

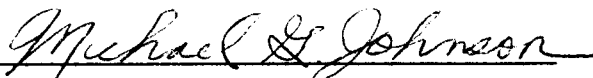
To the Graduate Council:

I am submitting herewith a dissertation written by William Deal entitled "Expert-Novice Differences in Clinical Decision Making." I have examined the final copy of this dissertation for form and content and recommend that it be accepted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, with a major in Psychology.



Dr. Michael Nash, Major Professor

We have read this dissertation
and recommend its acceptance:


Dr. Michael Johnson
Dr. Steve McCallum
Dr. Michael Smith

Accepted for the Council:



Associate Vice Chancellor
and Dean of The Graduate School

EXPERT-NOVICE DIFFERENCES IN CLINICAL DECISION MAKING

**A Dissertation
Presented for the
Doctor of Philosophy
Degree
The University of Tennessee, Knoxville**

**William E. Deal
December 1995**

Acknowledgments

I would like to acknowledge my committee members:

Michael Nash, Ph.D., Committee Chair

Michael Smith, Ph.D.

Michael Johnson, Ph.D.

Steve McCallum, Ph.D.

Abstract

This paper reviews the cognitive science literature on expertise and presents methodology designed to examine expert-novice differences in the domain of clinical (psychology/psychiatric) diagnosis. The study examines differences between experienced and novice clinicians on a clinical diagnostic task. 10 novice psychology graduate students and 9 experienced psychologists read two case histories, each divided into a series of clinical vignettes. Subjects were asked to "think aloud" about the case formulation and later asked to recall of information about the case. Recorded protocols were transcribed and coded according to frequencies of diagnostic hypotheses, symptom hypotheses, speculative ideas, case references, and case features. Results showed significant differences between experienced and novice subjects' use of diagnostic and symptom hypotheses. Experienced subjects also generated a greater proportion of case reference to case feature statements. No differences were found in case information recalled. These results support predictions generated from a cognitive representation model of expertise.

Table of Contents

1	Introduction
11	CHAPTER 1 Review of Expertise Literature
12	The Absolute Performance View
14	The Process in Performance View
15	The Expertise Approach
17	Chess Expertise
19	"Enhanced Recall"
21	"Forward Reasoning"
25	Physics Problem Solving Expertise
31	Medical Diagnostic Expertise
37	Theories of Expertise
38	The Associationist View
43	The Representational View
48	Two Empirical Themes Revisited
50	CHAPTER 2 Methods and Procedures
50	Overview
51	Subjects
53	Materials
56	Procedure
63	Protocol Analysis
66	CHAPTER 3 Rational for Coding System
67	Figure 1. Abstraction of Coding Categories
72	Variables
72	Hypothesized Predictions
73	Table 1. Variables
76	CHAPTER 4 Results
77	Table 2. Case 1
78	Table 3. Case 2
79	Table 4. Cases Combined
82	CHAPTER 5 Discussion
90	References
95	Appendix 1 Protocol Coding System
114	Appendix 2
114	Case 1
123	Case 2
132	Appendix 3
133	Table 5. Case One Experienced
134	Table 6. Case One Novice
135	Table 7. Case Two Experienced
136	Table 8. Case Two Novice
137	Vita

Introduction

Expertise is a broad and complex subject matter which integrates various perspectives on social and cognitive science. Empirical and theoretical developments in expertise, concept formation, and education research have provided insight into the role of experience and training in improving performance on specific cognitive and non-cognitive tasks. Due to nature of the subject matter and the need for systematic methodology, the domains in which these cognitive concepts have been developed are understandably limited. Expertise researchers have typically investigated domains involving straight forward and easily operationalized tasks. Chess and physics problem solving, for example, are excellent domains for the study of expertise because, while the tasks are complex, the rules and principles involved are finite. Both domains also have the advantage that tasks can be defined within a more or less self contained system. Any particular chess problem, for example, gives rise to only a few viable single move options. The complexity of the task arises from vast array of possible combinations of single move choices. Relatively

close-ended task of this sort simplify the experimenter's selection of criteria for judging one particular response as better, or more correct, than another because all possible single move responses are known in advance and outcomes of these responses can be predicted. Despite this advantage, however, not all domains of interest to researchers may be operationalized within close-ended predictable systems. The current study is an effort to develop methodology which may be used to differentiate experts and novices in the relatively open-ended domain of clinical psychology/psychiatric diagnosis.

Although the study of expertise has been driven by an interest in the general mechanisms of expert performance, the interest of individual researchers has also been domain specific. Researchers have attempted to expand the concepts and methods developed for the study of expertise in the ideal "laboratory" domains of chess and physics problem solving to include the study of expertise in other areas. There has also been interest in developing methods within various domains for the study of expert performance on more ecologically valid tasks¹ (eg. Scardamalia & Bereiter,

1991). The result of various researchers adapting methodologies to their particular areas of interest has been the growth of the expertise research paradigm to include domains involving more open-ended tasks. While necessitating more complex methodology, this development into "open-ended" domains has led to a better understanding of expertise. The domain of medical diagnosis is a primary example of this development.

Medical diagnostic expertise is a more complex research domain than chess or physics problem solving expertise because the tasks involved are typically less well defined. This means that the experimenter cannot as easily predict the subject's responses, making it more difficult to determine whether a particular response is more or less correct than another. However, as a domain for expertise research, medical diagnosis has some particular advantages. Medicine as a science operates on a practical theoretical foundation of cause and effect relations. The medical

¹See Neisser (1987) for an enlightened discussion of the concept of ecological validity. For example, solving physics textbook problems is not a particularly ecologically valid task for assessing how a physics expert exercises his or her expertise in day to day employment.

professional can therefore observe signs and symptoms as criteria for a disease, and from these may make firm inferences about the existence of that disease. This is because the physician's symptom criteria for identifying the presence of an illness are distinct from the epidemiological criteria used to define the existence of the illness. Since the existence of disease is defined by structural difficulties in the organism (an established theory of epidemiology), while the presence of the disease is indicated by functional signs and symptoms observed by the physician, the cause and effect reasoning approach presents no problems. This theoretical basis allows the expertise researcher to operationalize medical diagnostic tasks for which a particular research subject's responses may be considered as correct or incorrect. Medical diagnosis is also a domain in which much credence is traditionally placed on authority as a criterion for truth. Thus, the expertise researcher may fairly utilize the "expert's expert" as a formal baseline for assessing research subject production. The use of authority as an expertise criterion is made possible because of the general unity of scientific perspective within the field of physical medicine. This

unity of perspective, however, does not prevail in the "mental medicine" domain of psychology/psychiatry.

