessed to the police: " Sometimes I [think] about blowing her head off....Sometimes I wanta put a shotgun in her mouth and blow the back of her head off...."

In a forensic psychiatric evaluation, Hank revealed that his mother was the object of his most intense sexual fantasy:

"I was gonna string her up by her feet, strip her, hang her up by her feet, spin her, take a razor blade, make little cuts, just little ones, watch the blood run out, just drip off her head. Hang her up in the closet, put airplane glue on her, light her up. Tattoo 'bitch' on her forehead...."

Hank's mother had beaten and mocked her son, a bed-wetter until age 13, calling him "pissy pants" in front of guests. One of her husbands punished him mercilessly, forcing him to drink urine and burning a cigar coal into his wrist. When his mother tried to intervene, his stepfather smashed her head into a plaster wall. From that point on, she joined in the active abuse of her children. As far back as he could remember, Hank had nightmares of being smothered by nylon stocking material and being strapped to a chair in a gas chamber as green gas floated into the room.

Hank began to burglarize with an older brother at age 7, and at age 12 was put on probation. A year later he was sent to the California Youth Authority

for committing "lewd and lascivious acts" with a six-year-old girl. As a teen-ager he faced charges of armed robbery and auto theft. A habitual truant, he was suspended from high school at 17 with F's in five academic subjects and F's in five categories of "citizenship." That same year he married for the first time.

Often knocked unconscious in fights, he was comatose twice, briefly at age 16, and for over a week at age 20. A computerized tomography brain scan revealed "abnormally enlarged sulci and slightly enlarged ventricles." A Halstead-Reitan neuropsychological battery and a Luria-Nebraska neuropsychological battery showed "damage to the right frontal lobe."

Hank married seven times. He beat each of his wives, sometimes badly. Most of the marriages lasted no more than a few months. One wife described him as "dominant," and said, "he's got to be in control." Another, who had had clumps of hair yanked from her head, called him "a Jekyll and Hyde." Yet another said he was "vicious." When she told him she wanted out, he took revenge by beating her parents. His first marriage ended when he beat his wife with a hammer. When she left him, she replaced his mother in his central fantasy. They had married five days after the birth of a baby daughter, and a custody battle ensued. In spite of his lengthy record of assaults, thefts, and parole violations, Hank won.

When he was 23, Hank went on a crime spree that eventually covered five states. Stealing license plates and cars, holding up bars and drugstores, he elude capture until caught and convicted for the armed robbery of a motel. Sent to prison for five years to life, he molested his six-year-old daughter for the first time during a conjugal visit.

Upon release, Hank went to live with his mother, who had not visited him during his three and a half years in prison. While there, he got involved with a woman whom he impregnated and whom he once kicked out of bed, literally, when she refused him anal intercourse. He chose not to marry her, she later recalled, as "he didn't want the responsibility." Thirteen days after she gave birth, he married another woman, his fifth wife. He was 28-years-old.

Hank and his fifth wife separated when he was released from parole. He took up residence with his 13-year-old daughter, whom he soon impregnated.

She had an abortion. His daughter had, by this time, replaced his first wife in his favorite fantasy, and he often raped her in the back of the van to which he and Jody would later lure victims. For the next six years, Hank assaulted her at least once a week. When a friend of hers arrived for a two-week visit, he also raped her.

He was 30 years old, and his divorce from his fifth wife had not been finalized when he moved in with Jody. By the time they met, Hank had been arrested on 23 separate occasions. The following summer Hank was fired from his job as a driver. He had been fired often, and it was an event that usually left him sexually impotent. An employer at the time termed him "inadequate." A week earlier he had celebrated his birthday by sodomizing his 14-year-old daughter. When his daughter finally informed authorities of the 6 years of abuse, felony charges were filed against Hank for incest, unlawful sexual acts, sodomy, and oral copulation. Hank responded by changing his name. Using the stolen driver's license of a state police officer, he obtained a new birth certificate and Social Security number, and he and Jody moved to another town.

Shortly before his final arrest, Hank, a gun enthusiast, owned a sem-automatic assault rifle, an automatic pistol, two revolvers, and a derringer. He was working as a bartender. A co-worker described him as a ladies' man, and said that women called him at work at all hours. After hanging up, he would rate them. Several women referred to him as "Mr. Macho." He was also a heavy drinker. Jody once cautioned him as he drank and drove that the combination was illegal. "Fuck the law," he answered. For his crimes, he eventually received multiple death sentences (Sptizer, et al., 1989, pp. 32-35).

Item C

Leon is a 45-year-old postal employee who was evaluated at a clinic specializing in the treatment of depression. He claims to have felt constantly depressed since the first grade, without a period of "normal" mood for more than a few days at a time. His depression has been accompanied by lethargy, little

or no interest or pleasure in anything, trouble concentrating, and feelings of inadequacy, pessimism, and resentment. His only periods of normal mood occur when he is home alone, listening to music or watching TV.

On further questioning, Leon reveals that he cannot ever remember feeling comfortable socially. Even before kindergarten, if he was asked to speak in front of a group of family friends, his mind would "go blank." He felt overwhelming anxiety at children's social functions, such as birthday parties, which he either avoided or, if he went, attended in total silence. He could answer questions in class only if he wrote down the answers in advance; even then, he frequently mumbled and couldn't get the answer out. He met new children with his eyes lowered, fearing their scrutiny, expecting to feel humiliated and embarrassed. He was convinced that everyone around him thought he was "dumb" or "a jerk."

As he grew up, Leon had a couple of neighborhood playmates, but he never had a "best friend." His school grades were good, but suffered when oral classroom participation was expected. As a teen-ager, he was terrified of girls, and to this day has never gone on a date or even asked a girl for a date. This bothers him, although he is so often depressed that he feels he has little energy or interest in dating.

Leon attended college and did well for a while, then dropped out as his grades slipped. He remained very self-conscious and "terrified" of meeting strangers. He had trouble finding a job because he was unable to answer questions in interviews. He worked at a few jobs for which only a written test was required. He passes a Civil Service exam at age 24, and was offered a job in the post office on the evening shift. He enjoyed this job since it involved little contact with others. He was offered, but refused, several promotions because he feared the social pressures. Although by now he supervises a number of employees, he still finds it difficult to give instructions, even to people he has known for years. He has no friends and avoids all invitations to socialize with co-workers. During the past several years, he has tried several therapies to help him get over his "shyness" and depression.

Leon has never experienced sudden anxiety or a panic attack in social situations or at other times. Rather, his anxiety gradually builds to a constant high level in anticipation of social situations. He has never experienced any psychotic symptoms (Spitzer, et al., 1989, pp. 53-54).

Item D

Paddy O'Brien is a 26-year-old bachelor, living with his mother and two older brothers on the family farm in the west of Ireland. He is interviewed as part of a family study of mental disorders being conducted in Ireland.

Paddy is described by his mother as having been a "normal" youngster up until the age of 14. He was average to slightly below average in schoolwork. He had friends he played with after school, and he helped his brothers and father with the chores around the farm. When he was 14, he began to "lose interest" in his schoolwork. His teacher noted that he was "staring into space" while in class, and rarely followed the work. Soon thereafter, his mother noticed that he no longer played with his friends after school, but would just come home and sit in front of the turf fire. It also became harder and harder to get him to do the farm chores. Sometimes he would come in and say the work was finished. Only hours later would they notice that only some of the cows had been milked, or only some of the eggs collected.

When he was 16, because his condition had become progressively worse, Paddy was withdrawn from school and was admitted to the county psychiatric hospital. The hospital records indicate that he was socially withdrawn and had a flat affect. It was not possible to interest him in ward activities. No psychotic symptoms could be elicited. Paddy has been in psychiatric care intermittently ever since that time. For the last year and a half, Paddy has been attending the local day center two days a week.

When interviewed by the research team, Paddy is observed to be an obese, rather disheveled young man. He replies to most questions with a "yes," or "no," or "could be." He denies any psychotic symptoms, feelings of depression or elation, or difficulty with appetite or energy. He does, however,

admit to unspecified problems with his "nerves," and problems in sleeping. On Probing, he admits to feeling uncomfortable around "people," except his family. Eye contact is poor: he looks at the floor during most of the interview. His affect is flat. Despite all attempts, the interviewer is unable to establish rapport with him.

According to Paddy's family, when he is not at the day center, he sits all day in front of the fire at home. Occasionally, he can be encouraged to help with a farm chore, but he usually stops after about 15 minutes and returns to his chair by the fire. Unless prompted, he will not wash or change his clothes. He refuses to attend any social functions, and his childhood friends have long ago stopped calling at the house for him.

At the day center, Paddy sometimes works for brief periods of time at simple tasks in occupational therapy, but then soon quits and goes to sit by himself in the day room. Both the family and staff note that he is quite aware of what is going on around him, as reflected by an occasional perceptive comment. Neither his family nor any other the psychiatric staff who care for Paddy has ever been able to elicit any psychotic symptoms (Spitzer, et al., 1989, pp. 207-208).

Item E

A 26-year-old unemployed woman was referred for admission to a hospital by her therapist because of intense suicidal preoccupation and urges to mutilate herself by cutting herself with a razor.

The patient was apparently well until her junior year in high school, when she became preoccupied with religion and philosophy, avoided friends, and was filled with doubt about who she was. Academically she did well; but later, during college, her performance declined. In college she began to use a variety of drugs, abandoned the religion of her family, and seemed to be searching for a charismatic religious figure with whom to identify. At times massive anxiety swept over her, and she found it would suddenly vanish if she cut her forearm with a razor blade.

Three years ago she began psychotherapy, and initially rapidly idealized her therapist as being incredibly intuitive and empathetic. Later, she became hostile and demanding of him, requiring more and more sessions, sometimes two in one day. Her life became centered on her therapist, to the exclusion of everyone else. Although her hostility toward her therapist was obvious, she could neither see it nor control it. Her difficulties with her therapist culminated in many episodes of cutting her forearm and threatening suicide, which led to the referral for admission (Spitzer, et al., 1989, p. 233).

Item F

Jane Berenson, a 36-year-old vice-president of a Detroit department store, responded to an advertisement describing a new clinic specializing in the treatment of sleep problems. She feels "mentally hyperactive" at bedtime, and is unable to stop thinking about significant experience of the day, particularly her interactions with dissatisfied customers. When she feels she has accomplished too little during a particular day, she feels she does not "deserve" to go to bed. Any evening excitement, e.g., an interesting movie or a lively party, leaves her unable to simmer down for hours thereafter. Occasionally, in the middle of the night, she awakens feeling wide awake and again finds herself ruminating about the day's events. When she sleeps poorly, she feels "high-strung" and tense the following day. The insomnia has worsened during the past year, coincident with more stress at work. She notes that she has not read a novel in over a year, an activity she previously enjoyed.

Her business involves occasionally "wining and dining" other executives, but she finds that late meals or alcohol intake aggravates the insomnia. She has noticed that on days when she has cocktails with dinner, she invariably awakens in the middle of the night, feeling wide awake and slightly sweaty. Business travel also worsens her sleep. She finds herself in a state of unrelieved overstimulation when her job requires "running from city to city" for extended periods.

Ms. Berenson was divorced three years ago after ten years of marriage. She has a wide circle of friends and enjoys socializing with them. Relaxing alone, however, has long been considered "dead time."

Both of her parents and a sister have had problems with alcohol. She is the only one in her family to be steadily employed.

During the last year she has been in once-a-week psychotherapy to try to understand "why I am so driven." This has not helped her insomnia. She has also tried sleeping pills, which leave her "hung over" the following day (Spitzer, et al., 1989, pp. 42-43).

Item G

Matthew is a 34-year-old single man who lives with his mother and works as an accountant. He seeks treatment because he is very unhappy after having just broken up with his girl friend. His mother had disapproved of his marriage plans, ostensibly because the woman was of a different religion. Matthew felt trapped and forced to choose between his mother and his girl friend, and since "blood is thicker than water," he had decided not to go against his mother's wishes. Nonetheless, he is angry at himself and at her and believes that she will never let him marry and is possessively hanging on to him. His mother "wears the pants" in the family, and is a very domineering woman who is use to getting her way. Matthew is afraid of her and criticizes himself for being weak, but also admires his mother and respects her judgment - "Maybe Carol wasn't right for me after all." He alternates between resentment and a "Mother knows best" attitude. He feels that his own judgment is poor.

Matthew works at a job several grades below what his education and talent would permit. On several occasions he has turned down promotions because he didn't want the responsibility of having to supervise other people or make independent decisions. He has worked for the same boss for ten years, gets on well with him, and is, in turn, highly regarded as a dependable and unobtrusive worker. He has two very close friends, who he has had since early

childhood. He has lunch with one of them every single workday and feels lost if his friend is sick and misses a day.

Matthew is the youngest of four children and the only boy. He was "babied and spoiled" by his mother and elder sisters. He had considerable separation anxiety as a child - difficulty falling asleep unless his mother stayed in the room, mild school refusal and unbearable home sickness when he occasionally tried "sleepovers." As a child he was teased by other boys because of his lack of assertiveness and was often called a baby. He has lived at home his whole life except for one year of college, from which he returned because of homesickness. His heterosexual adjustment has been normal except for his inability to leave his mother in favor of another woman (Spitzer, et al., 1989, pp. 123-124).

Item H

The patient is a 45-year-old lawyer who seeks treatment at his wife's insistence. She is fed up with their marriage: she can no longer tolerate his emotional coldness, rigid demands, bullying behavior, sexual disinterest, long work hours, and frequent business trips. The patient feels no particular distress in his marriage, and has agreed to the consultation only to humor his wife.

It soon develops, however, that the patient is troubled by problems at work. He is known as the hardest-driving member of a hard-driving law firm. He was the youngest full partner in the firm's history, and is famous for being able to handle many cases at the same time. Lately, he finds himself increasingly unable to keep up. He is too proud to turn down a new case, and too much of a perfectionist to be satisfied with the quality of work performed by his assistants. Displeased by their writing style, and sentence structure, he finds himself constantly correcting their briefs, and therefore unable to stay abreast of his schedule. People at work complain that his attention to details and inability to delegate responsibility are reducing his efficiency. He has had two or three secretaries a year for 15 years. No one can tolerate working for him for very long because he is so critical of any mistakes made by others.

When assignments get backed up, he cannot decide which to address first, starts making schedules for himself and his staff, but then is unable to meet them and works 15 hours a day., He finds it difficult to be decisive now that his work has expanded beyond his own direct control.

The patient discusses his children as if they were mechanical dolls, but also with a clear underlying affection. He describes his wife as a "suitable mate" and has trouble understanding why she is dissatisfied. He is punctilious in his manners and dress and slow and ponderous in his speech, dry and humorless, with a stubborn determination to get his point across.

The patient is the product of two upwardly mobile, extremely hard-working parents. He grew up feeling that he was never working hard enough, the he had much to achieve and very little time. He was a superior student, a "bookworm," awkward and unpopular in adolescent social pursuits. He has always been competitive and a high achiever. He has trouble relaxing on vacations, develops elaborate activities schedules for every family member, and becomes impatient and furious if they refuse to follow his plans. he likes sports, but has little time for them and refuses to play if he can't be at the top of his form. He is a ferocious competitor on the tennis courts and a poor loser (Spitzer, et al., 1989, pp. 80-81).

Item I

A wealthy and beautiful 34-year-old woman presented with a "marital problem." She was an heiress of a wealthy European family, and her husband was the president of a small importing company. She felt he was being insensitive and demanding; and he, apparently, accused her of being self-centered, impulsive, and a "compulsive" liar. Over the course of their ten-year marriage, each had had numerous affairs, most of which eventually came out into the open. Both would resolve to deal with their marital frustrations and to stop having affairs, and a brief period of reconciliation would follow; but soon one or the other would again surreptitiously begin an affair.

The patient also described a special problem that worried her and that she had never disclosed to her husband. Periodically she experienced the urge to walk into one of the more elegant department stores in the city and steal an article of clothing. Over the course of the previous three or four years she had stolen several blouses, a couple of sweaters, and a skirt. Since her husband's income alone was over \$250,000 a year and her investments worth many times that, she recognized the "absurdity" of her acts. She also indicated that what she stole was rarely very expensive and sometimes not even enough to her liking for her to wear.

The patient would become aware of the desire to steal something several days before she actually did so. The thoughts would increasingly occupy her mind until, on impulse, she would walk into a store, pluck an item off the rack, and stuff it under her coat or into a bag she happened to be carrying. Once out the door, she would experience a sense of relaxation and satisfaction; but at home she would feel anxious and guilty when she realized what she had done. She was caught on one occasion, but gave a long, involved story about intending to pay after she had gone elsewhere in the store and then "forgetting" to do so. She was released by the store security officers with a warning and suspiciously raised eyebrows.

She spent considerable time describing her own accomplishments, talents, and abilities. Her affairs, she said, proved that she was indeed beautiful and of superior "stock." She thought that she and her husband, who was handsome, aggressive, and successful, should be a perfect match. According to her, the problems with her husband stemmed from the little attention he paid her and the expectations he seemed to have that she should be at his beck and call. The frequent arguments they had upset her greatly, and thus it was her idea that they seek professional help. Regarding the charge that she was a compulsive liar, she admitted that she often found it easier to tell "white lies" than to face up to something "stupid" that she had done (Spitzer, et al., 1989, pp. 182-183).

Item J

The young lady, aged thirty, carefully dressed in black, who comes into the hall with short, shuffling steps, leaning on the nurse, and sinks into a chair as if exhausted, gives you the impression that she is ill. She is of slender build, her features are pale and rather painfully drawn, and her eyes are cast down. Her small, manicured fingers play nervously with a handkerchief. The patient answers the questions addressed to her in a low, tired voice, without looking up, and we find that she is quite clear about time, place and her surroundings. After a few minutes, her eyes suddenly become convulsively shut, her head sinks forward, and she seems to have fallen into a deep sleep. Her arms have grown quite limp, and fall down as if palsied when you try to lift them. She has ceased to answer, and if you try to raise her eyelids, her eyes suddenly rotate upwards. Needlepricks only produce a slight shudder. But sprinkling with cold water is followed by deep sigh; the patient starts up, opens her eyes, looks round her with surprise, and gradually comes to herself. She says that she has just had one of her sleeping attacks, from which she has suffered for seven years. They come on quite irregularly, often many in one day, and last from a few minutes to half an hour.

Concerning the history of her life, the patient tells us that...she was educated in convent schools, and passed the examination for teachers. As a young girl, she inhaled a great deal of chloroform, which she was able to get secretly, for toothache. She also suffered from headaches, until they were relieved by the removal of growths from the nose. She very readily became delirious in feverish illnesses. Thirteen years ago she took a place as governess in Holland, but soon began to be ill, and has passed the last seven years in different hospitals, except for a short interval when she was in a situation in Moravia.

It would appear from the statements of her relations and doctors that the patient has suffered from the most varied ailments, and been through the most remarkable courses of treatment. For violent abdominal pains and disturbances of menstruation, ascribed to stenosis of the cervical canal and retroflexion of

the uterus, recourse was had five years ago to the excision of the wedge supposed to cause the obstruction, and the introduction of a pessary. At a later period loss of voice and a contraction of the right forearm and the left thigh set in, and were treated with massage, electricity, bandaging, and stretching under an anaesthetic. Heart oppression and spasmodic breathing also appeared, with quickly passing disablements of various sets of muscles, disturbances of urination, diarrhea, and unpleasant sensations, now in one and now in another part of the body, but particularly headaches. Extraordinarily strong and sudden changes of mood were observed at the same time, with introspection and complaints of want of consideration in those about her and in her relations, although the latter had made the greatest sacrifices. Brine baths, Russian baths, pine-needle baths, electricity, country air, summer resorts, and finally, residence on the Riviera - everything was tried, generally with only a brief improvement or with none at all.

The immediate cause of the patient being brought to the hospital was the increase in the "sleeping attacks" two years ago. They came on at last even when the patient was standing, and might continue for an hour. The attacks continued in the hospital, and spasmodic breathing was also observed, which could be influenced by suggestion.

After spending eight months here, the patient went away at first to her sister's. But after a few months she had to be taken to another asylum, where she stayed about a year, and then, after a short time spent with her family, came back to us.

During her present residence here, so-called "great attacks" have appeared, in addition to her previous troubles. We will try to produce such an attack by pressure on the very sensitive left ovarian region. After one or two minutes of moderately strong pressure, during which the patient shows sharp pain, her expression alters. She throws herself to and fro with her eyes shut, and screams to us loudly, generally in French, not to touch her. "You must not do anything to me, you hound, cochon, cochon!" She cries for help, pushes with her hands, and twists herself as if she were trying to escape from a sexual assault. Whenever she is touched, the excitement increases. Her whole body

is strongly bent backwards. Suddenly the picture changes, and the patient begs piteously not to be cursed, and laments and sobs aloud. This condition, too, is very soon put an end to by sprinkling with cold water. The patient shudders, wakes with a deep sigh, and looks fixedly round, only making a tired, senseless impression. She cannot explain what has happened.

The physical examination of the patient shows no particular disturbances at present, except the abnormalities already mentioned. There is only a well-marked weakness, in consequence of which she often keeps to her bed or lies about. All her movements are limp and feeble, but there is no actual disablement anywhere. She often sleeps very badly. At times she wanders about in the night, wakes the nurses, and sends for the doctor. Her appetite is very poor, but she has a habit of nibbling between her meals at all kinds of cakes, fruit, and jam, which are sent at her request, by her relations.

With her growing expertness in illness, the emotional sympathies of the patient are more and more confined to the selfish furthering of her own wishes. She tries ruthlessly to extort the most careful attention from those around her, obliges the doctor to occupy himself with her by day or by night on the slightest occasion, is extremely sensitive to any supposed neglect, is jealous if preference shown to other patients, and tries to make the attendants give in to her by complaints, accusations, and outbursts of temper. The sacrifices made by others, more especially by her family, are regarded quite as a matter of course, and her occasional prodigality of thanks only serves to pave the way for new demands. To secure the sympathy of those around her, she has recourse to more and more forcible descriptions of her physical and mental torments, [dramatic] exaggeration of her attacks, and the effective elucidation of her personal character. She calls herself the abandoned, the outcast, and in mysterious hints makes confession of horrible, delightful experiences and failings, which she will only confide to the discreet bosom of her very best friend, the doctor (Spitzer, et al., 1989, pp. 450-451).

Item K

Bob, a 21-year-old man, comes to the psychiatrist's office, on the advice of his college counselor, accompanied by his parents. He begins the interview by announcing that he has no problems. His parents are always overly concerned about him, and it is only to get them "off my back" that he has agreed to the evaluation. "I am dependent on them financially, but not emotionally."

The psychiatrist was able to obtain the following story from Bob and his parents. Bob had apparently spread malicious and false rumors about several of the teachers who had given him poor grades, implying that they were having homosexual affairs with students. This, as well as increasingly erratic attendance at his classes over the past term, following the loss of a girl friend, prompted the school counselor to suggest to Bob and his parents that help was urgently needed. Bob claimed that his academic problems were exaggerated, his success in theatrical productions was being overlooked, and that he was in full control of the situation. He did not deny that he spread the false rumors, but showed no remorse or apprehension about possible repercussions for himself.

Bob is a tall, stylishly dressed young man with a dramatic wave in his hair. His manner is distant, but charming, and he obviously enjoys talking about a variety of intellectual subjects or current affairs. However, he assumes a condescending, cynical, and bemused manner toward the psychiatrist and the evaluation process. He conveys a sense of superiority and control over the evaluation.

Accounts of Bob's development were complicated by his bland dismissals of its importance and by the conflicting accounts about it by his parents. His mother was an extremely anxious, immaculately dressed, outspoken woman. She described Bob as having been a beautiful, joyful baby, who was always extremely gifted and brilliant. She recalled that after a miscarriage, when Bob was one year old, she and her husband had become even more devoted to his care, giving him "the love for two." The father was rugged-looking, soft-spoken, successful man. He recalled a period in Bob's early life when they had been very close, and he had even confided in Bob

about very personal matters and expressed deep feelings. He also noted that Bob had become progressively more resentful with the births of his two siblings. The father laughingly commented that Bob "would have liked to have been the only child." He recalled a series of conflicts between Bob and authority figures over rules, and that Bob had expressed disdain for his peers at school, and for his siblings.

In his early school years, Bob seemed to play and interact less with other children than most others do. In fifth grade, after a change in teachers, he became arrogant and withdrawn and refused to participate in class. Nevertheless, he maintained excellent grades. In high school he had been involved in an episode similar to the one that had led to the current evaluation. At that time he spread false rumors about a classmate with whom he was competing for a role in the school play.

In general, it became clear that Bob had never been "one of the boys." He liked dramatics and movies, but had never shown an interest in athletics. He always appeared to be a loner, though he did not complain of loneliness. When asked, he professed to take pride in "being different" from his peers. He also distance himself from his parents and often responded with silence to their overtures for more communication. His parents felt that behind his guarded demeanor was a sad, alienated, lonely, young man. Though he was well known to classmates, the relationships he had with them were generally under circumstances in which he was looked up to for his intellectual or dramatic talents.

Bob conceded that others viewed him as cold or insensitive. He readily acknowledged these qualities, and that he had no close friends; but he dismissed this as unimportant. This represented strength to him. He went on to note that when others complained about these qualities in him, it was largely because of their own weakness. In his view, they envied him and longed to have him care about them. He believed they sought to gain by having an association with him.

Bob had occasional dates, but no steady girl friends. Although the exact history remains unclear, he acknowledged that the girl whose loss seemed to

have led to his escalating school problems had been someone whom he cared about. She was the first person with whom he had had a sexual relationship. The relationship had apparently dissolved after she had expressed an increasing desire to spend more time with her girl friends and to go to school social events (Spitzer, et al., 1989, pp. 197-198).

Item L

Clara Cole, a 34-year-old, black mother of three children, aged 13, 11, and 5, was referred by a juvenile court for psychiatric evaluation pending termination of her parental rights for her two oldest children, Tyrone and Tanya.

Ms. Cole described her Tyrone as a difficult child who had been hyperactive from birth. She has trouble toilet training him and difficulty disciplining him. When he was two and a half, Tyrone was treated for second-degree burns on his ankles, posterior calves, buttocks, and penis. Ms. Cole was alone with him at the time, but claimed this was an accidental injury that could have happened to anyone. She explained that he had turned on the hot water by himself when she was out of the room.

Since that injury there have been many other multiple injuries to all three children. Each of them has been observed to have bruises, welts, and marks consistent with their stories that they were hung upside down and beaten with rubber hoses, the youngest child, Winnie, has been treated for ongoing hallucinations of a female voice telling her to hit other children. Both of the two older children are in residential treatment facilities at this time because of the severity of their behavior problems. When each of these children entered residential treatment, they were frightened, particularly of adult women. Both Tyrone and Tanya have said that their mother threatened further physical abuse if they told anyone what was going on at home.

Ms. Cole is insulted that the juvenile court is involved in her case. She says her children were abused in the past by her ex-husband, who was also physically and emotionally abusive of her. She denies she has ever abuse them and cannot explain why the children appear to be selectively frightened of

women. Ms. Cole acknowledges that she disciplined her children by whipping them with a belt, but denies that she has ever hung them upside down. She believes that her oldest son began making up stories about how she mistreated them to get back at her for setting limits on him. She does not believe that the younger two children verified his stories (which they did). Ms. Cole says that recently, when she understood that her parental rights were to be terminated, she stopped spanking her youngest child for fear of losing custody of her as well. Ms. Cole admits disciplining the children to the point of leaving bruises and welts on them for infractions such as talking back to her, not coming home immediately after school, not getting grades as good as she expected, and being disrespectful.

Ms. Cole's parents were both alcoholic, but she has never had any problems with drugs or alcohol. Her mother beat her frequently without telling her what she had done wrong. She recalls being frightened that her mother would "get weird on me and hit me with whatever was handy." At times, she needed stitches from injuries caused by her mother. These were taken care of at home, as her family was too poor to afford medical help. Her mother accused her of being responsible for the misbehavior of her siblings. She recalls that by the time she was in high school, she was so hostile toward women that she was assigned only male teachers. Nevertheless, she obtained good grades and was active in the Reserve Officers Training Corps.

Ms. Cole is steadily employed as a supervisor in a shipping department, a job she has held for three years. She has had two promotions during this time. She is a large, attractive, neatly dressed woman. On formal mental status exam, she appears to be above-average intelligence, is fully oriented, and has no difficulty with concentration or abstractions. Although she is very hostile, and critical of the child protective services and juvenile court, she is calm and pleasant with the examiner.

Ms. Cole's ex-husband is frightened of her. He says it is true that she has a bad temper, but she was the one who instigated the physical fights that they had when they were married. He would like to have custody of the children, but will not seek custody unless her parental rights have already been terminated.

Ms. Cole does not see any need for psychiatric intervention. She does not believe that she has a problem for which she needs treatment. She also says that although Tyrone has had multiple behavior problems for a number of years, if his custody were returned to her, she would stop his psychotherapy. She believes his problems would be resolved by replacing him back in her care. If that proves not to be the case, she is prepared to give custody to Social Services (Spitzer, et al., 1989, pp. 109-110).

Item M

Alexi, a somewhat overweight 23-year-old man, who looks more like 17, is brought to the Moscow District Mental Health Center by his father for an evaluation. Alexi's father had been estranged from his wife and child since the patient was five years old. He is a retired army officer who knows about Alexi mostly from his wife's letters. Until four weeks ago, Alexi lived with his mother in a single room, sharing kitchen and bath with three other families in a communal apartment. Four weeks ago, his mother died suddenly of a heart attack. A neighbor, who had known both Alexi and his mother for years, somehow found means to communicate this new to the father, who came to bury his wife, and who had spent the last four weeks with his son.

The father is alarmed by Alexi's condition. According to the father, Alexi spends all of his time alone at home. He sleeps during the day and spends his nights reading and taking copious notes from "strange books." He is a vegetarian, and amazes his father by his total lack of interest in any food other than boiled potatoes and sweet tea. According to the neighbor, on the morning following his mother's death, Alexi shaved his head, referring vaguely to some Eastern rites of mourning. This was incomprehensible to the father until he explored Alexi's bookshelves, where he discovered numerous books on Eastern religion, natural healing, and astronomy.

The father has tried to engage his son in talking about the future, but Alexi has no interest in education, work, or the "things all young people should be interested in." Over the last four weeks, all the father's attempts to discuss

his son's future have caused Alexi to become irritable and either withdraw to his bed or leave the apartment, to wander the streets until his father is asleep.

The young man is interviewed by a young psychiatrist, who starts the interview with great enthusiasm and a friendly attitude. Soon, the psychiatrist is amazed at failing to make any connection with the patient, who remains impassive, virtually silent, and answers the psychiatrist's questions only when presented with the same question at least twice. Most of his answers are monosyllabic, and his face betrays no emotion. He is informed that he will be admitted to the hospital for "further evaluation." He greets this news with only one comment: "Will you allow my father to bring me some of my books?"