Within the larger domain of medical diagnostics, little previous research has attempted to operationalize expertise in clinical psychology/psychiatry. In contrast to other domains in which expertise has been successfully studied, clinical psychological/psychiatric diagnosis involves tasks which are difficult to define with precision. Comparable psychological/psychiatric experts often generate discrepant clinical views of the same patient data based on widely differing perspectives. The primary reason for this is that, unlike other branches of medicine involving diagnosis, tasks within the domain of psychology/psychiatry are inherently open-ended. While the advent of the self-contained DSM system has improved reliability and predictability of diagnoses somewhat, this has not solved the fundamental problem. Unlike non-psychiatric medical diagnosis, the theoretical use of cause and effect reasoning leads to logical problems and irrational conclusions in the domain of psychiatry. The core of this difficulty arises from the classic dilemma of the "mind-brain paradox," a thorough examination of which is beyond the scope of this

overview. Simply put, however, the assumption of a direct one-to-one relationship between physical (brain) disease and the mental and behavioral criteria for identifying such disease is highly problematic and in many cases false. This difficulty is due not to a lack of scientific understanding of the relationship between structure and function within the brain. It is due to a fundamental difference in category between physical and mental phenomena. While DSM provides a common descriptive language with fairly precise sign and symptom criteria, the lack of a firm theoretical foundation for mental "disease" means that the system in essence defines disorders into existence.² On a purely descriptive clinical level, firm cause and effect theoretical grounding is not necessary: mental "disease" as such serves as a euphemism for what in fact are naturally occurring clinical syndromes.³ While this is adequate for

²This problem has led to the highly politicized nature of the DSM and the to the controversies surrounding the process of determining scientific truth through committee and lobby groups.

³A medical syndrome is a cluster of signs or symptoms occurring together with sufficient regularity that the grouping has practical (clinical) utility. Kaplan and Sadok (1989) note that most DSM disorders are in fact clinical syndromes rather than disease entities per se.

clinical purposes, it does not aid the researcher in assessing diagnostic expertise. Expertise assessed from purely descriptive criteria would not tap the expert subjects' presumed deeper knowledge base and would simply measure proficiency within the DSM system. Since different professionals in different settings have varying degrees of allegiance to descriptive psychiatry for conducting their clinical practice, the use of a DSM proficiency task would not effectively tap the underlying mechanisms of clinical expertise. Expertise in the domain of clinical psychological/psychiatric diagnosis cannot be meaningfully operationalized by limiting experts to an artificially close-ended task.

For the researcher attempting to operationalize diagnostic expertise in the domain of clinical psychology/psychiatry, the problem of task open-endedness must be addressed methodologically. Since the examiner cannot easily predict subjects' diagnostic ideas about a particular clinical case, the notion of one idea as more correct than another for a particular clinical example becomes more subjective. This is compounded by the fact that psychology/psychiatry genuinely lacks a unified

scientific perspective, making any appeal to authority, or an "expert's expert," of questionable validity.

As a result of the difficulty of operationalizing clinical psychology/psychiatry expertise, methodology in past research has generally been limited to either clinical outcome studies, or to phenomenological descriptions of individual cases. Both of these methodologies circumvent the problem of task open-endedness by adopting pragmatic criteria for the "correctness" of subject behavior. Clinical outcome research, while primarily a method for comparing different types of treatment, may also be used for assessing the effectiveness of clinicians of differing levels of expertise. When measuring clinical outcome as a dependent variable, the "correct" subject responses are those leading to the most favorable clinical results. This approach is similar to the absolute performance conceptualization of expertise discussed in the next section. Experimental outcome comparisons of clinician's performance is difficult research to conduct and does not assess underlying mechanisms of clinical expertise.

Like the clinical outcome approach to the study of expertise, phenomenological case study methodologies address

the problem of task open-endedness by adopting pragmatic criteria for "correctness."⁴ These methodologies are used to describe thematic material as it emerges in the context of subject performance of ecologically valid tasks. The assessment of "correctness" is related to the effectiveness of clinical interventions in vivo. This approach embraces the open-ended nature of the domain specific tasks, but limits itself mostly to description rather than to analysis of underlying mechanisms. Phenomenological "correctness" is simply a function of achieving desired ends in an expedient manner.⁵ While useful for examining the process of experts engaged in meaningful tasks, Phenomenological methods are generally not practical for generating systematic comparisons between groups.

⁴William James, well known for his influence in the development of philosophical pragmatism and psychological functionalism, is also considered one of the originators of phenomenology. The connection between pragmatism, functionalism, and phenomenology embodied in James's work is both compelling and of interesting historical note.

⁵De Groot's (1978) original chess expertise study adopted aspects of this open-ended approach by asking subjects to mention the "next best move" when presented with a mid-game chess problem. Subsequent research in the "expertise approach" domain, however, did not follow this methodology. DeGroot's work is discussed in greater detail in the next section.

The purpose of this project is to develop a preliminary methodology for the study of expert-novice differences on a clinical psychology/psychiatry diagnostic task. The methodology is an extension of the cognitive science "expertise approach" for examining differences in how expert and novice clinicians cognitively process clinical information. Given the open-ended nature of the diagnostic task, however, the method has been modified in an attempt to maintain clinical relevance and ecological validity, while also allowing for a systematic comparison between groups. This is accomplished through the use of open-ended interview techniques followed by systematic quantitative analysis of recorded protocols. The specific rationale for the methodology is discussed in later sections.

CHAPTER 1

Review of Expertise Literature

In light of the complexity resulting from the various cognitive approaches to the study of expertise, the purpose of this section is to touch upon highlights of theory and seminal research, and to outline major themes relevant to this project. This overview begins with a brief description of the distinction between the absolute performance verses process in performance view in the clinical expertise literature. This is followed by a review of the more general cognitive science "expertise approach." This review will center around the two empirical themes of "enhanced recall" and "forward reasoning." These themes will then be traced through the domains of chess, physics problem solving, and medical diagnostic expertise. Following this is an outline of expertise theory and a review the early associationist perspective. This will then be linked to the more comprehensive theory of cognitive representations. Finally, the two empirical themes of "enhanced recall" and "forward reasoning" will be reexamined in light of the theory of cognitive representations.

The Absolute Performance View

The evaluation of expertise based on predictive accuracy of outcome is a perspective which originates from outside the domain of cognitive psychology. Camerer and Johnson (1991) and Garb (1989) review numerous studies from behavioral decision research which examine experts' ability to make accurate predictions within their domain of expertise. This line of inquiry begins with Sarbin (1944), who outlined the distinction between clinical reasoning and actuarial judgement, and suggested that many clinicians fail to attend to important actuarial variables when making clinical predictions. Meehl (1954) followed up on Sarbin's work through a review of dozens of studies which conclusively demonstrated that cross-validated actuarial models can predict outcomes more accurately than can human clinical experts. Even when compared with novices, experts predict outcomes with only slightly higher accuracy (Camerer and Johnson, 1991). The notion that long years of clinical experience make for enhanced ability to predict outcomes is refuted by the behavioral decision literature. Analysis of

expert verses novice performance indicates that, while extent of training has some small effect on the accuracy of clinical predictions, there is little correlation between length of clinical experience and accuracy of clinical judgements (Garb, 1989). Camerer and Johnson (1991) link the inability of experts to use their extensive knowledge to make accurate outcome predictions with heuristic biases, such as "representativeness" errors (Kahneman et al, 1982), and the extensive use (or misuse) by experts of configural rules.⁶ Experts use cue weights inconsistently, combine cues incorrectly, overgeneralize from small samples, attribute meaning to random error, use inappropriate implicit theories, and over-stress evidence confirming their own prior beliefs (Camerer and Johnson, 1991).

While these findings from the behavioral decision perspective on clinical prediction are compelling and provocative, little actual follow-up research has been stimulated by this approach. A likely reason for this is that operationalizing expertise in terms of predictive

⁶"representativeness" errors are heuristic errors involving subjects overgeneralizing from particular instances. Configural rules involve "if, then" propositions which result in multilevel decision trees.

accuracy has little ecological validity. Certainly the suggestion that lightly trained novices could or should replace experts in clinical domains is an idea which goes against the grain of traditional judgement regarding clinical science and practice. While opinions differ regarding the clinical effectiveness of experts verses novices, it is clear that expertise involves more than simply the ability or inability to make accurate predictions.

The Process in Performance View

In light of the dismal performance of experts in studies examining predictive accuracy and clinical judgement, it might have been expected that interest in the cognitive study of expertise would diminish. In spite of such studies, however, the fact that experts continue to be regarded in their various domains, and to be recognized for their superior performance, suggests that there is more to the notion of expertise than is accounted for by the behavioral decision literature. Clearly an alternative approach to predictive accuracy of outcomes as an

operational variable is necessary to study expertise. Within the field of cognitive science the alternative has been to examine the cognitive processes of experts and novices on tasks where performance is attributable to stable individual characteristics (Ericsson and Smith, 1991). The evaluation of performance in this sense relies less on the notion of predictive accuracy and more upon the assessment of performance on tasks specifically designed to capture what experts actually do within their domain. Most commonly, this assessment of performance involves expert-novice comparisons on applied domain specific tasks, such as, problem solving, or diagnostic case formulation. The mainstream application of this view in cognitive science has been the classical "expertise approach."

The Expertise Approach

According to Ericsson and Smith (1991), the expertise approach is characterized by efforts to descriptively evaluate performance under operational conditions, to analyze these descriptions, and to identify aspects of the process which make for superior performance. In contrast to

the absolute performance view, the expertise approach assumes superior performance as a defining characteristic of expertise and seeks to operationalize tasks which capture the expert's abilities. Once tasks which suitably differentiate expert from non-expert performance are operationalized, then analysis of these differences are used to yield process variables consistent with superior performance.

In a general description, research utilizing the expertise approach starts by defining tasks which are representative of the domain and which "capture" relevant aspects of expert performance. Once the task is defined, the processes resulting in superior performance are systematically analyzed such that underlying cognitive mechanisms can be postulated. These mechanisms are then linked to a theoretical account which encompasses the development of expertise in terms of learning mechanisms. Within domains which have "captured" expert performance in this manner, such performance has been found to include faster response times, greater planning ability, and superior memory performance (Ericsson and Smith, 1991).

Chess Expertise

The study which serves as the foundation for the classical expertise approach was conducted in the domain of chess by de Groot (1946/1978). De Groot cites a Russian experiment from the 1920's in which 8 chess grand masters were tested for cognitive and perceptual abilities and compared with a control group. No differences between groups were found on measures of these general abilities, suggesting that the skills underlying chess expertise must be domain specific. De Groot pioneered the use of "think aloud" protocol analysis, asking both grand masters and players of lesser ability to verbally describe the "best next move" after studying a mid-game board position. De Groot believed that, in contrast to measures of general cognitive ability, having subjects' "think aloud" about the "best next move" more effectively elicited cognitive processes specific to chess expertise. De Groot's initial hypothesis was that experts explore more extensive move combinations and consider options of greater complexity than do novices. He found, however, that the extent of search for move options did not differentiate these subjects

according to chess expertise. What did differentiate grand masters from players of lesser skill was that the former group typically noticed the "next best move" during their initial familiarization with the board, while the latter group did not. On a recall task, however, de Groot found that board positions were reported by the grand masters more accurately and in larger groupings of pieces than were those reported by the lesser skilled players. From these results, de Groot inferred that chess experts are able to integrate the essential characteristics of a board position into a single whole, whereas (relative) novices cannot do this. De Groot theorized that experts recognize characteristic board positions from past representations stored in memory and evaluate possible moves based on solutions which have worked in the past. Thus, he hypothesized that expertise (in chess) is a function of long experience which allows for the recognition of previously seen patterns and the building up of associations between characteristic board positions and appropriate move responses. De Groot's findings set the stage for the study of expertise and established the germs of two themes of empirical research which continue to pervade the expertise literature: enhanced recall and

forward reasoning.

"Enhanced Recall"

De Groot's work highlights the theme of enhanced recall performance of experts compared to (relative) novices. In the domain of chess, the superiority of experts' memory for game relevant material has been well documented. In their classic study following up on de Groot's work, Chase and Simon (1973) focused explicitly on expert memory performance. They presented expert and novice chess players with brief exposures to board positions. Some of these board positions were actual mid-game arrangements, while others were random placements. Subjects were asked to recall the board arrangements as accurately as they could. The experts exhibited a clear superiority over novices on this task, but only for the mid-game arrangement condition. There was no difference between groups for recall of random board arrangements. Chase and Simon interpreted these results as evidence linking experts' superior memory performance to their recognition of meaningful relations between pieces. In addition to pure recall effects, Chase

and Simon noticed that subjects tended to verbally recall board positions in groupings, rather than in a random fashion. The number of pieces per grouping was significantly larger for the experts than for the novices. This result is consistent with an information processing model of cognition. The authors inferred that these groupings represent "chunks" of information based upon meaningful relations or "patterns." The model which emerges from these results, interpreted in an information processing context, is that expertise involves a cognitive "pattern matching" process in which previously learned configurations serve as a kind of template for evaluating new information. Experts are able to demonstrate enhanced recall because new information is processed in part through recognition of previously learned patterns. The presence of these patterns in an established knowledge base allows information to be "chunked" in an economical fashion, resulting in greater storage and recall capacity.