In the hospital, the young man's psychiatrist and other members of the staff fail to make any emotional contact with him, and describe him as "cold, but not hostile." He appears to be profoundly disinterested in his surrounding, and shows some air of contentment only when he is allowed to read on of the books his father has brought from home.

On further investigation of school records, information from the patient's pediatrician and elementary - and secondary - education teachers is located. the young man is described as having been a "perfect infant;" he could play in his crib for hours alone, and was "not demanding on his mother." His mother once remarked to the pediatrician, "this kid never even tried to climb out of the playpen." Elementary - and secondary school teachers described the young man as a loner who stayed away from other children and who was able to work academically with an average level of achievement up until the 7th grade. At that time, he gradually lost interest in his studies, but was never defiant when scolded by his teachers or teased by his classmates, who called him "retard." He absolutely refused to take physical education and, only after being confronted by the principal, mumbled, "I don't want to undress."

Throughout most of the young man's life, his mother worked the night shift at a factory and slept most of the day. The boy never complained about being left alone most of the time. At the age of 11, he was given a pet, a little hamster, to which he became attached in a matter of hours. He took great care of his pet; but when the pet died three months later, he showed an astonishing

lack of emotion. When his mother brought him several tropical fish to console him, he never looked at the aquarium, refused to feed the fish, and impassively watched them die one by one.

In the hospital, where the patient remained for four weeks, he was given small doses of neuroleptics, to which he immediately developed an acute dystonic reaction. A senior physician from another department was called as a consultant. He evaluated the record, and having failed to elicit more than a few monosyllabic responses from the patient over a 40-minute attempted interview, the consultant suggested eliminating all medication. A diagnosis that qualified the patient for permanent disability was established, and he was referred to a sheltered workshop. Several days following his discharge, when his father came to the ward to collect some of the things his son left behind, one of the nurses remarked: "I already have difficulty remembering his name. I'm sure we will not remember his face in a week or so" (Spitzer, et al., 1989, pp. 424-425).

Item N

John is a 50-year-old retired policeman who seeks treatment a few weeks after his dog has been run over and died. Since that time he has felt sad, tired, and has had trouble sleeping and concentrating.

John lives alone, and has for many years had virtually no conversational contacts with other human beings beyond a "Hello" or "How are you?" He prefers to be by himself, finds talk a waste of time, and feels awkward when other people try to initiate a relationship. He occasionally spends some time in a bar, but always off by himself and not really following the general conversation. He reads newspapers avidly, and is well informed in many areas, but takes no particular interest in the people around him. He is employed as a security guard, but is known by fellow workers as a "cold fish" and a "loner." They no longer even notice or tease him, especially since he never seemed to notice or care about their teasing anyway.

John floats through life without relationships except for that with his dog, which he dearly loved. At Christmas he would buy the dog elaborate gifts and,

in return, would receive a wrapped bottle of scotch that he bought for himself as a gift from the dog. He believes that dogs are more sensitive and loving than people, and he can, in return, express toward them a tenderness and emotion not possible in his relationships with people. The loss of his pets are the only events in his life that have caused him sadness. He experienced the death of his parents without emotion, and feels no regret whatever at being completely out of contact with the rest of his family. He considers himself different from other people, and regards emotionality in others with bewilderment (Spitzer, et al., 1989, pp. 249-250).

Item Q

A 34-year-old psychiatrist is 15 minutes late for his first appointment. He had recently been asked to resign from his job in a mental health center because, according to his boss, he had frequently been late for work and meetings, missed appointments, forgot about assignments, was late with his statistics, refused to follow instructions, and seemed unmotivated. The patient was surprised and resentful - he thought he had been doing a particularly good job under trying circumstances and experienced his boss as excessively obsessive and demanding. Nonetheless, he reported a long-standing pattern of difficulties with authority.

The patient had a childhood history of severe and prolonged temper tantrums that were a legend in his family. He had been a bossy child who demanded that other kids "play his way" or else he wouldn't play at all. With adults, particularly his mother and female teachers, he was sullen, insubordinate, oppositional, and often unmanageable. He had been sent to an all-boys' preparatory school that had primarily male teachers, and he gradually became more subdued and disciplined. He continued, however, to stubbornly want things his own way and to resent instruction or direction from teachers. He was a brilliant but erratic student, working only as hard as he himself wanted to; and he "punished" teachers he didn't like by not doing their assignments. He

was argumentative and self-righteousness when criticized, and claimed that he was not being treated fairly.

The patient is unhappily married. He complains that his wife does not understand him and is a "nitpicker." She complains that he is unreliable and stubborn. He refuses to do anything around the house and often fails to complete the few tasks he has accepted as within his responsibility. Tax forms are submitted several months late; bills are not paid. The patient is sociable and has considerable charm, but friends generally become annoyed at his unwillingness to go along with the wishes of the group (for example, if a restaurant is not his choice, he may sulk all night or "forget" to bring his wallet) (Spitzer, et al., 1989, pp. 107-108).

Item P

Maryann is an attractive, 35-year-old single woman, originally from San Diego, now working as a magazine editor and living by herself in a deteriorating Boston neighborhood. She was referred for psychotherapy by her female family doctor, who suggested she needed to work on problems in her relationships with men. Maryann resisted following through on the referral for a year, saying, "I don't like getting help. I like giving it."

When interviewed, Maryann appeared to be highly intelligent; she was affable and articulate and spoke in a breathy, girlish voice. She has metal-black hair, was dressed in all black - leather skirt and jacket and black top - and wore "punkish" glasses. She said, at the beginning of the interview, that she didn't want a male therapist because she was mistrustful of men, who, in her experience, wanted only to exploit women. However, with the exception of her family doctor, she had no close women friends.

Her story was that she had just extricated herself from a "destructive" relationship with a man, "my outlaw love," who was a heroin addict; and she was fighting her wish to return to him. Once, four years earlier, he had hit her and made her cry, but she told him that if he did that again, she'd leave, and it never recurred. She claimed she was not frightened of him, and actually

blamed herself for attacking her. "I often tell him things he should know about himself, and he gets furious. I only do it to motivate him. I hit his soft spot."

Her lover's addiction persisted, and Maryann continued to support him financially whenever he needed help. She said she received many indications that this relationship could not make her happy. The man had gone out with other women while dating Maryann, served a brief jail sentence for selling drugs, and never wanted to engage in mutually entertaining activities, except sex, which was enjoyable. Maryann had gone to a university, but her lover had never completed high school. She felt that he was like a little child who needed mothering. He would tell her to get lost when she insisted he stop using drugs, but she continued to call him regularly in spite of his ungrateful behavior. She felt resentful and embittered because of all she had done for him, but always helped him when he, typically, came back to her, late at night, asking for money or assistance. As a result, she said she felt "more like a Mother Theresa than a girl friend."

Maryann is now seeing another "exciting" man, also a substance abuser. Although she considers herself "left-wing," her new friend is a collector of Nazi memorabilia. She knew that he treated his previous girl friend cruelly by being unfaithful and abusive, but didn't think about whether this might happen to her. She has seen this man on and off for a year. He insisted he wanted a close relationship, but did not tell her he was seeing one of her acquaintances on the side. When she found out about this she was very upset but continues to have an intense interest in him. A number of nicer men who had monogamous intentions have frequently tried to date her, but she has avoided them because they all seemed "boring."

In her other relationships, Maryann always gives help, but never asks for it, even when she is in real need. Most of her friends and ex-boyfriends have been drug addicts, or ex-addicts. She herself has never abused drugs. She often visits these people in jail and offers to help them; but when they are released, they hardly ever visit her.

At her job Maryann is hardworking and good at solving disputes, but she has sometimes gotten into trouble with her boss for arranging to use the

magazine's resources to raise money for needy groups. She feels that her female colleagues "gang up on her" because of envy of her abilities and capacity for hard work, in spite of all the benefits that she has helped them obtain.

Maryann is the oldest of four children, and often had to grudgingly care for her young sibs. She became a "Good Two Shoes," while her younger brothers were permitted to "act up." In church and school she did well and won many awards, until, in her teens, she rebelled and left home. Her parents predicted she would "go to hell." She went through a period of "sexual liberation" during which she had about fifty lovers, often in one-night stands, which she rarely enjoyed "because I didn't love those guys." As a young adult she was always involved in some worthy cause for the underprivileged, the poor or the politically disadvantaged (Spitzer, et al., 1989, pp. 175-176).

Item Q

The patient, a single, unemployed, 19-year-old male was referred for psychiatric evaluation before undergoing orthognathous surgery for a protruding mandible. The procedure was to create a new facial look and improve both function and aesthetics. The evaluation was requested to determine if there were any psychiatric contraindications to surgery.

The patient says that his jaw has been protruding since childhood; he feels it may have protruded because as a child he frequently stuck his tongue out, and "maybe this stretched my jaw." He knows his molars are in place, but the teeth on the side are "pointed." His friends don't tease him about his jaw, but they do say, "You got a mug," and this upsets him. He describes himself as shy and feels it is partly from this self-consciousness about his jaw. He has difficulty talking and eating, as his teeth underbite and his tongue protrudes; thus, he cannot bite, but has to tear, his food. He has wanted to have his jaw fixed for a long time, but was "too shy" to ask about it. He says that, as a result, he hasn't seen a dentist for the last four years. He is aware that some teeth will have to be removed and that he will have his jaw wired for six weeks and will be

on a liquid diet. He is uneasy about being unable to eat solid food. He hopes the surgery will correct his chewing problem, and that he will feel better about his face and become more comfortable with other people.

The patient did well in school until he reached high school then he started to cut classes and dropped out of the tenth grade. He worked for two years as a security guard. He is now unemployed, but wants to go back to school and become an auto mechanic.

The patient is the third in a family of eight children. His parents separated when he was 14 years old. he lives with his mother and siblings. He argues with his siblings about doing household chores and, as a result, doesn't spend much time with his family; he just comes and goes and spends time with friends. He restrains himself from telling his friends not to comment on his "mug," preferring to "keep it inside." He hopes that if the operation is successful, his friends will stop remarking on his looks.

When examined, the young man was noted to have mild acne and a very visibly protruding jaw with an underbite. His manner was somewhat awkward, There were no gross abnormalities of thinking, perception, or overt behavior. He denied ever having any problems with mood, sleeping, eating, or in the use of alcohol or other drugs (Spitzer, et al., 1989, p. 172).

Item R

A 28-year-old junior executive was referred by a senior psychoanalyst for "supportive" treatment. She had obtained a master's degree in business administration and moved to California a year and a half earlier to begin work in a large firm. She complained of being "depressed" about everything: her job, her husband, and her prospects for the future.

She had had extensive psychotherapy previously. She had seen an "analyst" twice a week for three years while in college, and a "behaviorist" for a year and a half while in graduate school. Her complaints were of persistent feelings of depressed mood, inferiority, and pessimism, which she claims to have had since she was 16 or 17 years old. Although she did reasonably well

in college, she consistently ruminated about those students who were "genuinely intelligent." She dated during college and graduate school, but claimed that she would never go after a guy she thought was "special," always feeling inferior and intimidated. Whenever she saw or met such a man, she acted stiff and aloof, or actually walked away as quickly as possible, only to berate herself afterward and then fantasize about him for many months. She claimed that her therapy had helped, although she still could not remember a time when she didn't feel somewhat depressed.

Just after graduation, she married the man she was going out with at the time. She thought of him as reasonably desirable, though not "special," and married him primarily because she felt she "needed a husband" for companionship. Shortly after their marriage, the couple started to bicker. She was very critical of his clothes, his job, and his parents; and he, in turn, found her rejecting, controlling, and moody. She began to feel that she had made a mistake in marrying him.

Recently she has also been having difficulties at work. She is assigned the most menial tasks at the firm and is never given an assignment of importance or responsibility. She admits that she frequently does a "slip-shod" job of what is given her, never does more than is required, and never demonstrates any assertiveness or initiative to her supervisors. She views her boss as self-centered, unconcerned, and unfair, but nevertheless admires his success. She feels that she will never go very far in her profession because she does not have the right "connections," and neither does her husband; yet she dreams of money, status, and power.

Her social life with her husband involves several other couples. The man in these couples is usually a friend of her husband. She is sure that the women find her uninteresting and unimpressive, and that the people who seem to like her are probably no better off than she.

Under the burden of her dissatisfaction with her marriage, her job, and her social life, feeling tired and uninterested in "life," she now enters treatment for the third time (Spitzer, et al., 1989, pp. 37-40).

Scoring Key Pretest and Posttest Items

Item A

DSM-III-R Diagnosis:

Axis I: No Diagnosis or Condition

Axis II: Paranoid Personality Disorder, Moderate

Axis II features (not adequate to full diagnosis):

Schizoid Personality Traits

(Spitzer, et al., 1989, p. 164)

Item B

DSM-III-R Diagnosis:

Axis I: Sexual Sadism

Axis II: Personality Disorder Not Otherwise Specified

(Sadistic Personality Disorder)

(Antisocial Personality Disorder)

(Spitzer, et al., 1989, p. 36)

Item C

DSM-III-R Diagnosis:

Axis I: Dysthymia, Primary Type, Early Onset

Social Phobia, Generalized Type

Axis II: Avoidant Personality Disorder

(Spitzer, et al., 1989, p. 55)

Item D

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Schizotypal Personality Disorder, Severe

(Spitzer, et al., 1989, p. 209)

Item E

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Borderline Personality Disorder

(Spitzer, et al., 1989, p. 233)

Item F

DSM-III-R Diagnosis

Axis I: Primary Insomnia

Axis II: No Diagnosis

Axis II traits (not adequate to diagnosis):

Obsessive Compulsive Personality Traits

(Spitzer, et al., 1989, p. 45)

Item G

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Dependent Personality Disorder, Mild

(Spitzer, et al., 1989, p. 124)

Item H

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Obsessive Compulsive Personality Disorder, Moderate
(Spitzer, et al., 1989, p. 81)

Item I

DSM-III-R Diagnosis

Axis I: Marital Problems

Kleptomania

Axis II: No Diagnosis

Axis II Traits (not adequate to diagnosis):
Narcissistic Personality Traits

(Spitzer, et al., 1989, p. 184)

Item J

DSM-III-R Diagnosis

Axis I: Somatization Disorder

Axis II: Histrionic Personality Disorder

Axis II features (inadequate to full diagnosis):
Narcissistic Personality Traits

(Spitzer, et al., 1989, p. 453)

Item K

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Narcissistic Personality Disorder (Provisional)
(Spitzer, et al., 1989, p. 199)

Item L

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Personality Disorder Not Otherwise Specified (Sadistic Personality Disorder)

Axis II features (inadequate to full diagnosis):

Antisocial Personality Traits.

(Spitzer, et al., 1989, p. 111)

Item M

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Schizotypal Personality Disorder, Severe

Axis I features (not adequate to full diagnosis):

Pervasive Developmental Disorder Traits

(Spitzer, et al., 1989, p. 427)

Item N

DSM-III-R Diagnosis

Axis I: Adjustment Disorder with Depressed Mood

Axis II: Schizoid Personality Disorder

Axis II Characteristics (insufficient for full diagnosis):

Schizotypal Personality Traits

(Spitzer, et al., 1989, p. 250)

Item Q

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Passive-Aggressive Personality Disorder, Moderate
(Spitzer, et al., 1989, p. 108)

Item P

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: Personality Disorder Not Otherwise Specified
(Self Defeating Personality Disorder)
(Spitzer, et al., 1989, p. 177)

Item Q

DSM-III-R Diagnosis

Axis I: No Diagnosis or Condition

Axis II: No Diagnosis
(Spitzer, et al., 1989, p. 173)

Item R

DSM-III-R Diagnosis

Axis I: Dysthymia, Primary Type, Early Onset (p. 232)

Axis II: No Diagnosis.