"Forward Reasoning"

The domain of chess has served as a particularly useful medium for the study of expertise. One of de Groot's original hypotheses for chess expertise was that experts would demonstrate a superior ability to search "ahead" of the game in terms of possible sequences and planned move combinations. He discovered, however, that even the most expert chess subjects only plan a few moves ahead and do not engage in "deep search" procedures as he had predicted. Based on de Groot's original hypothesis that chess experts search more deeply than do novices, Charness (1981) utilized a protocol analysis technique to examine expert verses novice subjects' planning of chess moves. This study found that there was a predictable relationship between novice subjects' depth of search and their level of skill. This relationship disappeared, however, once subjects reached an intermediate skill level. For subjects at a higher level of expertise, Charness found that depth of search may actually decrease (Charness, 1991). These results suggest that there is a necessary relationship between adequate depth of search and skill at chess, but that depth of search is not a

sufficient condition for chess expertise. In chess, there is a trade off between knowledge based evaluation of board positions (ie pattern matching) and the depth of search necessary to play well. While a certain depth of search is necessary to play chess well, this level of "planning ahead" is not particularly deep, even for experts. Thus, the depth of search employed by individuals familiar with the game is probably more a function of human processing limitations (Miller, 1956) than of chess expertise per se. At an intermediate level of chess skill and above, the selectivity of which move combinations are planned ahead by the subject, becomes more significant than the general depth or extent of search (Charness, 1991). Knowledge is used to constrain search parameters to a level manageable within the limits of human processing ability. Chess experts are able to use their extensive knowledge of chess to better evaluate which characteristics of a board position are significant, thereby narrowing the search parameters to move combinations which are likely to lead to victory⁷.

⁷Note that this trade off between evaluation and search applies mainly to human chess experts. Computer chess programs tend to operate through extensive search routines which require greater processing capacity than humans are

The trade off between knowledge based evaluation and depth of search as complimentary strategies for chess highlights the second major theme which has emerged from the expertise literature: forward reasoning. This theme has been conceptualized and labeled differently across various domains, but the basic idea encompasses familiar findings from several areas of research. Essentially, forward reasoning, as opposed to backward reasoning, refers to a descriptive application of Charness's distinction between an knowledge based evaluation focused verses a depth of search focused inference strategy⁸. To work in a forward direction through the problem space, a subject must first evaluate which of the presented data are relevant to achieving the

capable. Although computers can do this very quickly, their search is not guided by knowledge and so is inefficient. Modern chess programs are beginning to incorporate sophisticated information about chess into their programming, but this is not yet at the level of human evaluation ability and these programs still rely on extensive search capacity to match human player's abilities.

⁸See footnote number 7 for an example of this distinction with regard to expert human chess players verses advanced chess programs. In this sense, chess programs play a novice game of chess, but they play it better than any human novice is capable. This ability is, of course, a result of the computers ability to handle large amounts of data and to process this data much more quickly than is possible for people.

final goal. Once the key data are identified, the subject must combine this information based on specific knowledge of the underlying principles which will lead forward from the data to the final goal⁹. Both evaluation of presenting data and working forward from data to solution involve a certain familiarity with the problem type and the sort of solutions expected. To work in a backward direction means that the subject does not evaluate the presenting data in terms of specific knowledge, and so must begin with predetermined goals and engage extensive search routines in an attempt to link each goal to the presented data. Note that this is a general method which does not require specific knowledge of the domain and which operates through exhausting the possibilities of search rather than through execution of an efficient plan. Forward reasoning utilizes search guided by knowledge such that the need for extensive search of non-productive avenues is eliminated. This is what Charness (1991) means by a trade off between evaluation and search at

⁹The goal varies depending upon which domain is being researched. In chess, the goal is to defeat one's opponent, in medical diagnosis the goal is to treat the patient, in physics problem solving the goal is to arrive at the correct solution. Anzai (1991) addresses this issue.

the expert's level of information processing. As one can see, the concept of forward reasoning is evident in de Groot's (1946/1978) original study of chess expertise. If experts are able to identify the "best next move" while familiarizing themselves with the board, this eliminates the need for extensive search and allows the expert to highlight only salient aspects of the board position and to search only the move combinations related to those aspects.

Physics Problem Solving Expertise

The concept of forward reasoning as a variable differentiating experts from novices has been developed in the domain of physics problem solving expertise. One of the early related concepts in this domain is the distinction between the use of strong and weak method problem solving strategies. Weak methods are defined as general, knowledge independent, approaches to problem solving which utilize a generate and test or trial and error procedure to reduce problems to basic elements (Anzai, 1991). The use of weak methods is characteristic of novices, but is also the strategy of choice for both experts and novices on poorly

defined or novel tasks.

A weak problem solving strategy is an approach which reflects the subject's lack of specific knowledge by which to organize the particular problem. Weak methods are non-problem-specific and involve the use of rote procedures to generate a means-end analysis of the problem space. This begins with the subject skipping over initial information about the problem and examining first the end-goals. There is no "pre-processing" of information in terms of problem formulation because the use of a weak strategy does not rely on organizing principles of physics. Beginning with the problem end-goal, a weak method approach proceeds through a systematic search of various possibilities to link the end-goal with the presented data. Thus, for example, a subject with little understanding of physics principles who is presented with a physics problem will use a weak method and begin by looking at the type of variables asked for as a problem solution. By first categorizing the solution variables, the subject may proceed by searching "backwards" through the presented information for variables linked to solution variables through formulas, units of measure, or other surface features. Through meticulously continuing

this procedure the subject will eventually hit upon a combination of formulas and variables which link the solution variables to the presented data. Note that this "backward" approach requires no specific knowledge as to the principles embodied in the formulas being used, and the particular formulas and intermediate variables have no particular meaning other than the fact that they link the presented data with the end-goal of the problem. The drawback of a weak method, however, is that it is an inefficient approach involving a haphazard search down many "blind alleys" in hopes of finding workable connections.

A strong method for physics problem solving is an approach which makes use of domain specific knowledge to evaluate the presented data in terms of an expedient solution. To continue the example above, a subject's use of a strong method for solving a physics problem would involve their recognition of the applicability of basic physics principles within the initial problem presentation. This "pre-processing" of presented information, prior to any consideration of the solution variables, allows the subject to formulate the problem into a meaningful configuration (ie problem representation). Once the subject understands and

identifies the principles involved, he or she simply retrieves the appropriate formulas from memory, applies the formulas to the presented data, and calculates the solution. Note that in order to work "forward" in this manner, the subject must have specific knowledge of the principles which apply to the particular problem at hand. Thus, even subjects with great physics expertise may not be able to apply a strong method to a particularly difficult or novel problem. This is in contrast to weak methods which may be applied to any problem type.

The example above illustrates another aspect of weak verses strong methods for solving physics problems. Subjects using strong methods to formulate and solve problems are conceptualizing information on a more abstract cognitive level than subjects applying weak methods. While the use of weak methods relies primarily on extensive systematic search routines, the use of strong methods relies explicitly on abstract conceptualization. This contrast between expert and novice subjects in terms of their operation at different levels of cognitive abstraction will be addressed in later discussion of expertise theory.