Axis II features (inadequate to diagnose):
NOS (Self Defeating Personality Features)
(Spitzer, et al., 1989, p. 41)

Pretest and Posttest Answer Sheet Example**Item # A Answer Sheet**

Axis II Diagnosis(es)_____

Axis II Features (inadequate to full diagnosis)

Axis I Diagnosis(es)_____

Axis I Features (inadequate to full diagnosis)

No Diagnosis (place an X here if no diagnosis exists)_____

Appendix C

Instructions for Participants of Problem-Solving vs. Decision Tree for DSM Diagnosis Experimental Teaching Sessions

Session # 1 (pretest)

Subjects will meet at 8:00 a.m. in the testing room predesignated by the school site. The proctor will briefly greet subjects and explain procedures.

Proctor: Welcome to the experimental study of psychodiagnostic instruction. My name is, xxxx, and I will be one of your monitors for this study. For the purpose of this study, the exact variables to be measured will not be disclosed until after the experiment has been finished.

I or another monitor will be with you throughout the experiment. I will start this session with a brief orientation to the day and instructions for this session. Each session will begin with instructions from the monitor. If, after instructions have been given, you need to ask questions about procedure, please do. However, I cannot answer questions about content in any session. All data gathered from subjects in this study will be held in strict confidence.

During the next four hours, you will take a pretest, see a brief film, work for two hours on a computer learning program, and take another test. Breaks have been built in between each session, so please take advantage of them. Except for the two hours you will spend in the computer lab, with the learning program, I will ask that you do not leave the room until the break unless it is an absolute emergency. You may take breaks during the computer lab session, as it lasts for two hours.

Today you will begin with a pretest. The test is designed to document how much knowledge individuals in this group may already have about DSM Axis II diagnoses, personality disorders. In front of you, on the table is a folder, inside of which is the exam. The examination contains a questionnaire. Please open the folder and fill out page 1 completely. Do not turn the page.

Pause to wait for demographic questionnaire to be completed (estimated time: 2 min).

Proctor: The test you are about to take contains 9 case studies of potentially disordered individuals drawn from the DSM-III-R Casebook. Each case will be presented separately and will be followed by a one-page answer sheet for that case study. After you have read the case study, please fill out the answer sheet. (using a flip chart, proctor turns to the illustrated page, designated, "answer sheet illustration").

On each answer sheet there are two lines designated for identification of any personality disorder diagnosis or diagnoses that you believe exists. This space is to be used only to identify personality syndromes that you believe meet an adequate number of criteria from the DSM-III-R to be considered a full diagnosis.

Next, you find spaces provided for Axis II features (inadequate to full diagnosis). If you believe that the individual in the case study has features of a personality disorder, but does not evidence enough signs and symptoms to qualify for a full diagnosis, the diagnosis which has the features you recognize, that do not qualify for full diagnosis, should be written in this space.

Next, you will find space for Axis I features (inadequate to full diagnosis). If you have already taken course work enough that you recognize symptomology inadequate for an Axis I diagnosis but present in the client of the case study, the name should be written here.

Next you will find a space for Other Diagnosis (i.e.: Axis I Diagnosis). If you have already taken course work enough that you recognize a diagnosable syndrome from Axis I, you should write it in this space. If you have not yet received enough information to diagnose, but recognize that some of the symptoms or signs that the case study client displays belong to some other syndrome than personality disorders, Axis II diagnoses, place an "X" on the first line of Other Diagnosis.

Finally, you will see a line which says No Diagnosis. If you believe that the case study client does not show signs or syndromes of a psychological or personality disorder, place an "X" in the space provided next to No Diagnosis.

Case studies have been carefully selected to represent a wide range of cases from simple to complex. Cases may contain, no diagnosis, one diagnosis - either Axis I or Axis II), more than one full diagnoses - including either more than one Axis I, more than one Axis II, or a combination of Axis I and Axis II diagnoses, as well as features of Axis I or Axis II categories that do not meet the full criteria for diagnoses. Write in as many as you believe exist in each category, or check no diagnosis. This is a pretest. Its purpose is only to explore the extent of your knowledge of diagnosis before the experiment. You are to perform at whatever level you can; difficulty with answers on part or all of the test is expected for anyone not already trained in DSM diagnosis.

As soon as you have completed the first case, quickly go to the second and continue exactly as you did for the last. Remember, there are 9 cases to read and diagnose. You will have 30 minutes to complete the pretest. After you are finished, please close your folder with all test material inside and wait quietly for others to finish. You will have a break after you receive instructions as to where to go for the next session.

You may begin.

Session #1: Closing instructions

Proctor: It is (time of day). You are finished with the first phase of this study. For the next phase, you have been assigned to one of two groups. You will see four lists posted for your convenience on the walls both here and in the hallway. Your group assignment is listed on each of those four sheets. As you leave, please go to one of the sheets and look for your social security number. It will be listed under a room number. If you are not sure where you should go, please ask myself or the other monitor (introduce other monitor). Please go to that room after you have taken a short break. Because of room schedules, only 15 minutes have been allotted for your break and to go to the assigned room.

Please be in the room and ready to go at 9:00 a.m. sharp. Do not discuss the test with other participants until after the entire study is finished.

Please leave the folder with your test on the desk as you leave. Thank you. You are dismissed. We will see you at 9:00.

Session #2 (film)

Proctors be sure to count your group and make sure all are present a few minutes before 9:00. If there are missing subjects, please go to the hall and look for them.

Proctor: Can I have your attention, please. This next phase of the study will be short, approximately 15 minutes. Please stay for the whole video and quietly watch, respecting the need of others in the room. After I dim the lights, the movie will be shown. Afterward, I will turn up the lights and give you the next set of instructions. Enjoy the video.

Approximately 35 minutes for video. Turn up the lights and give instructions.

Proctor: Now that we have finished viewing the video we will go to the computer lab for the two hour computer-assisted class. I will accompany the group so that it is easy to find the next room. It will take us a few minutes to get to that room. Please wait for instructions before taking a break. You will have up to 2 hours in the computer lab and may take a breaks as needed.

We will start the computer-assisted learning lab at 9:30. Please be prompt. Remember, you will have the opportunity for breaks as needed throughout the two hours. However, I ask that you make breaks brief, as you will need most of the two hours to complete the computer-assisted learning lab. Please come with me.

Session #3 (computer-assisted learning lab)

Proctors, be sure to count subjects a few minutes before you begin and look for stragglers.

Proctor: May I have your attention. Please make sure your computer is booted up. If you need help, please raise your hand.

The next phase of this study will be a computer-assisted learning lab. You will have up to 2 hours to study, but may terminate this session whenever you feel that you have attained competence with the study material and are ready to take the posttest. When you are ready to leave, please let a lab assistant know and return to the room where you took the pretest at the beginning of the study.

I will introduce you to Hyperaxis II, the computer program you will be using for this lab. To aid in this instruction, you will find a one-page instructional sheet at your station. You will also find a diskette containing the Hyperaxis II program. It will contain all that you need for this exercise. Please pick up your instruction sheet for reference.

When you are told to begin, put your diskette entitled Hyperaxis II into the computer at your station. Macintosh has been chosen for this learning lab, because of the ease of learning to use this program. If your computer program does not show the first screen immediately, double click, with your mouse, on the Hyperaxis icon. You will see an introductory page, which will look like this.

Proctor using either display monitor or flip chart, flip to introductory screen display.

Proctor: After reading the instructions, click on the finger located at the lower right hand corner of the screen (this icon will repeat throughout Hyperaxis II - it signifies forward). The next screen will be a questionnaire. It was originally intended for social workers, so a few changes are in order. Instead of telephone number, please fill in the last four digits of your social security

number. Also, "social work experience" can be considered to be how many months of counseling experience you've had, including internship. When you click on forward a map like this one appears. Clicking forward will bring up the main menu. From this screen information about each personality disorder can be reviewed by clicking on the icon with the name of the personality disorder. A clinical presentation may be followed with a differential diagnosis, a case example and other illustrations of the disorder simply by clicking on forward. Further information can be accessed by clicking directly on any icon which offers further information. The screen will explain how to get back to the previous screen. When finished with each, you can return to the main screen by pressing the arrow in the lower left corner of the screen. When you have reviewed all cases, you can click on the test case icon to take a practice test. At any time during the program you can return to any screen you wish to review by clicking the icon representing where you wish to go, including during the practice test. When you are finished, please do not attempt to exit the program but raise your hand and let a lab assistant know you are ready to go take the posttest.

You may begin.

Session #4 (posttest)

As individual participants enter the room, ask for social security code number and give the copy of the posttest with that code number on the front. Give the participant the following instructions

Proctor: Here is your posttest. Procedure for taking the test is exactly as the pretest except that, when you finish, please bring the test to the front desk and let me know you are finished. There will be a short interview at that time, and you will be finished. Please put away any notes or handouts you have before you begin the test. You may begin.

Proctors be sure to record time each test is handed in on the folder

Instructions for Use of Hyperaxis II

- >Place diskette into computer disc port.
- >Double click on Hyperaxis icon.
- >If you get a second set of icons, click on Hyperaxis icon.
- >After reading instructions, click on finger in lower right hand corner (forward)
(At any time that you wish to move to the next frame, click on finger).
- >Fill out "User Information," substituting Social Security # for telephone #
and months of internship and professional counseling experience for
"Social Work Experience."
- >The next frame is a map of the Hyperaxis II program. Notice that the next
window will be the main menu. To access information about any
personality disorder, click on the icon bearing the name of the disorder.
The next three windows following the main menu contain information
about the disorder be accessed by clicking on the "forward" icon.
- >After you have studied a disorder (which ends with "case example" on your
program map) you can return to the main menu by clicking on the
icon which contains an arrow (located in the lower left corner of the
screen). At any time that you wish to look at a previous window you may
do so by clicking the arrow icon. To return to Axis II categories during the
practice test, click the button to the far right of the diagnosis you wish to
review.
- >During study of "clinical presentation" arrows pointing up and down appear to
the right of the screen. These are used to scroll, so that you can read the
whole text.

- >Buttons will appear periodically offering added information about particular DSM categories. To access, click directly on the button containing the offer of information. The information frame will tell you how to return to previous frame.
- >If at any time you need help, click on the icon with a question mark (located the bottom, middle of the screen) and instructions will appear. The help frame will tell you how to return to the previous frame.
- >After you have studied the cases, you can move forward and take a sample test. When finished move forward to the last window and then exit. You can't totally shut down. When you are finished indicate to a lab assistant that you are finished.

Appendix D

Structured Posttest Interview

As subjects finish the post test, they will bring the test to one of the monitors. The monitor will accompany the subject to a quiet place and conduct a interview. All interviews will be recorded.

Monitors will work with a script throughout the study so as not to bias the experiment.

First, the monitor will choose a case presented in the post test. Please choose one which the participant has answered and that is over one page long. The monitor will hand the subject the page which contains the chosen case (cases are separated from answer sheets). Subjects will be asked:

“Please talk your way through this case as you remember taking the test.”

The question will be posed with as few cues as possible as to the desired information so as not to bias answers from subjects. The monitor will sit quietly and attend to the answer of the subject. If the subject has not answered the query in such a way that the monitor is sure that the question of concern - what cognitive process the subject used to make the diagnostic decision - has been answered, follow up questions will be asked as necessary until the question is answered.

“You said....What were you thinking when you made that decision?”

If the subject still has not supplied ample information, the monitor will inquire:

“How did you reach that decision?”

If, again the subject has not supplied the necessary information, the monitor will inquire:

“How did you choose the diagnosis(es) that you chose?”

Appendix E

Problem Solving Approach to Diagnosis

>Diagnostic researchers have discovered that proficient clinicians approach the diagnostic interview in particular ways. The way we will approach this process is based on a philosophy of decision making called either Problem Solving.

>Problem Solving is based in research of learning.

>In the diagnosis of human problems, accurate decision making by clinicians is a process.

>Diagnosticians learn the method of thought that is related to accurate client assessment very early in practice of diagnosis.

>The process, Problem Solving, is the result of study of physicians, experienced and inexperienced who were renowned as proficient diagnosticians.

THE METHOD

>Begin to hypothesize about the nature of the problem very early, while listening to the patient, within less than a minute of the beginning of the interview or case review.

>Many diagnostic hypotheses are generated throughout the whole interviewing process. Several hypotheses can be carried at 1/4 and 1/2 way through the interview.

LEARNING HOW TO RETAIN MULTIPLE HYPOTHETICAL DIAGNOSES

>We remember in "chunks" of 5 to 7.

>Learners can discover what is a chunk for them by memorizing quickly 5, then 6, then 7 and so on until, when they attempt to recall, they can only recall part of the grouped data.

THE DSM

>The disorders are divided into larger categories, Axis I and Axis II Disorders.

>Axis I disorders are divided into several categories, developmental, organic, substance use, psychotic, mood, anxiety, somatoform, dissociative, sexual, sleep, factitious, impulse, adjustment and those affecting physical function.

>Axis II disorders are divided twice, by cluster and disorder.

>From the beginning of an interview or case review, the diagnostician should begin to entertain as many hypothetical diagnoses as make sense, based on the symptomatic cues given by the client.

>Continue to think as broadly about what might be going on with the client as long as is practical.

>It is important to learn early in your career to think broadly and keep an open mind.

>Think of as many possible solutions as possible as early as possible and don't rule anything out until you are convinced that a particular hypothesis can no longer be considered in light of the information given.

>Learn to filter which client cues are most meaningful

>Learn to note which diagnoses occur most often.

>The same process that was related to accurate diagnosis has been shown to exist in client case management decisions. Accurate diagnosticians use the same thought process when considering the needs of their clients and so made decisions based on multiple hypotheses, both of what problems exist and what should be done about them.

>Development of the ability to think about multiple hypotheses and to allow that to guide reflections is an important skill to develop.

>In psychological counseling there may be more than one diagnosis.

>Elicit counseling flow from a client, developing an empathetic relationship and move a client toward reflectivity.

PRACTICING PROBLEM SOLVING SKILLS

>To practice, as you read a case study, consider the responses you would make to get the client to talk about the presented problems in such a way that you can gather more information and refine the several hypotheses at the same time.

>Practice interviewing the client as you read the case study, sentence by sentence.

>Students who begin to learn problem solving as early as possible reach proficiency early in their careers.

>Practice is important for developing the skills, learning to pose questions to self and elicit cues from clients, pay attention to appropriate detail, notice the frequency of occurrence of particular disorders.

>It is not what is taught, but what is practiced that determines what is learned.

The problems to avoid include:

- Being too general with hypotheses (psychosis rather than schizophrenia),
 - Discarding information that the client attempts to offer,
 - Ignoring potential new diagnoses (to avoid having to generate new hypotheses),
 - Assigning exaggerated importance to justify retention of existing hypotheses.

Appendix F

Problem-Solving Video Script

The Diagnostic and Statistical Manual of Mental Disorders, The DSM, has become the most widely accepted and used instrument for categorization of human abnormal behavior. However, although there is almost universal agreement that the manual is the most usable instrument for diagnosis and classification of psychopathology, clinicians regularly disagree on the diagnosis given to individuals who seek help from the mental health community. And because most treatment is based on diagnosis, a client could be treated for different disorders depending on the orientation of the clinician. So, it has become important for researchers to develop methods of diagnosis which will bring about more reliability between diagnosticians.

Diagnostic researchers have discovered that proficient clinicians approach the diagnostic interview in particular ways. Based on the practice of experienced clinicians, then, interview styles and ways of thinking about client signs and symptoms have been developed to make the process more consistent. The way we will approach this process is based on a philosophy of decision making called either Problem Solving or Patient Management Problem. We'll use the term Problem Solving, since it is the most commonly used reference to this thought process.

Problem Solving is not a system, but rather a theory of accurate diagnosis, based in research of learning. In the diagnosis of human problems, accurate decision making by clinicians is a process. Diagnosticians learn the method of thought that is related to accurate client assessment very early in practice of diagnosis. This is a whole method of learning as well as the practice of client assessment.

The process, Problem Solving, is the result of study of physicians, experienced and inexperienced who were renowned as proficient diagnosticians. The method of diagnostic decision making can be explained relatively easily.

The first important feature of Problem Solving is that While interviewing patients during an assessment , we begin to hypothesize about the nature of the problem very early, while listening to the patient. In other words, the first idea about what the diagnosis may be usually appears within less than a minute of the beginning of the interview.

-Next, many diagnostic hypotheses are generated throughout the whole interviewing process. In fact, several hypotheses can be carried throughout a good portion of the interview with a client. Even at 1/4 and 1/2 way through the interview, there may be several diagnostic hypotheses consistently being considered.

How many? Well, that depends on the individual human capacity to remember. In fact, we remember in what is called "chunks". The average human remembers in chunks of 5. Some people can remember as high as 7 ordered things without beginning to forget part of the data.

(camera pans to poster of numbers chunks)

If you think about numbers we commonly use, most number sets are set up in such a way that we seldom get over 5. Even phone numbers are set up in a pattern of 3 followed by 4, separated. Why separated? Because, it has also been discovered that we can remember more if we make subsets of items, dividing them up so that we can remember smaller quantities together. For some reason, then, we can remember much larger quantities at once, because they are divided into chunks.

(camera pans back to instructor)

Now, consider the Diagnostic and Statistical Manual of Mental Disorders. The disorders are divided into larger categories, Axis I and Axis II Disorders. Axis I disorders are divided into several categories, developmental, organic, substance use, psychotic, mood, anxiety, somatoform, dissociative, sexual, sleep, factitious, impulse, adjustment and those affecting physical function. Axis II disorders are divided twice, by cluster and disorder.

Most of what we attempt to remember is divided in such a way that it reflects the way the mind works. So, one of the things each of us has to discover is how many is a chunk for us. When we stretch this clustering, we

tend to have things “fall off” and so lose pieces what we try to learn, forget. Most of the physicians studied tended to entertain 5 hypotheses at one time. In other words, from the beginning of the interview and until quite late in the interview, it is possible to use problem solving thought process, think of about 5 possible potential diagnoses, based on the symptomatic cues given by the client and continue to think as broadly about what might be going on with the client as long as is practical.

One system is to look at several hypotheses at once, to find ways to subdivide what we want to look at, so we can group things that had aspects in common together. In other words, we can clump together general diagnoses (like mood disorders instead of thinking of all of the type of mood disorders, until they were sure they were really observing a mood disorder). That way they could continue to look at as much as they needed and discard nothing until they felt the need to.

The problems to avoid include:

- Being too general with hypotheses (psychosis rather than schizophrenia),
- Discarding information that the client attempts to offer,
- Ignoring potential new diagnoses (to avoid having to generate new hypotheses),
- Assigning exaggerated importance to justify retention of existing hypotheses.

So, it is important to learn early in your career to think broadly and keep an open mind, to think of as many possible solutions as possible as early as possible and not to rule anything out until you are convinced that a particular hypothesis can no longer be considered in light of the information given.

On the surface, this would seem to be rather inefficient. It would seem that considering so many possibilities would be confusing, take too much time, and potentially leave an undecided clinician. Interestingly, it was discovered that the opposite was true. Professionals who learned early to use problem solving techniques proved to be efficient and accurate. It is also important to learn to filter which client cues are most meaningful, and to begin note which

diagnoses occur most often. Nonverbal cues have also been shown to important cuing systems.

The same process that was related to accurate diagnosis has been shown to exist in client case management decisions. Accurate diagnosticians also tended to use the same thought process when considering the needs of their clients and so made decisions shown to be more effective in therapy. Development of the ability to think about multiple hypotheses and to allow that to guide reflections is an important skill to develop.

(camera pans to models of generating hypotheses)

Chronologically, this process looks something like this: (Point to chart)
(camera pans back to instructor)

Notice that in psychological counseling, the assumption is that there may be, and commonly is, more than one diagnosis.

In fact, the assumption in the problem solving approach to client management is that diagnostic classification is a multidimensional description of client dimensions of behavior, not a rigid classification and so should be to some degree flexible. The reason is that we are describing patterns of behavior so that we can conceptualize, plan and discuss counseling cases. Diagnosis, like counseling, is not a hard science, but a phenomenological interaction in which the counselor takes care not only in guidance but in planning. So the best description is probably the one that describes as many dimensions as are important to consider. Grob stated: "Classification systems are neither inherently self-evident nor given. On the contrary, they emerge from the crucible of human experience; change and variability, not immutability, are characteristic. Indeed, the ways in which data are organized at various times reflect specific historical circumstances. Empirical data, after all, can be presented and analyzed in endless varieties of ways."

This, of course, fits well with the counseling skills considered to be the most effective. The ability to elicit counseling flow from a client, develop an empathetic relationship and move a client toward reflectivity and motivation to change or to invest in a therapeutic regimen seems most congruent with the description of problem solving as direct associative retrieval. Associative

retrieval is basically the process of noticing the common features between otherwise very divergent psychopathologies and allowing that broad based and reflective recognition to guide ruling out and reconceptualizing until understanding is achieved.

It is also important for the student to find a way to “verbalize” the process. This may not always be easy to do, but the effect of talking one’s way through cases is to again practice using the methods of problem solving, thinking reflectively about the cues, developing hypotheses and generally practicing at becoming more aware. It also helps to develop skills of interview that would assist with information gathering.

For instance, as you read the case study, consider the responses you would make to get the client to talk about the presented problems in such a way that you can gather more information and refine the several hypotheses at the same time. Consider the counseling skills you have learned in practicum, internships or counseling skills classes. Take time, considering as you go how you would respond to the information if you were hearing it directly from the client, you may find that you begin to integrate the skills of interactive counseling and reflective response, but now with the addition of improved descriptions and more organized ways to conceptualize without jumping to conclusions. This might be a slower way to read, but the counselor will develop better interactive skills while developing better diagnostic skills.

But shouldn't a student first learn the concrete information from the DSM before attempting to concentrate on problem solving? Well, so far research has suggested strongly that students who begin to learn problem solving as early as possible reach proficiency early in their careers, applying the principles as they practice counseling techniques (early in sessions with clients) as well as in assessment courses.

It was also discovered that practice was an important component of developing the skills, learning to pose questions to self and elicit cues from clients, pay attention to appropriate detail, notice the frequency of occurrence of particular disorders. Whether students were watching films of diagnostic sessions, reading case studies, or actually performing client interviews, practice

appeared to be a most important component of learning, and using problem solving early was correlated with learning accuracy, just as practice resulted in better use of problem solving skills.

It is not what is taught, but what is practiced that determines what is learned, and methodically thinking (or talking) one's way through cases is the most effective way to learn.

So, let's now consider a case and illustrate before you go to the learning lab. Because of the complexity of diagnosis, which is fairly typical of real life situations, we will use a client presentation of symptoms of mood changes to illustrate this way of organizing the diagnostic process. The DSM-III-R Casebook provides us with such a case on page 90. Let's listen to the case, using problem solving techniques while reading to illustrate responses and develop hypotheses as we read. (Read Sickly, p. 90-91). (Read first sentence)

It would easy to assume that the client is suffering from a mood disorder and begin asking questions, but that may not be so. You may not be aware of this, but there are 19 separate diagnoses in the DSM that mention depressed mood as a criteria for diagnosis. 19 is too many for a brain chunk, but when we consider that 3 are organic, 4 are substance induced, one is psychotic, 6 are mood disorders, 2 are stress related, 3 are related to other Axis I diagnoses, and 1 is a sexual diagnosis, we now have usable chunks. We also will not assume all of them, anyway. But as your knowledge of the DSM develops, you should be aware that the symptom exists in many disorders. You could hypothesize right away that there may be an organic reason, that the client may have a mood disorder or that she may have a stress related disorder or that it may not have a reality base (may be psychotic). There are some important things to hear. But we can find out much of this by inviting the client to tell more about the depression, then to reflect the affective material in a client centered manner. This will probably provide significant information in a contextual fashion. (how might you have invited more information from the client?) (Read the rest of the sentence)

It appears that the client is oriented, as she is able to provide context within a month (time orientation) and recognizes change to the situation.

Though we cannot rule out completely psychosis, we can backshelve it. This certainly sounds like a mood disorder, but there is no indication whether it is longstanding or perhaps cyclical, affects her function, or is in response to something organic or environmental. So, as the psychotic possibilities begin to fade from the set of hypotheses, historical and stress related material might be considered alongside mood disorder. What else might you wish to consider?

We might, then request information from these areas. The questions can be general or we might simply wish to continue to show empathy, allowing the client to formulate her story. The client's agenda is to take us back into discomforts she has experienced for some time.

(Read 2 sentences, second paragraph)

We have discovered the possibility of organic causes or maybe the result of stressors from father's leaving. We haven't yet lost most of our hypotheses, but we have confirmed that the client is within a reality base, psychosis is doubtful. It also appears that we can discard mood disorders that are manic in nature, at least for now.

We can also establish that some organic hypotheses, like dementia probably don't exist, since her memory is intact both long and short range. It is odd that the physician did not discover an organic base for the childhood depression or for the sickly nature but prescribed a "tranquilizing agent" - the wine. We might at this point add some hypotheses of not only stress related etiology but also other potentially neurotic diagnoses. We now do know that the client has experienced depression as a reaction to circumstance at least once before and need to place a specific hypothesis or recurring mood alongside those we have.

How would you proceed then to evoke more information? There are several ways to proceed, but no clear indicators of exact nature of interview. You might reflect back to the client that the physician did not prescribe anything at the time, and ask her how long the symptoms persisted. The client then takes us past the time of father's desertion and fills in more of the pictures. Minimal Encouragers, clarifiers nonverbal listening cues might be in order as she speaks.

(Read paragraph 2 to “department store)

Well, we now know that the sickly symptoms spontaneously disappeared. We can begin to suspect that organic causes are improbable. We also see fairly rapid development, graduation early and a responsible position, young. We could assume this as precocious or we could wonder about the family structure. Is she driven, is she just intelligent. What does this have to do with the depressions she experiences now? She doesn't go on to college. What does she do and why?

Notice that two things are happening to our reflection and to the resultant interview. First, we have enough hypotheses that we can't just run down any single alley of reasoning, asking specific questions to prove or disprove a single diagnosis. Because we have so many potential diagnoses, and only the assessment session with which to work, we need to encourage the client to talk broadly. However, we also are being guided away from areas. Notice that we have discarded several broad areas of potential concern as we add others and refine those we already have. This fits well with what we have learned of counseling skills. In a safe and empathetic environment, the client will provide information. Yet, we see the advantage of this structure to guide the interview by dictating what we reflect and don't reflect back to the client. We don't need to check cognition but do need to get pictures of potentially continuing episodes of depression and need a better systemic context. So, how would you encourage more information? You could note that it sounds like the depression at least temporarily subsided and that she functioned well but shyly in high school and let her respond.

(Read last sentence, second paragraph)

We get a systemic picture of discord but also add a new hypothesis of sexually related material. It would be easy to sidetrack on one thing here, except that we have the advantage of other hypotheses to keep our agendas from interfering with listening. Adding sexual diagnoses and family discord begins to refine our set of hypotheses. How would you reflectively respond not to get too narrow but to find out more of the client's picture of this period of time? You could simply reflect back that she was having trouble in her marriage as

had her mother and ask how she coped. Then, You could quietly encourage the therapeutic flow.

(Read paragraph 3)

So, we now have added another hypothesis, and since the client has resolved her own defense mechanisms so that we get a clear picture of developed alcoholism, including symptoms and treatment. So, we can not only add the hypothesis but pretty well confirm it. She coped with a dysfunctional family setting and with sexual indifference and pain by drinking. It does make sense, then, that Dad abandoned the family. It would fit that he was a drinker and that could be checked out later. But for now, we can refine our hypothesis of mood disorder further. The client has had recurring depressive episodes and has been treated for them even after cessation of drinking. In fact, they were severe enough that she received major forms of therapy. She is still in reality with her story, so psychosis does not reenter the picture, but the depression appears to be deep, if therapies don't work. We do know however that the client has informed us, she has been depressed for a month, so recurring rather than chronic enters the hypothesis refinement. It also brings up a natural question, "If therapy brought no relief, what did and what did you experience when the depression lifted this time?" Although depression is severe enough to retain and begin to refine as to which kind, the natural cessation suggests some neurotic affect. I've responded to the client, with the request for information about her experience of natural cessation of the depression. What would you have reflected. What hypotheses would you have discarded, added or retained? Again, I will encourage and reflect as necessary to keep flow. The client's response is...

(Read paragraph 4)

It is interesting that she readily admits to ebb and flow of depression and replacement with somatic (physical) discomforts, pains and nervousness. These are not light weight symptoms the client is experiencing. She has physical problems all over her body, but has not had confirmation from the physician, except a hysterectomy (which has not relieved symptoms) and an abscess of the throat. Since there are 2 major diagnostic clusters which

somatize, we need only consider severity of symptoms. Histrionic personality disordered individuals will externalize and live on drama, but tend to have vague symptoms, while somatoform disorders tend to be more specific and definitely more physically focused (pain and medically focused without a lot of manipulation of others). We know that we have a rather large cluster of somatic complaints, changing over time, but that most recently include some very specific symptoms we can use as we read the DSM after the interview and reflect on what we've heard. Vomiting, food intolerance, weakness, fatigue, chest pain, fainting, anxiety attack all are occurring concurrently.

So, as we refine our hypotheses further, it is pretty obvious that we have an individual that gets severely depressed at times, but who is quite neurotic with a fixation of physical health. We also are aware that she is alcoholic and that the alcoholism is in remission. Remission has not allowed symptoms to subside, so other diagnoses are in order. We have now only to gather adequate information to be able to later reflect as we sit alone in the office or consult with another professional and "talk" or think our way through what we have heard. You may wish to know other things about her present circumstances outside of the depression. You have an adequate history, but need to know more about family and any other environmental circumstance she may feel is important. You also need to confirm the hypothesis of both family dysfunction with her reaction being neurotic rather than dealing with things in a conscious fashion, particularly the circumstances of her father's abandonment of the family.

(Read last paragraphs)

We have confirmed the alcoholism in the family, which reinforces our diagnosis of her own. We also know that she felt the need to overcome, but ended up in another relationship much like the one with her father. She was overcome by symptoms so that she became more dependent. However, she has not enjoyed her relationships and has had difficulties enjoying sex. There has been little question for some time during this interview that the client has substantial depressive symptoms that have recurred periodically, and been severe enough to be treated in an inpatient setting. This we can not overlook.

The client continues to seek a physical explanation of everything including her depression and wants magic bullets, rather than dealing with her psyche. We now have enough information to reflect, so can terminate the interview, set a practical time for the next session (soon, to deal with depression) and take time to look at what is and is not present in the client's presentation. Note that we have not confirmed diagnoses in session but need time to reflect and consider the competing hypotheses I still retain. You can then read through several diagnostic criteria on may to ruling out some and confirmation of others.