One of the classic studies examining physics problem

solving expertise was conducted by Chi et al (1981). This study illustrates some important expert-novice differences with regard to the problem solving strategies discussed above. These researchers presented physics experts and novices with a number of physics problems. Subjects were asked to categorize problems according to which problems "go together." The results indicated that novices tend to categorize problems according to characteristic similarities in the problem presentation such as the presence of incline planes, of wheels and pulleys, and other surface features. Experts, on the other hand, categorize problems indiscriminate of these characteristic problem features. Instead, the experts in the study grouped problems according to the abstract principles of physics used to obtain problem solutions (eg conservation of energy, Gauss's law, etc). This finding suggests that physics experts are cognitively representing problems in accordance with strong methods, while novices are representing problems in accordance with weaker methods. Other studies in the domain of physics problem solving have supported this notion.

Simon and Simon (1978) collected "think aloud" protocols from experts and novices asked to solve physics

problems. These researchers found that experts predominately work in a forward direction through the problem, while novices work in a backward way. Larkin (1980), using a similar protocol analysis technique, found that experts are able to begin making inferences about a problem from the presented data, while novices attempt to begin solving the problem by examining the end-goal of the problem. Larkin argues that experts learn forward reasoning as a strategy to limit the scope of problems in terms of acquired domain-specific knowledge. Anzai and Yokoyama (1984) used verbal protocols to find that experts generate initial problem representations during their initial exposure to problem data. The expert subjects also maintained these representations throughout the problem solving process. This study operationalized initial representations in terms of subject descriptions of the salient problem features and the principles invoked to arrive at a solution (ie a theory based framing of the essential aspects of the problem). Novices in this study had trouble generating initial representations and often shifted representations as initial attempts at problem solutions failed. Anzai (1991), in an extensive review of

the physics expertise literature, emphasizes that forward reasoning as an inference strategy is a direct function of the expert's superior ability to represent problems in terms of expedient solutions. Experts use their well organized abstract knowledge of the domain to construct problem representations which highlight key features and lead directly to efficient solutions based on specific procedural knowledge.

Medical Diagnostic Expertise

Research on medical diagnostic expertise has also examined the themes of enhanced recall and forward reasoning. The results of these studies, however, have been less conclusive than those looking at expert-novice differences in chess or physics problem solving. Specifically, protocol analysis studies of medical diagnostic expertise have called into question the generality of enhanced recall effects. There are several possible reasons for this. Traditional thinking about medical reasoning has assumed that physicians employ direct cause and effect logic to the problem of generating a

diagnosis. This approach is embodied in the classic "hypotheticodeductive" method (see Elstien et al, (1978)). This method involves the physician postulating hypotheses which he or she then deductively evaluated according to the data. Adopting the terminology of physics expertise, however, one can see that this is an example of a weak general method of problem solving. The "hypotheticodeductive" reasoning approach is a backward inference strategy because it begins with a hypothesized disease entity, the presence of which is subsequently confirmed or disconfirmed through a deductive search for effects (symptoms) associated with the disorder. Given that the task of generating a medical diagnosis is less well-defined than, for instance, solving a pre-determined physics problem, the use of weak methods might be appropriate, even for the expert diagnostician (Larkin, 1977). The idea that the use of backward reasoning may be the norm for medical diagnostic reasoning is also suggested by the fact that the task of arriving at a medical diagnosis usually occurs in an interactive setting (Anzai, 1991). Because of this the diagnostician's information input typically involves an active information gathering procedure using a question and

answer technique. This task variable leads to the potential occurrence of radical shifts in the doctor's perspective as additional elements of information are processed. In addition, since medical information is usually gathered from various sources, the information presented to the physician evaluating the case may be disorganized and possibly contradictory. As a result, reasonable preliminary diagnostic ideas may frequently have to be discarded as new variables are considered. Finally, the goal of treating a patient is fundamentally different and more complex than simply deriving correct answers to well-defined or circumscribed problems.¹⁰

Patel and Groen (1991) suggest that enhanced recall as an empirical reflection of medical diagnostic expertise is invalid. These researchers demonstrated, using protocol

¹⁰Hammond (1980) discusses a theory of medical expertise hypothesizing a theoretical cognitive continuum spanning from pure rote clinical analysis (algorhythmic reasoning) to pure clinical intuition. While the development of clinical expertise involves a gradual shift towards the clinical intuition end of the continuum, the type of reasoning most appropriate for solving problems is a function of the task itself. Thus, expertise involves not only the ability to use clinical intuition, but also an ability to judge whether clinical intuition or clinical analysis is the appropriate means of addressing the problem.

analysis methodology, that for a clinical diagnostic task there is a negative relationship between recall and diagnostic accuracy (Patel and Groen, 1987). In their study, verbal protocols were broken down into proposition elements using formal linguistic methods. The protocols were used to evaluate expert and novice subjects' performance on a difficult clinical diagnostic task involving a written medical vignette. By categorizing certain linguistic elements from the protocols as "linking" propositions, the researchers were able to identify the step by step formal logic used by the subject to reason through the case material. "Linking" propositions logically connect facts and ideas (coded as formal propositions) used by subjects in their "think aloud" protocols. These "links" between other types of propositions in the protocol indicated the subject's use of either an inductive or deductive reasoning strategy. Inductive logic operationalized in this manner corresponds to forward reasoning because it involves a direct movement from presented data to an inferred task solution or diagnosis. Deductive logic, in contrast, corresponds to backward reasoning because it involves the testing of different

possible solutions or diagnoses through a systematic evaluation of presented data (symptoms). Through comparison of each protocol with an established "expert's expert" analysis of the case, the authors were able to assess the predominate type of reasoning employed by each subject. The study found that on the task of reconstructing the case from memory, experts' recall was significantly less detailed and explicit than the novices' recall. This result is explained by the authors as a function of the experts focusing attention only on case relevant material, thereby not processing other case data with the same level of attention. The novices, on the other hand, could not easily differentiate relevant from irrelevant information and therefore processed all of the information with more or less equal attention. This resulted in the novice subject's greater recall performance. Patel and Groen (1991) argue that the theme of enhanced recall is irrelevant to medical diagnostic expertise because what differentiates experts is their ability to identify relevant information, not their ability to store large amounts of unimportant data.

Patel and Groen's (1987) primary finding using their propositional analysis technique was that forward reasoning

is highly correlated with diagnostic accuracy. This result goes against the traditional wisdom of the "hypotheticodeductive" method as a model for how medical experts reason through clinical cases. The authors argue that diagnostic expertise is defined by acquired, schematic, domain-specific knowledge which allows experts to hone in on relevant information and to quickly represent cases in terms of probable diagnostic hypotheses. Even on the open-ended interactive task of diagnosing an actual patient, the expert will initially adopt a data-driven (forward) search to generate likely hypotheses, followed by some backward reasoning in order to account for unexplained data (Patel and Groen, 1991)¹¹. These authors also argue that the high correlation between the use of forward reasoning and diagnostic accuracy is evidence that expert diagnosis is

¹¹Not all researchers in the domain of medical diagnostic expertise agree with Patel and Groen's conclusions regarding the "hypotheticodeductive" method. Elstein et al (1978) present a study demonstrating the use by expert subjects' of the "hypotheticodeductive" method on medical case analysis. Elstein and Bordage (1988) explicitly question Patel and Groen's conclusions regarding the link between forward reasoning and diagnostic accuracy, and offer an alternate interpretation of their data. A thorough exploration of this dispute is beyond the scope of the present study, however.

more than simply a matter of the experts recognizing prelearned associative patterns. This point is predicated on the vast complexity of potential medical case presentations and on the authors' objection to the notion that diagnostic expertise can be reduced to simple production rules. The empirical finding that forward reasoning is correlated with accurate clinical diagnosis suggests that experts have established knowledge in the form of schemata which automatically "filter" out irrelevant material (Patel and Groen, 1991). This theoretical alternative to the "pattern matching" theory of expertise is discussed in the next section.

Theories of Expertise

De Groot's original study of chess expertise generated two significant empirical findings: 1) experts demonstrate enhanced recall for actual mid-game board positions and 2) experts typically noticed the "best next move" during their initial familiarization with the board. Although de Groot noted that experts tended to represent the board differently than (relative) novices (ie in terms of the "best move"),

his theoretical explanation was based primarily on experts' enhanced recall ability. It was the finding that experts are able to reconstruct mid-game board positions more accurately than novices and that experts' recall of information occurs in larger groupings, which served as the basis for the development of the first theory utilizing the "expertise" approach. Enhanced recall effects fit well within the predominate information processing paradigm of the time and follow up studies, such as that by Chase and Simon (1973), were devoted to exploring recall effects as conceptualized from an information processing model of cognition. It was Chase and Simon who delineated the first working theory of expertise.¹²

The Associationist View

The Chase and Simon theory of expertise seeks to explain how experts can seem to process information beyond the limits of human short term memory capacity as established by Miller (1956). In order to account for the

¹²Chase and Simon (1973) present their theory as a preliminary "first attempt" at explaining the data.

phenomenon that chess experts can reconstruct a mid-game board position comprised of dozens of pieces, virtually without error, Chase and Simon hypothesize that experts "chunk" information according to meaningful arrangements. The observation from Chase and Simon's (1973) study that subjects verbally recalled piece arrangements in groupings, the size of which differentiated experts from novices, supports the postulation that experts encode information in larger "chunks." Chase and Simon's perspective, consistent with standard information processing theory, is that cognition is "static." Information is processed in a veridical fashion and acted upon by internal mechanisms determined by the structure of the cognitive system. This model of mind is loosely based on an analogy to the well understood mechanisms by which computers process and store information. Thus, the Chase and Simon theory explains the empirical findings of the chess expertise literature in terms of information storage and the simple principle of association. Experts develop expertise as a function of years of experience during which they accumulate vast knowledge in the form of a storehouse of key patterns. Through experience and study, this vast reservoir of

information is stored in meaningful configurations which are associated in memory with similarly stored procedural knowledge of "best" solutions. When the expert is confronted with a mid-game board position, as in de Groot's study, key configurations within the position are recognized, or matched with memory patterns, and these serve as retrieval cues for eliciting associated "best" solutions. In this way general domain knowledge is linked to the specific procedural knowledge necessary for superior performance. Experts demonstrate enhanced recall over novices because the encoded patterns in their memory are larger and more complex. Experts notice the "best next move" from these recognized patterns because such concrete procedural knowledge is encoded in memory through associated links with the encoded patterns.

Although the associationist theory of Chase and Simon addresses the empirical findings of de Groot's original study, as well as follow up studies in the chess expertise domain, there are some shortcomings to this theoretical approach. One shortcoming is that the theory assumes that experts encode meaningful patterns as a result of their extensive experience within the domain. To encode specific

information sets in this fashion, along with the necessary associational retrieval links, would require a vast storage capacity along with a complex hierarchical central control system in order to maintain the information in a retrievable state. While this is theoretically possible, the inconsistency and apparent context dependence of memory phenomena, suggests that this is an overly cumbersome model for human mental functioning.

Another shortcoming of the associationist theory of expertise is that, while it predicts superior expert performance on well-defined tasks, such as solving physics problems, it does not predict differences between experts and novices on the performance of novel tasks¹³. This is problematic because common understanding of the term "expertise" suggests that experts should excel in particular at difficult and novel tasks. Experts select difficult tasks for themselves (Scardamalia and Bereiter, 1991) which "push

¹³One could argue that the findings from the behavioral decision literature, that experts fare poorly on clinical prediction tasks, is evidence that experts indeed are no more effective than novices on novel sorts of tasks. I feel, however, that what differentiates experts from novices in this sense is that experts ask better questions than novices, an ability which results from their operating at a more abstract cognitive level.

the limits" of the domain in new directions and expand the frontiers of ability and understanding. It is by this criteria that individuals are heralded as experts in society and for which their efforts are valued. Such aspects involving the creative application of knowledge and experience to wholly new contexts is poorly accounted for by a simple pattern matching theory of expertise.

Yet another, potentially more serious shortcoming of the associationist theory of expertise, is that it ubiquitously predicts that experts will demonstrate enhanced recall ability. This empirical finding, however, has been called into question (Ericsson and Smith, 1991) in domains where the task is less well defined than chess. Enhanced recall has not been a prominent finding in the domain of physics problem solving (Anzai 1991) and the medical diagnosis literature has demonstrated negative recall effects (Patel and Groen, 1987). As noted above, Patel and Groen (1991) argue that, for the domain of medical diagnosis, recall effects are irrelevant to expert performance. These cross-domain findings suggest that the nature of the task may have a large influence on how experts function within their domain (Hamm, 1988). Findings such as

enhanced recall effects may be more domain specific than previously thought.