My first diagnosis will be that of alcohol dependence in remission. There is ample evidence to easily confirm, since the client came from an alcoholic gene pool, lost control and needed treatment. Her continued abstinence also aids in the diagnosis.

In looking back at the information received about the depressive symptoms you can see that the client shows four symptoms of major depression (one month duration, diminished interest, insomnia, and poor concentration). Normally this would be considered to be Dysthymia symptoms - minor depression, but there is ample evidence of more severe depressions requiring hospitalization in the past, and so we can diagnose the client with major depression, recurrent. Dysthymia tends to recur consistently with minor symptoms and this is not the case with this client. However, we have no evidence of suicidal tendencies, psychosis or other signs that suggest this major depression is one of the more severe, so we can add the descriptor, mild, and state the client is suffering from a Major Depression, recurrent, mild.

Now, to consider the somatic complaints I will reread the personality disorder criteria and rule that out because the client is not shallow (she feels deeply, argues with her husband) and she has much detail to complaints. She appears not to take center stage. So we begin to look at somatoform disorders and discover that she has multiple complaints about many organ systems. The condition has gone on for many years and has not proven to be organic. The client is not concerned about physical appearance, is not incapacitated like a conversion disorder, preoccupied with a major disease as a hypochondriac would be, nor body dismorphic. She fits the characteristics of Somatization

disorder better than any other. However, she only has 11 rather than the 13 required symptoms for the disorder (sexual indifference, pain during intercourse, chest pain, abdominal pain, backaches, extremity pain, vomiting, unconsciousness, food intolerance, weakness, and colitis or diarrhea). So, I would give the diagnosis but as a trait rather than a full blown syndrome. It would read, Somatization disorder (provisional).

We now have a diagnosis that describes the client as we know her so far. we can do one more thing with this case diagnostically. It is uncomfortable ruling out entirely the personality disorders, since we have not known the client long enough to be sure of the Somatization diagnosis. In fact with a provisional diagnosis, we would be wise to retain a hypothesis of at least a potential personality disorder. After all, the client does have a long term and pervasive view of her own body function and it does appear to be a neurotic defense against dealing consciously with her social and psychological problems (my own conceptualization), so we can defer diagnosis on Axis II, personality traits and continue to investigate as sessions continue.

Ok, I have illustrated the use of problem solving, multiple competing hypothesis and how they guide not only the counseling session but the reflective process of diagnosis. You will in a moment move to the computer learning lab and spend a time getting acquainted with the personality disorders from the DSM-III-R, the Axis II diagnostic categories. As you study, practice the process of competing, multiple hypotheses. Also, consider the kind of counseling responses you would use to gather information from your client.

During the computer-assisted lab you will be presented with Axis II and an opportunity to pay particular attention to the common elements of differing diagnosis can be practiced. Next, during a review of case studies, you will have the opportunity to notice not only features that coexist between the personality disorders, but also between the Axis II diagnoses you will study and other disorders with which you are already familiar. The common verbal cues that clients tend to use, may help develop recognition of cues later. Finally, a practice test is included in the program. This is a most important part of development of problem solving skills, because, not only can you practice your

own competing hypotheses, including multiple diagnosis in appropriate cases, but you will get immediate feedback of the outcome of your diagnosis (the only restriction to this learning lab, is that the computer only recognizes the one most predominate diagnosis, even if other diagnoses or features of other diagnoses are present). Nevertheless, if you diligently take advantage of the feedback, reconsider cues from cases you fail to diagnose correctly and even return to earlier sections to review when you feel you need it, you are practicing both reflectivity and associative recall.

Decision Tree Video Script

The Diagnostic and Statistical Manual of Mental Disorders, The DSM, has become the most widely accepted and used instrument for categorization of human abnormal behavior. However, although there is almost universal agreement that the manual is the most usable instrument for diagnosis and classification of psychopathology, clinicians regularly disagree on the diagnosis given to individuals who seek help from the mental health community. And because most treatment is based on diagnosis, a client could be treated for different disorders depending on the orientation of the clinician. So, it has become important for researchers to develop methods of diagnosis which will bring about more reliability between diagnosticians.

Diagnostic researchers have discovered that experienced clinicians approach the diagnostic interview in particular ways. Based on the practice of experienced clinicians, then, interview styles and ways of thinking about client signs and symptoms have been developed to make the process more consistent. The way we will approach this process is based on the decision tree. The decision tree is a series of signs or symptoms for which the clinician checks in the diagnostic interview. Questions are formulated for the purpose of discovery of whether or not enough symptomology exists to give the client a particular diagnosis.

We are not, in this study going to actually practice diagnostic interviewing. But the process of interview very much parallels the process of case conceptualization. In other words, most counselors think about the client in much the same way that they seek information, and in fact approach the interview with methods congruent with their own diagnostic decision making. This makes sense, because it allows the counselor to stay organized and efficient to the task of ruling out improbable diagnoses and moving systematically toward a decision to explain what the counselor observes. So, decision tree will be presented in an interview format so you can see both how it is used in the assessment interview and how the clinician uses it to make diagnostic decisions.

The decision tree method is a highly organized process, which focuses on problems (symptoms and signs) that the client presents during the initial parts of the assessment interview and singles each significant symptom out for investigation as to whether the suspected syndrome that is hypothesized by the counselor (from the cue of the stated sign or symptom) is in fact the key to diagnosis. Once a major symptom is singled out, the counselor uses the DSM decision tree as a systematic focus both to systematically narrow down diagnostic possibilities by the questions asked and to stay organized building a set of symptoms that will ultimately result in a diagnosis.

Of course different settings dictate the length of time available to the clinician to perform a diagnostic interview. In an inpatient or hospital setting, a counselor may have 2 or 3 hours for an assessment interview. In an outpatient setting, whether in a community, mental health, chemical dependence clinic or in a school setting the standard time allotted for a session of any sort is 45-55 minutes. Usually, by the end of the first session, the counselor is expected to have ample information to develop at least a preliminary, differential diagnosis.

Rather lengthy instruments have been developed for hospital settings, particularly the Structured Clinical Interview for the DSM, developed by the New York Psychiatric Institute with grant money from the National Institute for Mental Health. This instrument, however, takes 2-3 hours to administer in its entirety. As this is somewhat impractical in many settings, the authors have suggested that particular modules - parts of the test designed to investigate a suspected diagnosis be used whenever time restrictions dictate a shorter version.

Diagnostic literature agrees on one point, that any diagnostic interview should begin with an open-ended interview which investigates with the client past history, particularly the presenting problem in the client's own words, prior treatment, social and occupational functioning, and the context of the development of the presenting problem or psychopathology. At the point that a counselor begins to suspect that a particular pattern of symptoms and signs are being presented that suggest a particular group of diagnoses, a line of questioning is developed, using the decision tree model, to direct the interview. It is helpful, at first, to ask general questions until symptomology directs the

counselor toward a diagnostic area. At that point the decision tree, then, becomes the instrument for use, whether the counselor asks questions from a structured interview format or from one's own design, based on the decision tree.

Some points of importance should be made about the decision tree based diagnostic interview. First, the diagnostic interview must not only start open-ended, but must use the structured approach as a guide to interview, not a firm protocol. In other words, since many questions in an interview of this sort are specific, they may elicit a YES or NO response from the client, which may not be adequate to explain the circumstances and context of the answer. It is important to ask the client to elaborate.

(camera should pan to complete decision tree chart for mood disorders)

Because the decision tree is highly organized (consider the decision tree for mood disorders) it is quite easy to organize both the interview and particularly the case conceptualization for decision making. This is obviously important since there are so many possible diagnoses cued by that original presented symptom or group of symptoms.

(camera pans back to instructor)

The sequence of questions should be designed to approximate the differential diagnosis process. As Reid & Wise said in their textbook teaching decision tree diagnostic process, "Parsimony (attempting to fit all symptoms into one disorder - simplicity) is a good tool." Obviously, in the initial stage of an interview organization is impractical. The client must start globally to express him or herself. But as cues begin to be identified by the counselor, the decision tree organization can begin to be imposed on the interview.

Because diagnostic categories have traditionally selected a minimum amount of criteria of symptoms and signs required to actually make a diagnosis, it is a relatively simple process to learn not only the categories and criteria for recognizing syndromes, named diagnoses, but also to organize the thought and interview process to confirm or to rule out specific syndromes.

Of course, whether a person is using case studies as we will be doing in the learning lab, interviewing a client during intake or reading other clinicians'

notes, the counselor should use all of the information available. The counselor should also organize the assessment session, whether it is an interview or a case study to fit within the time frame allotted, which means that organization should be practiced from the very beginning of learning diagnosis.

The counselor should, if he or she is interviewing, use clarifying questions whenever it seems important to gather further detail about a client's response to decision tree questions. So, as you learn the Axis I categories today in the learning lab, you should begin to formulate the kinds of questions you might ask the client to find out if the client has experienced the signs and symptoms you are learning. In other words, you should begin to ask yourself the questions you will ask both the client and yourself later to organize your decision making. In fact, as you read case studies, you should also look for the signs and symptoms in the same organized fashion. You will find that as you practice, the organizing principles of decision tree and its simplistic YES or NO criteria will become easier and to most quite natural.

Other principles of decision tree that are important to discuss are the concepts of "absent, subthreshold or threshold." If a symptom doesn't exist the client does not meet that criterion for diagnosis. If the client has some but not all of the DSM required criterion for full diagnosis, the clinician should consider the the diagnosis as subthreshold and report the diagnosis as "traits" rather than full diagnosis. For instance, to declare a person as diagnosed with a Major Depressive Episode, the individual must meet 5 criteria from the list of 9 (depressed mood, diminished interest in activities, weight loss or gain when not dieting, insomnia or hypersomnia, psychomotor agitation or retardation, fatigue, feelings of worthlessness or guilt, diminished concentration, or recurrent thoughts of death). 5 qualify for threshold or full diagnosis and 4 qualifies for traits or "provisional" diagnosis.

Clients tend to come in the counselor's door with a single most pressing problem, but this does not mean that they present with only one diagnosis. Many do, but just as often clients will have more than one diagnosis or have a full diagnosis and one or more provisional or traits diagnosis. So, once the counselor has decided on the major diagnosis, the responsible clinician will

consider if other symptoms exist that are not explained by the primary diagnosis and if so choose another decision tree track for exploration for further diagnostic consideration. This may, in clinical practice, require that during the second interview, new questions be posed, based in the formulation and writeup of the first assessment interview. With case studies, like you'll be presented with today, this is even easier. Once symptoms are organized the question can be asked and new diagnoses formulated. However, be careful not to over diagnose. Once symptoms are explained, new diagnoses will only confuse and complicate the case.

Since we are not practicing interviewing during this learning laboratory, a few other things should be discussed about decision tree diagnostic interviewing, so that you have a full picture of how it would be done in practice.

1. Clients should receive a brief explanation of the purpose of the interview.
2. The counselor should not apologize for use of structured interview techniques. The practice has been done regularly in psychiatric communities and clients may appreciate the thoroughness.
3. The counselor should be sure to start with an overview including presentation of problem and history from which cues will be taken to formulate decision tree questions.
4. The counselor should not ask for details of the overview until ready to do the structured interview. Once the client has presented adequate information, the counselor should begin to use the decision tree questions.
5. Enough information should be present from the overview to provide a context for later answers.
6. Specific questions about symptoms should not be asked until the counselor is ready to use the decision tree.
7. The overview should be organized around problem presentation and specific questions about history, such as family structure, development, education, etc.
8. Wandering through information wastes time and leaves little for the structured interview.
9. Once into the structured decision tree interview, clarifying questions should be asked as appropriate.

10. The structured interview shouldn't be so simple that it becomes purely a true/false test.
11. The counselor should use judgment and should confront the client (gently) about incongruent information.
12. The counselor need not accept contradictory information.
13. The counselor should make sure that clients understand questions.
14. The counselor should use language the client understands.
15. The counselor should understand the time frame of client reports.
16. Clients tend to clump symptoms, not necessarily in chronological order, so the counselor should clarify when necessary when the client experienced certain things.
17. The counselor should be sure to ask all the questions pertinent to checking the criteria for suspected disorders.
18. The counselor should not over interpret, particularly with psychotic symptoms.
19. The counselor should be sensitive to subcultures.
20. Only significant symptoms should be noted.
21. Excluded symptoms should not be documented.
22. Always use criteria to diagnose.
23. Counselors should not make diagnoses that do not fit DSM categories or criteria.
24. Sequential decision tree interview and thinking is important to establishing diagnosis.
25. Learn not to skip criteria from the decision tree both in the interview and formulation stages of diagnosis.
26. Remember, although we are only dealing with the first and particularly the second Axis today, there are 3 more Axes important to conceptualizing and describing the client.

Let's look at a particular diagnostic track, using the decision tree found in the back of the DSM (because the learning packages we will use today were developed for the DSM-III-R, and software has not been developed to teach DSM-IV, all materials used will be from the DSM-III-R for consistency). Because

of the complexity of diagnosis, which is fairly typical of real life situations, we will use a client presentation of symptoms of mood changes to illustrate this way of organizing the diagnostic process. The DSM-III-R Casebook provides us with such a case on page 90. First let's listen to the case, then, using decision tree process questions, we'll diagnose the client. (Read Sickly, p. 90-91).

The first necessity is to separate the presenting problem from more longstanding problems. The client complains of depressive symptoms. Using questions developed in the SCID, we can ask the client, "In the last month has there been a period of time when you were feeling depressed or down most of the day nearly every day? (What was it like?)

YES

How long did it last? (As long as two weeks?)

YES

What about losing interest of pleasure in things you usually enjoyed?

YES

Did you lose or gain any weight?

YES

(What about compared to your usual appetite?) (Did you have to force yourself to eat:)

YES

How were you sleeping?

(Trouble falling asleep, waking frequently, trouble staying asleep, waking too early, or sleeping too much? How many hours a night compared to usual?)

Was that nearly every night?

YES

Were you so fidgety or restless that you were unable to sit still?

NO

What was your energy like?

WORSE THAN USUAL

How did you feel about yourself? (Worthless?)

NO

Did you have trouble thinking or concentrating: (what kinds of things did it

interfere with?) Nearly every day?

YES, recently.

And so on.

The client has also reported to us that she has been sickly since childhood and that she drank heavily for 8 years of her early adulthood. We will, using decision tree take each of these reports separately and use decision tree questions to investigate them.

After investigating depressive criteria, the next to be investigated are symptoms the sickly ones. These questions will follow a logical order through the decision tree so that we may rule out as we go. First, we will ask some questions to establish history. For instance:

Over the last several years, what has your physical health been like?

NOT TOO GOOD

How often have you had to go to a doctor because you weren't feeling well?

(What for?)

YES, A LOT, pain, nervousness.

Was your doctor always able to find out what was wrong, or were there times when the doctor said there was nothing wrong but you were still convinced that something was wrong? Do you worry much about your physical health? Does your doctor think you worry too much? Some people are very bothered by the way they look. Is this a problem for you?

YES

Tell me about it.

During the discussion of specific physical symptoms the clinician may ask questions like

Did you tell a doctor about (symptom)

What was the diagnosis? (What did the doctor say was causing it?)

Was anything abnormal found on tests or x-rays?

Were you taking any medications, drugs, or alcohol around the time of the more recent symptoms?

How much trouble have (physical symptoms) caused?

How much have (physical symptoms) interfered with your life? (has it made it hard for you to do your work or be with friends?)

The client's response was, according to the case report, that she did seek consultation but was diagnosed with nervous disorders such as "spastic colon" and has experienced symptoms most recently including vomiting, but with no identified etiology pronounced by the physician.