The Representational Model

In de Groot's original study of chess expertise, the finding which strikingly differentiated the grand masters from the players of lesser skill was that the grand masters noticed the "best next move" during their initial familiarization with the board. Both de Groot and subsequent researchers attributed this finding to encoded patterns associated with concrete procedural knowledge. An alternative explanation, however, is that the expert when faced with a chess problem, literally "sees" the chess board differently than does the novice. This idea stems from the notion that domain knowledge is learned by the expert in a "dynamic" fashion involving a gradual cognitive shift over time from reasoning based on over learned concrete procedures to reasoning based on abstract concepts¹⁴. For

¹⁴This idea bridges the expertise literature with research on concept formation. An interesting parallel to the findings from the physics expertise literature in this respect is evident in studies of concept formation in

experts, abstract domain specific knowledge is used to generate cognitive task representations which, as a function

children conducted by Kiel (1987). In these studies both younger and older children were given word stimuli and asked to provide verbal definitions. The younger children tended to focus on word typicality, or the typical features associated with words. For example, words like "uncle" were defined by descriptors like "man who brings me presents" or "daddy's pal." The older children, however, had a different understanding of these same words. They tended to focus on defining characteristics of words, such as, "uncle is daddy's or mommy's brother." This developmental shift from typical characteristics to defining features of concepts is a transition from concrete to abstract linguistic representations. This description is similar to the novice-expert differences reported by Chi et al (1981) in which subjects differentially categorized physics problems by either surface features or underlying principles of physics. This finding suggests that in learning concepts individuals initially focus on typical concrete characteristics which are used to form cognitive "prototypes," or representations of good typical examples based on features correlated in experience. As additional knowledge is gained, however, the features which are selected as salient for defining categories shift to a more functional basis. Features which were previously seen as correlated with the concept (uncle brings me presents) are re-evaluated in terms of their logical necessity (grandmother also brings me presents, but she's not uncle). In other words, as a more sophisticated knowledge base is developed, cognitive representations are formed based upon abstract understandings. These are used to highlight relevant features and to de-emphasize irrelevant features (lots of people bring me presents, but only daddy's brother is my uncle). Thus, the defining features of the concept become established as personal theories of how the world operates are developed. This model for how concepts are learned and mastered by children applies equally well to the development of expertise.

of their schematic organization, automatically incorporate a means-end analysis of a solution. Experts use schematized abstract theories¹⁵ to frame their cognitive apprehension of the task such that ready solutions become immediately apparent. In this sense, experts literally apprehend a different chess board than novices because their immediate means of conceptualizing, or making sense of, the pattern arrangements is different.

On an empirical level, a representational theory as described above does not predict enhanced recall effects as a central research finding. Such a theory does, however, predict superior expert performance on novel tasks. On a theoretical level, a representational theory does away with a "two stage" view of cognition, positing instead that individuals apprehend tasks directly. Rather than having greater memory stores or enhanced processing ability, experts have implicit functional theories about how things work, knowledge which they use to make better sense of what they observe. The problem space offers the expert more

¹⁵Lakeoff (1987) in discussing concept formation refers to schematic theories of this type as idealized cognitive models.

opportunity than is afforded the novice because the expert has acquired the cognitive machinery to relate directly with this environment. The novice, on the other hand, must plod through indirect rote procedures¹⁶ in a search for possible connections to the desired goal. This formulation suggests an analogy between the expert's direct cognitive apprehension of the problem space, and a general model of direct perception.¹⁷

¹⁶Dreyfus and Dreyfus (1986) propose a theory of expertise which involves a five stage process of development along a dimension from analytic reasoning to intuition. Having no conceptual knowledge upon which to draw, the stage one novice uses purely rule bound algorithmic procedures to solve problems and to handle decisions. With experience, these procedures become more automated until, at the level of the expert, pure intuition is used to assess and solve problems, as well as to make all decisions. While this theory is overgeneralized, it does capture the notion that expertise involves a development from concrete procedural operations to a use of abstract cognitive models.

Another related theory proposed by Hammond (1980), known as cognitive continuum theory, also makes use of a theoretical dimension between analytic reasoning and intuition. This theory posits an optimum balance between algorithmic procedures and intuition depending upon the specific nature of the task. Expertise is thus conceptualized in terms of the ability to "read" this optimum balance and to apply the appropriate cognitive strategy.

¹⁷The ecological perspective on concept formation is highly relevant to this argument. The "classical" cognitive psychology

view of categorization is that category membership is defined simply by necessary and sufficient features. This view was displaced by the work of Rosch (1973) who developed a research paradigm demonstrating that abstracted "typical examples," or "prototypes" are formed within the cognitive apparatus to represent categories. Category membership is assessed through degree of similarity to these prototypes. Rosch showed that a "graded structure" can be empirically established to operationalize the goodness of fit between exemplars of a category and the idealized prototype. More recent work, however, has shown that these empirical "prototype effects" are not particularly stable (Barsalou, 1987). Medin and Wattenmaker (1987) argue that the notions of similarity to prototype, or correlated features, are insufficient to account for category cohesion without the ongoing operation of some kind of personal cognitive theory. Given the great variability of empirical graded structure ratings (prototype effects) and the problems of explaining category coherence in terms of prototypes, Lakeoff (1987) argues for the conceptualization of "idealized cognitive models." These are understood as functional structures resulting from the interaction of people's knowledge and expectations with environmental objects (ie objects and their ecological meaning are a function of environmental context, not entities existing independent of human experience). Kiel (1987) presents developmental research in support of the idea that individuals initially form concepts based on prototypes, but gradually replace these with idealized cognitive models as greater functional understanding is acquired. Neisser (1987) points out that categorization is subjectively an immediate process and that the spontaneous use of concepts is similar to the model of direct perception proposed by J.J. Gibson. This is to say that a two stage model of cognition, like a two stage model of perception, is redundant if we assess the actual information used to form concepts. Neisser, like Lakeoff (1987), argues that categories are similar to perceptual affordances (opportunities afforded the organism by the environment) in that category membership is defined by the relation between the object and the particular system of thought used to apprehend the object (Neisser 1987).

Two Empirical Themes Revisited

The associationist theory of expertise was built primarily upon the finding that experts demonstrate enhanced recall effects in the domains chess and physics problem solving. The representational model, in contrast, is based on the idea that forward reasoning differentiates experts and novices across the domains of chess, physics problem solving, and medical diagnostic expertise. It appears from more recent developments in the expertise literature that recall effects are a phenomena restricted to certain closed-ended tasks, and are not as central to a general understanding of expertise as was once thought. Forward reasoning, however, is a finding which appears to be a fairly consistent characteristic of expertise across domains. From the perspective of representational theory, forward reasoning may be understood in terms of the presence of working theoretical representations of domain specific material within the expert's cognitive frame. With an

Expertise in this context may be understood as a domain specific expert categorization ability.

understanding of relevant abstract principles, the expert is able to generate highly specific task representations, regardless of whether the problem type is familiar in a concrete sense. These representations, or working theories, are used to integrate information in a forward direction towards a problem solution. Without an understanding of the relevant abstract principles involved, a working theoretical representation cannot be generated and the subject must focus on concrete problem features, hoping through a search of possible "right answers" to hit upon a hypothesis which causally accounts for the data. While the notion of working theoretical models, and the subsequent use of forward reasoning, clearly applies to the study of expertise, this pattern of moving from "seeing" only concrete features to understanding abstract functional principles, is plausible for any task domain which requires time, training and experience to master.