Turning to the reported alcohol use, we may now establish a line of questioning that will allow diagnosis.

While you were drinking did you ever miss work or school because you were intoxicated or very hung over?

Did you ever drink in a situation in which it might have been dangerous to drink at all?

Did you often find that when you started drinking you ended up drinking much more than you were planning to?

Did you try to cut down or stop drinking alcohol?

Did you spend a lot of time drinking or hung over?

(We do not have enough case material to establish what exact information was retrieved by the counselor, so we can only surmise that the counselor asked adequate questions to establish the full extent of the potential problem, since a diagnosis was established)

Now, let's follow that line of questioning through the decision tree and watch how it not only guided the interview but organized the clinician's thought toward decision making.

(move to chart 1)

The client originally reported feeling depressed. So we can establish our immediate line of questioning around the mood disturbances

(point to first box, chart 1)

Because the client had established for us throughout the interview that no organic basis for her pains had been discovered we can answer the first decision tree question. Thus, we do not need to deal with the Organic Decision Tree.

So, we are lead to the questions establishing whether there have been one or more periods of persistently elevated, expansive or irritable mood, and associated symptoms. When the client has established that there have not been, we have eliminated a full half of the decision tree.

(Illustrate by showing full chart of decision tree and "wiping out" the right half.)
(move to p. 2 of mood disturbance chart)

So, we can concentrate on the line of questioning dealing with depressive disorders only. We can ask questions about the length of the depressive syndrome.

(move to char 3 of mood disturbance)

The client established for us that the depression has lasted for about a month, so we can find out the severity, by asking questions regarding potential psychosis, to which the client responds negatively. We now know that this is a mood disorder, rather than a schizoaffective disorder. So, again, we can cancel a whole line of questioning and concentrate on the mood disorder.

But we don't yet know if the client is psychotic (she certainly presents somewhat bizarre physical symptoms, are there accompanying delusional or hallucinatory symptoms?). Again we find none and so

(move to chart 4 of mood disturbance)

can establish that we are observing most of the features of a Major Depressive Disorder. We only have 4 of the 5 criteria required for diagnosis, we can say that she has experienced this before, and so may easily be seeing a recurrence of a Major Depressive Episode (one month of depressed mood, diminished interest, insomnia, and poor concentration). The case does not supply us with all the information we may need but we can establish that the client has been hospitalized in the past with major depression, so we are safer with that diagnosis than with Dysthymia.

(camera back to instructor, who moves toward charts - placed on other side - which do somatoform disorders)

So, we have established one diagnosis. The second line of questioning can be followed through the decision tree, watching the ruling out of lines of questions and establishment of diagnosis.

(camera moves to somatoform disorders page 1)

We have established a picture of physical complaints and anxiety about illness. We ask questions to establish potential organic findings and discover that the answers are for the most part negative. Although the client has had a hysterectomy, other complaints, from nervous stomach to nervous heart have not appeared in tests to have an organic basis.

Since the answer is NO, we can ask questions to establish if there is some conscious secondary gain from the somatic complaints. There appear to be none. So, we can ask the next questions, to establish if the client has been preoccupied with the belief that she has a serious disease for at least the last six months.

The client appears to have pretty general complaints, none of them of such a nature that they would qualify as a "serious disease" and most of her complaints being to some degree transient. For instance the vomiting has been occurring for about 3 months, so we can answer this question with a NO

(move camera to 6, which also page 2 of somatoform disorders)

If the client had been of the belief that the disorders were major, we would look at the potential of delusion, but they are not, so we abandon that line of reasoning and consider whether the client is preoccupied with pain, without other physical complaints. If she had, we would have arrived at a diagnosis of Hypochondriasis, but the client has other physical complaints, and in fact is more concerned with nervous conditions than with actual pain, so we can ask if the client is concerned about her physical appearance.

She does not present with symptoms nor signs of concern about physical appearance. Had she been concerned with appearance we would diagnose with somatoform pain disorder, but this is not the case, so

(move camera to 7, which is also page 3 of somatoform disorders)

we need to establish whether the client has had multiple physical complaints for the last 6 months. If she had not, we could then establish whether these signs were a response to psychosocial stressors, but she has, so we need to establish only whether she has had the required 13 symptoms before age 30, which she has. Just to assure ourselves that we have the required 13

symptoms, let's count: sexual indifference, pain during intercourse, chest pain, abdominal pain, back aches, extremity pain, vomiting, periods of "unconsciousness," food intolerance, weakness, and diarrhea (shown as colitis). We only have 11, so we make the diagnosis provisional.
(camera back to instructor)

We could follow the same reasoning process with the alcohol abuse, but the authors of this case did not provide us with adequate material to follow their entire line of reasoning. Besides the complication of the case could bog us down in detail. Be aware that the authors did diagnose the client as Alcohol Dependent, in Remission.

Due to the rather bizarre nature of the client's presentation, the authors also deferred diagnosis on personality disorder, in case the Somatization disorder, which you will remember was provisional proved later to be features of a Histrionic Personality Disorder. The diagnosis Histrionic Personality Disorder is not given even as provisional because somatization disorder describes the client better and so takes the place of the Histrionic disorder. (clinicians don't give two competing diagnoses in the diagnostic phase). Histrionic individuals show signs of somatic complaints, but they tend to be general in nature.

However, remember that we have a subthreshold diagnosis of somatization, so we will defer the Axis II until we confirm somatization.

OK, we've discussed the basic philosophy, interviewing protocols, systematic decision making process and criterion based method for decision tree diagnostic process. In a few minutes you will be introduced to a computer-assisted learning laboratory which is both easy and complete in presentation of DSM-III-R Axis II (personality disorder) categories and criteria.

Remember that the learning laboratory is not just for the purpose of memorizing the 12 categories, but becoming proficient at diagnosis, by practicing the principles of learning to conceptualize each diagnostic category uniquely in the ordered appearance of the criteria and developing a decision tree of your own both to think about and to formulate interview questions you would ask a client in a real assessment session. This association should also help you learn the categories more easily and practice skills needed later. If

someday you wish to access and use a structured interview, they can be ordered through the American Psychiatric Association or from the New York Psychiatric Institute and are free.

The computer learning lab will present you with the personality disorder diagnostic categories. As you read these categories, consider what questions you would ask to discover if the client displays or experiences each sign or symptom. The learning lab will also present case studies. As you read, rehearse the questions you developed. See if they fit with the signs and symptoms the client displays or experiences in context of the case study. Perhaps you'll develop more questions or refine the ones you've developed.

Finally, the computer learning lab provides you a practice test to gauge your skills. You should be able to diagnose most if not all of the disorders more easily if you are methodical, using the rule out techniques you've learned. Start with personality clusters and attempt to confirm the disorder you suspect. If you can't move to the next. If you make a mistake, you will get immediate feedback from the computer. What did you fail to consider? What decision tree step might you have missed?

Skills practiced early appear to affect proficiency later, so use this session to practice development of a decision tree mentality. Please wait for further instructions from the guide assigned to your group.

Appendix G

TABLE 34. Pretest Individual Item Mean Scores Using Three Point Scale

Test Item	Participant <i>n</i>	Mean Score	Standard Deviation
A	21	.5714	.7464
B	40	.8750	.9920
C	0	.	.
D	19	.1053	.3153
E	21	.4286	.9258
F	39	.3590	.7776
G	41	.3171	.4711
H	20	.6500	.9333
I	60	.1333	.3428
J	0	.	.
K	39	.1795	.4514
L	19	.1053	.3153
M	41	.1463	.4220
N	41	.2439	.4889
O	39	.1282	.5221
P	19	.0000	.0000
Q	41	1.6098	1.5146
R	39	.1026	.3074

TABLE 35. Posttest Individual Item Mean Scores Using Three Point Scale

Test Item	Participant <i>n</i>	Mean Score	Standard Deviation
A	39	1.5641	.8206
B	20	1.2000	1.0052
C	60	1.3833	1.1061
D	41	1.0769	1.2223
E	39	.4286	.9258
F	21	.5238	.9284
G	19	2.1579	1.1187
H	40	2.1750	1.1959
I	0	.	.
J	60	.8000	.9531
K	21	1.4762	1.3274
L	41	.5854	.7738
M	19	.7895	.9177
N	19	.7895	.9177
O	21	1.5714	1.3256
P	41	.1707	.3809
Q	19	2.0526	1.4327
R	21	.0952	.3008

TABLE 36. Pretest and Posttest Interitem Correlations Using Three Point Scale

Pretest Items Above Diagonal, Posttest Items Below Diagonal																			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	
r=		.16	.00	.00	-.01	.00	.14	.00	-.03	.00	.00	.00	-.18	.31	.00	.00	-.19	-.07	A
p=		.2			.3		.27		.45				.22	.09			.2	.39	
n=		21	0	0	21	0	21	0	21	0	0	0	21	21	0	0	21	21	
A			0	-.15	.48	.53	.35	.0	.06	.0	.69	-.15	.04	.01	.0	.0	.05	.04	B
				.26	.01	.01	.06		.36		.0	.26	.42	.48			.42	.42	
			0	19	21	19	21	0	40	0	19	19	21	21	0	0	21	39	
B				.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	.0	C
	-.22																		
	.18			0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
20																			
C					.0	-.17	.0	.0	-.21	.0	-.16	-.12	.0	.0	.0	.0	.0	.0	D
	.31	.26				.24			.2		.25	.32							
	.03	.14			0	19	0	0	19	0	19	19	0	0	0	0	0	0	
39	20																		
D						.0	.56	.0	.03	.0	.0	.0	-.11	-.15	.0	.0	.26	.43	E
	.28	.17	.15				.00		.46				.32	.26			.12	.03	
	.12	.24	.17			0	21	0	21	0	0	0	21	21	0	0	21	20	
20	20	41																	
E							-.31	-.23	.17	.0	.19	-.17	.13	.0	.27	.0	-.05	.0	F
	19	.04	-.05	-.01			.1	.16	.15		.13	.24	.3	.5	.05		.42		
	.12	.43	.39	.48			20	20	39	0	39	19	20	20	39	19	20	19	
39	20	39	20																
F								.01	.01	.0	.51	.0	-.11	.09	-.23	.0	-.10	.42	G
	.0	.0	-.1	.23	.0			.48	.48		.01		.24	.29	.16		.26	.03	
			.34	.16				20	41	0	20	0	41	41	20	0	41	20	
0	0	21	21	0															
G									.09	.0	.13	.0	.49	-.22	-.18	.0	-.32	.0	H
	62	.0	.59	.0	.45	.0			.36		.3		.02	.18	.23		.09		
	.00		.00		.03				20	0	20	0	41	41	20	0	41	20	
19	0	19	0	19	0														
H										.0	.31	-.21	.13	.05	-.11	.0	-.12	-.16	I
	59	.0	.42	-.01	.39	-.04	.91				.29	.2	.22	.37	.26		.24	.17	
	.00		.00	.5	.05	.43	.0			0	39	19	41	41	39	19	41	39	
19	0	40	21	19	21	19													
I											.0	.0	.0	.0	.0	.0	.0	.0	J
	.0	.0	.0	.0	.0	.0	.0	.0											
											0	0	0	0	0	0	0	0	
0	0	0	0	0	0	0	0												
J												-.17	.16	.25	-.1	.0	.03	.0	K
	.16	-.08	.41	.17	-.08	-.39	.08	.12	.0			.25	.26	.14	.27		.44		
	.16	.38	.00	.15	.32	.04	.38	.23				19	20	20	39	19	20	19	
39	20	60	41	39	21	19	40	0											
K													.0	.0	.0	.0	.0	.0	L
	.0	.0	.07	.19	.0	.11	.0	.09	.0	.1									
			.38	.21		.31		.36		.33			0	0	19	19	0	19	
0	0	21	21	0	21	0	21	0	21										
L2														-.06	-.17	.0	-.26	-.12	M
	.17	.45	.28	-.03	-.22	.19	.0	.06	.0	.2	.58			.36	.24				
	.24	.02	.04	.44	.18	.21		.39		.11	.0			41	20	0	41	20	
20	20	41	41	20	21	0	21	0	41	21									

Appendix H

TABLE 37. Pretest Individual Item Mean Scores Using One Point Scale

Test Item	Participant <i>n</i>	Mean Score	Standard Deviation
A	21	.4286	.5071
B	40	.5000	.5064
C	0	.	.
D	19	.1053	.3153
E	21	.2381	.4364
F	39	.2308	.4268
G	41	.3171	.4711
H	20	.4500	.5104
I	60	.1333	.3428
J	0	.	.
K	39	.1538	.3655
L	19	.1053	.3153
M	41	.1220	.3313
N	41	.2195	.4191
O	39	.0769	.2700
P	19	.0000	.0000
Q	41	.5366	.5049
R	39	.1026	.3074

TABLE 38. Pretest Individual Item Mean Scores Using One Point Scale

Test Item	Participant <i>n</i>	Mean Score	Standard Deviation
A	39	.8462	.3655
B	20	.6500	.4894
C	60	.7000	.4621
D	41	.4390	.5024
F	21	.5385	.5050
E	39	.3333	.4830
G	19	.8947	.3153
H	40	.8500	.3616
I	0	.	.
J	60	.4667	.5031
K	21	.6667	.4830
L	41	.4390	.5024
M	19	.5263	.5130
N	19	.4737	.5130
O	21	.7143	.4629
P	41	.1707	.3809
Q	19	.6842	.4776
R	21	.0952	.3008

TABLE 39. Pretest and Posttest Interitem Correlations Using One Point Scale

[illegible]

APPENDIX I

TABLE 40. Correlations of Scores to Explanatory Variables Using the 3 Point Scale
(Posttest Scores Above Diagonal, Difference Between Pretest and Posttest Below)

Line 1=r, Line 2=p, Line 3=n

Sc=Scores (Pretest Above Diagonal, Difference Between Pre and Posttest Below)

MP=Number of Multiple Hypotheses Written on Pretest

MO=Number of Multiple Hypotheses Written on Posttest

CP=Number of Multiple Hypothesis Cases Written on Pretest

CO=Number of Multiple Hypothesis Cases Written on Posttest

Ag=Age

GP=Undergraduate Grade Point Average

CX=Computer Experience

Ds=Classroom Exposure to DSM diagnosis

%C=Percent of Cases Reviewed on CAI software

%T=Percent of PracticeTest Items Reviewed on CAI software

%I=Percent Correct PracticeTest Items on CAI software

Ti=Time Spent on CAI software

I1=Number of Problem-Solving Participants Identified from Interviews by Rater 1

I2=Number of Problem-Solving Participants Identified from Interviews by Rater 2

Ge=Gender

	SC	MP	MO	CP	CO	Ag	EA	GP	CX	Ds	%C	%T	%I	Ti	I1	I2	Ge
SC		.43 .03 21	.07 .38 21	.34 .07 21	.04 .44 21	.64 .00 21	.78 .00 21	.27 .12 21	.40 .04 21	.51 .01 21	.13 .30 20	-.3 .09 21	-.1 .30 18	.14 .26 21	.08 .39 15	-.1 .38 11	-.2 .16 21
MP	.07 .38 21		.33 .07 21	.98 .00 21	.39 .04 21	.02 .46 21	.46 .01 21	-.1 .30 21	-.0 .47 21	.69 .00 21	-.1 .40 20	.26 .13 21	.02 .46 18	-.1 .41 21	.29 .15 15	.36 .14 11	-.2 .18 21
MO	.01 .48 21	.33 .07 21		.36 .05 21	.97 .00 21	.26 .13 21	-.0 .42 21	-.1 .38 21	-.0 .47 21	.31 .08 21	-.3 .14 20	.16 .25 21	-.0 .47 18	-.4 .02 21	.09 .37 15	.07 .42 11	.17 .23 21
CP	.34 .07 21	.98 .00 21	.36 .05 21		.43 .03 21	9.0 .48 21	.40 .04 21	-.2 .20 21	-.1 .41 21	.63 .00 21	-.1 .38 20	.28 .11 21	.01 .49 18	-.1 .37 21	.36 .10 15	.44 .09 11	-.2 .18 21
CO	-.0 .45 21	.39 .04 21	.97 .00 21	.43 .03 21		.20 .19 21	-.1 .33 21	-.0 .42 21	-.0 .46 21	.27 .12 21	-.3 .12 20	.27 .11 21	.03 .45 18	-.4 .06 21	.20 .24 15	.16 .32 11	.17 .24 21
Ag	.51 .01 21	.02 .46 21	.26 .13 21	-.0 .48 21	.20 .19 21		.50 .01 21	.08 .36 21	.16 .25 21	.19 .21 21	.11 .32 20	-.1 .30 21	-.2 .22 18	.07 .38 21	-.1 .42 15	-.5 .06 11	.02 .47 21
EA	.48 .01 21	.46 .02 21	-.0 .42 21	.40 .04 21	-.1 .33 21	.50 .01 21		.02 .47 21	.25 .14 21	.63 .00 21	.09 .36 20	-.3 .12 21	.09 .36 18	-.0 .46 21	-.3 .14 15	-.1 .36 11	-.3 .10 21
GP	.41 .03 21	-.1 .30 21	-.1 .38 21	-.2 .20 21	-.0 .42 21	.08 .36 21	.02 .47 21		.17 .23 21	.05 .42 21	-.0 .45 20	-.3 .11 21	.05 .42 18	.18 .22 21	.07 .41 15	-.2 .26 11	.03 .44 20
CX	.45 .02 21	-.0 .47 21	-.0 .47 21	-.0 .41 21	-.0 .46 21	.15 .25 21	.25 .14 21	.17 .23 21		.15 .26 21	.10 .33 20	-.3 .14 21	.18 .24 18	.42 .03 21	-.2 .21 15	-.1 .35 11	.32 .08 21
Ds	.30 .09 21	.69 .00 21	.31 .08 21	.63 .00 21	.27 .12 21	.19 .21 21	.63 .00 21	.05 .42 21	.15 .26 21		-.3 .10 20	-.1 .34 21	-.2 .17 18	-.2 .20 21	-.2 .22 15	-.3 .22 11	-.1 .29 21
%C	.10 .34 20	-.1 .40 20	-.3 .14 20	-.1 .38 20	-.3 .12 20	.11 .32 20	.09 .36 20	-.0 .45 20	.10 .33 20	-.3 .10 20		.01 .48 20	.35 .09 17	.34 .07 20	.01 .49 14	.47 .08 10	-.2 .26 20
%T	-.3 .13 21	.26 .13 21	.16 .25 21	.28 .11 21	.27 .11 21	-.1 .30 21	-.3 .12 21	-.3 .11 21	-.2 .14 21	-.1 .34 21	.02 .48 20		.02 .46 18	.17 .22 21	.22 .22 15	.23 .25 11	.06 .40 21
%I	-.2 .25 18	.02 .46 18	-.0 .47 18	.01 .49 18	.03 .45 18	-.2 .22 18	.09 .36 18	.05 .42 18	.18 .24 19	-.2 .17 18	.35 .09 17	.02 .46 18		.14 .29 18	-.2 .30 13	.33 .17 10	.46 .03 18

Ti	.09	-.0	-.4	-.1	-.4	.07	-.0	.18	.42	-.2	.34	.17	.14		.28	.26	.29
	.35	.41	.02	.37	.06	.38	.46	.22	.03	.20	.07	.22	.29		.16	.22	.10
	21	21	21	21	21	21	21	21	21	21	20	21	18		15	11	21
I1	-.2	.29	.09	.36	.20	-.1	-.3	.07	-.2	-.2	.01	.22	-.2	.28		.81	-.2
	.22	.15	.37	.10	.24	.42	.14	.41	.21	.22	.49	.22	.30	.16		.00	.13
	15	15	15	15	15	15	15	15	15	15	14	15	13	15		11	15
I2	-.4	.36	.07	.44	.16	-.5	-.1	-.2	-.1	-.3	.47	.23	.33	.26	.81		-.2
	.11	.14	.42	.09	.32	.06	.36	.26	.35	.22	.08	.25	.17	.22	.00		.24
	11	11	11	11	11	11	11	11	11	11	10	11	10	11	11		11
Ge	-.3	-.2	.17	-.2	.17	.02	-.3	.03	.32	-.1	-.2	.06	.46	.29	-.2	-.2	
	.12	.18	.23	.18	.24	.47	10	.44	.08	.29	.26	.40	.03	.10	.23	.24	
	21	21	21	21	21	21	21	21	21	21	20	21	18	21	15	11	

Table 41. Correlations of Scores to Explanatory Variables Using the 1 Point Scale
(Posttest Scores Above Diagonal, Difference Between Pretest and Posttest Below)

Line 1=r, Line 2=p, Line 3=n

Sc=Scores (Pretest Above Diagonal, Difference Between Pre and Posttest Below)

MP=Number of Multiple Hypotheses Written on Pretest

MO=Number of Multiple Hypotheses Written on Posttest

CP=Number of Multiple Hypothesis Cases Written on Pretest

CO=Number of Multiple Hypothesis Cases Written on Posttest

Ag=Age

GP=Undergraduate Grade Point Average

CX=Computer Experience

Ds=Classroom Exposure to DSM diagnosis

%C=Percent of Cases Reviewed on CAI software

%T=Percent of PracticeTest Items Reviewed on CAI software

%I=Percent Correct PracticeTest Items on CAI software

Ti=Time Spent on CAI software

I1=Number of Problem-Solving Participants Identified from Interviews by Rater 1

I2=Number of Problem-Solving Participants Identified from Interviews by Rater 2

Ge=Gender

	SC	MP	MO	CP	CO	Ag	EA	GP	CX	Ds	%C	%T	%I	Ti	I1	I2	Ge
SC		.41 .03 21	.39 .04 21	.31 .08 21	.38 .04 21	.62 .00 21	.46 .02 21	.25 .14 21	.24 .14 21	.46 .02 21	.03 .44 20	-.1 .30 21	-.2 .25 18	-.1 .40 21	.10 .36 15	-.1 .37 11	-.1 .26 21
MP	-.3 .10 21		.33 .07 21	.98 .00 21	.39 .04 21	.02 .46 21	.46 .01 21	-.1 .30 21	-.0 .47 21	.69 .00 21	-.1 .40 20	.26 .13 21	.02 .46 18	-.1 .42 21	.29 .15 15	.36 .14 11	-.2 .18 21
MO	.04 .43 21	.33 .07 21		.36 .05 21	.97 .00 21	.26 .13 21	-.0 .42 21	-.1 .38 21	-.0 .47 21	.31 .08 21	-.3 .14 20	.16 .25 21	-.0 .47 18	-.4 .02 21	.09 .37 15	.07 .42 11	.17 .23 21
CP	-.4 .03 21	.98 .00 21	.36 .05 21		.43 .03 21	-.0 .48 21	.40 .04 21	-.2 .24 21	-.1 .41 21	.63 .00 21	-.1 .38 20	.28 .11 21	.01 .49 18	-.1 .37 21	.36 .10 15	.44 .09 11	-.2 .18 21
CO	.02 .47 21	.39 .04 21	.98 .00 21	.43 .03 21		.20 .19 21	-.1 .33 21	-.0 .42 21	-.0 .46 21	.27 .12 21	-.3 .12 20	.27 .11 21	.03 .45 18	-.4 .06 21	.20 .24 15	.16 .32 11	.16 .24 21
Ag	.22 .17 21	.02 .46 21	.26 .23 21	-.1 .48 21	.20 .19 21		.50 .01 21	.08 .36 21	.16 .25 21	.19 .21 21	.11 .32 20	-.1 .30 21	-.2 .22 18	.07 .38 21	-.1 .42 15	-.5 .06 11	.02 .47 21
EA	-.1 .27 21	.46 .02 21	-.0 .42 21	.40 .04 21	-.1 .33 21	.50 .01 21		.02 .47 21	.25 .14 21	.63 .00 21	.09 .36 20	-.3 .12 21	.09 .36 18	-.0 .46 21	-.3 .14 15	-.1 .36 11	-.3 .10 21
GP	.42 .03 21	-.1 .30 21	-.1 .38 21	-.2 .20 21	-.0 .42 21	.08 .36 21	.02 .47 21		.17 .23 21	.05 .42 21	-.0 .45 20	-.3 .11 21	.05 .42 18	.18 .22 21	.07 .41 15	-.2 .26 11	.03 .44 20
CX	.07 .38 21	-.0 .47 21	-.0 .47 21	-.0 .41 21	-.0 .46 21	.15 .25 21	.25 .14 21	.17 .23 21		.15 .26 21	.10 .33 20	-.3 .14 21	.18 .24 18	.42 .03 21	-.2 .21 15	-.1 .35 11	.32 .08 21
Ds	-.2 .26 21	.69 .00 21	.31 .08 21	.63 .00 21	.27 .12 21	.19 .21 21	.63 .00 21	.05 .41 21	.15 .26 21		-.3 .10 20	-.1 .34 21	-.2 .17 18	-.2 .20 21	-.2 .22 15	-.3 .22 11	-.1 .29 21
%C	.08 .37 20	-.1 .40 20	-.3 .14 20	-.1 .38 20	-.3 .12 20	.11 .32 20	.09 .36 20	-.0 .45 20	.10 .33 20	-.3 .10 20		.01 .48 20	.34 .09 17	.34 .07 20	.01 .49 14	.47 .08 10	-.2 .26 20
%T	.00 .49 21	.26 .13 21	.16 .25 21	.28 .11 21	.27 .11 21	-.1 .30 21	-.3 .12 21	-.3 .11 21	-.2 .14 21	-.1 .34 21	.02 .48 20		.02 .46 18	.17 .22 21	.22 .22 15	.23 .25 11	.06 .40 21
%I	.05 .42 18	.02 .46 18	-.0 .47 18	.01 .49 18	.03 .45 18	-.2 .22 18	.09 .36 18	.05 .42 18	.18 .24 19	-.2 .17 18	.35 .09 17	.02 .46 18		.14 .29 18	-.2 .30 13	.33 .17 10	.46 .03 18

Ti	-.1	-.0	-.4	-.1	-.4	.07	-.0	.18	.42	-.2	.34	.17	.14		.28	.26	.29
	.40	.41	.02	.37	.06	.38	.46	.22	.03	.20	.07	.22	.29		.16	.22	.10
	21	21	21	21	21	21	21	21	21	21	20	21	18		15	11	21
l1	-.2	.29	.09	.36	.20	-.1	-.3	.07	-.2	-.2	.01	.22	-.2	.28		.81	-.2
	.27	.15	.37	.10	.24	.42	.14	.41	.21	.22	.49	.22	.30	.16		.00	.13
	15	15	15	15	15	15	15	15	15	15	14	15	13	15		11	15
l2	-.3	.36	.07	.44	.16	-.4	-.1	-.2	-.1	-.3	.47	.23	.33	.26	.81		-.2
	.17	.14	.42	.09	.32	.06	.36	.26	.35	.22	.08	.25	.17	.22	.00		.24
	11	11	11	11	11	11	11	11	11	11	10	11	10	11	11		11
Ge	-.0	-.2	.17	-.2	.17	.02	-.3	.03	.32	-.1	-.2	.06	.46	.29	-.2	-.2	
	.44	.18	.23	.18	.24	.47	.10	.44	.08	.29	.26	.40	.03	.10	.23	.24	
	21	21	21	21	21	21	21	21	21	21	20	21	18	21	15	11	

APPENDIX J

TABLE 42. Frequency of Correct Diagnoses
by Instrument Item and Test Version

Test Version	Pre/Posttest	Item	N	%/Participants Correct
1	Pretest	A	21	42.9
2	Posttest	A	20	85.0
3	Posttest	A	19	84.2
1	Pretest	B	21	47.6
2	Posttest	B	20	65.0
3	Pretest	B	19	52.6
1	Posttest	C	21	61.9
2	Posttest	C	20	60.0
3	Posttest	C	19	89.5
1	Posttest	D	21	38.1
2	Posttest	D	20	50.0
3	Pretest	D	19	10.5
1	Pretest	E	21	23.8
2	Posttest	E	20	60.0
3	Posttest	E	19	47.4
1	Posttest	F	21	33.3
2	Pretest	F	20	20.0
3	Pretest	F	19	26.3
1	Pretest	G	21	33.3
2	Pretest	G	21	30.0
3	Posttest	G	19	89.5
1	Posttest	H	21	76.2
2	Pretest	H	20	45.0
3	Pretest	H	19	94.7
1	Pretest	I	21	9.5
2	Pretest	I	20	5.0
3	Pretest	I	19	26.3
1	Posttest	J	21	47.6
2	Posttest	J	20	35.0
3	Posttest	J	19	57.9
1	Posttest	K	21	66.7
2	Pretest	K	20	10.0
3	Pretest	K	19	21.1

1	Posttest	L	21	47.6
2	Posttest	L	20	40.0
3	Pretest	L	19	10.5
1	Pretest	M	21	4.8
2	Pretest	M	20	20.0
3	Posttest	M	19	52.6
1	Pretest	N	21	23.8
2	Pretest	N	20	20.0
3	Posttest	N	19	47.4
1	Posttest	O	21	71.4
2	Pretest	O	20	15.0
3	Pretest	O	19	00.0
1	Posttest	P	21	19.0
2	Posttest	P	20	15.0
3	Pretest	P	19	00.0
1	Pretest	Q	21	61.9
2	Pretest	Q	20	45.0
3	Posttest	Q	19	47.4
1	Pretest	R	21	19.0
2	Posttest	R	20	10.0
3	Pretest	R	19	00.0

TABLE 43. Reliability Coefficients by Total Test Scores of Difficult and Easy Items with Items and With Items Removed

Item	Test Version	Pre/Posttest	α with Item	α with Item Removed
A	1	Pretest	.42	.46
A	2	Posttest	.38	.36
A	3	Posttest	.74	.68
H	1	Posttest	.57	.57
H	2	Pretest	-.23	.03
H	3	Posttest	.74	.68
I	1	Pretest	.42	.44
I	2	Pretest	-.23	-.22
I	3	Pretest	.48	.49
P	1	Posttest	.57	.59
P	2	Posttest	.38	.34
R	1	Pretest	.42	.38
R	2	Posttest	.38	.33

TABLE 44. % Frequency of Correct Responses to Test Items by
Test Version and Diagnostic Cluster

Cluster/ Item	Pretest 1	Posttest 1	Pretest 2	Posttest 2	Pretest 3	Posttest 3
A/A	42.9		85			82.4
A/D		38.1		60	10.5	
A/M	4.8		20			52.6
A/N	23.8		20			47.4
B/B	47.6			65	52.6	
B/E	23.8			60		47.4
B/I	9.5	5		26.3		
B/J		47.6		35		57.9
B/K		66.7	10	21.1		
B/L		47.6		40	10.5	
C/C		61.9		60		89.5
C/F		33.3	20		26.3	
C/G	33.3		30			89.5
C/H		76.2	45			94.7
C/O		71.4	15	0		
C/P		19	15	0		
C/R	19		10	0		