CHAPTER 2

Methods and Procedures

Overview

This study utilized a semantically based protocol analysis technique for examining cognitive differences between experienced and inexperienced psychology clinicians. The method involved the collection of "think aloud" protocols from semi-structured interviews between subject and examiner. Subjects were shown written vignettes of two cases segmented into eight categories of historical and clinical information. After reviewing each vignette segment, subjects were interviewed by the examiner. The procedure was designed so that the order of vignette segment presentation was systematically varied depending upon subject responses during the interview. Following the "think aloud" interview associated with the presentation of all the vignettes for a particular case, subjects were asked to formulate their diagnosis for the case. In the final phase of the interview, subjects were asked to recall all of the facts they could remember from the case material. Audio

recording equipment was used to capture the interview text verbatim. The tape for each interview was later transcribed into an exact record of the subject-examiner interaction. This procedure resulted in written protocols with the text divided into segments corresponding to the eight case vignette segments presented to each subject, with two additional segments corresponding to the diagnostic formulation and recall phases of the interview. Specific phrases and key words were highlighted from the written protocols and coded according to the scoring system presented in appendix one.

Subjects

Subjects for this study were 10 inexperienced and 9 experienced psychology clinicians. All inexperienced clinicians were 2nd year graduate students in the University of Tennessee Ph.D. training program in clinical psychology. These subjects had no previous professional experience as psychotherapists prior to the beginning of their 2nd year of training at the University of Tennessee. The graduate students had all begun their clinical training in their 2nd

year of the program. This training consisted of evaluating and treating (under supervision) patients from the community at an in-house psychology clinic. All 10 inexperienced subjects therefore had at least one year of academic training, but less than one year of professional clinical experience. The inexperienced subjects were all classmates and so had similar academic experiences up to the point of participation in this study. Participation in the study was purely voluntary with no incentives offered. The age of the inexperienced subjects ranged from mid-twenties to early forties. Seven of the inexperienced subjects were male, and three were female.

Experienced subjects for this study were all clinical psychologists working at least part time as private practitioners in the community. Most were also involved in the University of Tennessee clinical training program as faculty members or supervisors. All experienced subjects had worked in clinical settings for a minimum of three years post doctorate. There was, however, considerable variability in years of experience beyond this minimum cutoff. The training of the experienced subjects also varied somewhat depending on background, but all had

familiarity with various theoretical perspectives, including psychoanalysis. Participation in the study was purely voluntary and no incentives were offered. The age range of experienced subjects was from mid-thirties to late fifties. All of the experienced subjects were male.

Materials

The materials used for this study consisted of two clinical case vignettes divided into eight segments of historical and clinical information. The actual case vignettes are included in appendix two. A tape recorder and cassette tapes were also used in data collection.

The two cases used in this study were developed from "root" cases taken from the DSM-III-R Casebook (Spitzer et al 1989). These root cases were then elaborated by the author through the addition of fictional biographical data, symptoms, and contextual information for various symptoms present in the original case. The case elaborations were intentionally developed without specific diagnostic categories in mind. The focus in writing the cases was placed instead on creating a plausible story and character

background. Clinical aspects of the case were derived both from the original material and from the examiner's effort to present a complex, believable, and interesting case description. Thus, the written cases provided a story context from which various aspects could be interpreted as symptoms, or as having diagnostic relevance, depending upon what the subject had in mind. Obviously, since the cases depart from the natural phenomenology of the original clinical material and include narrative material from the author, this eliminated the possibility of assessing any particular diagnosis as "correct."¹ The cases were constructed in a manner so as to purposely not provide enough information by which subjects could generate a definitive category based diagnosis using the DSM system. Vignettes shown to the subjects were intended to encourage speculation about the patient depicted and about what additional information might be sought in order to clarify a subject's diagnostic thinking. Since several potentially

¹As discussed in the previous section, the task of selecting criteria for "correctness" for open ended tasks is problematic. Maintaining the natural phenomenology of the case was therefore sacrificed in favor of having experimenter control over the type of clinical picture presented to subjects.

plausible diagnostic lines of reasoning were at least partially supported by the limited cases material, the subject was free to "play" with various differential options and to exercise his or her expertise in formulating the "best" possible case formulation. Both cases used in the study were reviewed prior to data collection by two independent clinicians in order to assess plausibility and face validity of the symptom clusters. Revisions were made by the author based on the feedback from these individuals. It should be noted that while the cases were considered fictional, the subjects were not informed of this until after the interview data had been collected. The majority of subjects engaged the case material as though it portrayed an actual patient and reported afterwards that the cases were believable. In the few instances where discrepancies were noted (both correctly and incorrectly) subjects generally interpreted these within the framework of the case material, using the discrepancy to generate additional questions and hypotheses about the case rather than to question the plausibility of the presentation.

Procedure

The procedure employed in this study consisted of the data collection procedure and the protocol analysis. The data collection procedure involved a simple semi-structured interview between the examiner and subject. Subjects were asked to set aside either a single two hour time block, or two one hour time blocks. All of the interviews were completed within one of these two time frames. The interviews were scheduled by appointment and took place either in the examiner's or the subject's office. Subjects were told that there were two cases to be reviewed and discussed in the interview, and that each case would take about an hour to complete. The instructions read to the subjects at the beginning of the interview, and repeated in the second interview for subjects who completed the procedure in two sessions, were:

"I am studying differences between experts and novices on a clinical diagnosis task. The study is focused on process variables and is not a comparison of performance or diagnostic accuracy."

"I have here two cases which I am going to show you and ask you about. Each case is divided into segments of information which roughly correspond to standard sections in a case file. The information provided is more limited than in an actual case file, however, and some of the information is presented biographically rather than clinically. You may assume that this information is from a reliable source. I will show you each segment one at a time and then take away that segment when you have finished reviewing the material. Your task is to read each segment and then to 'think aloud' about the case. I will ask you after each segment, first, what stands out from the case after reading the segment, and second, what is the first question which occurs to you with regard to additional information you would look to next if this were an actual case file. Your purpose in examining each case is to diagnose the case in any manner you choose and to formulate the case material based on your diagnosis. This study is not concerned with a measurement of diagnostic accuracy and I am much more interested in your case formulation than with your

selection of diagnostic categories. After each case I will ask you to recall information about the case."

After hearing these instructions subjects were asked if they had questions. Some subjects asked for clarification in terms of what was meant by diagnosis and formulation. Subjects were simply told that they could diagnose and formulate the case in any manner they chose. Following the instructions all subjects were given the first case segment entitled "Demographics and Presenting Problem" (see appendix two). Some subjects began to speak before reading the entire segment and were told that they would be asked to repeat what they had said when the tape recorder was operating. This effectively discouraged subjects from "thinking aloud" before reading the entire segment. After the subject indicated that he or she had completed reviewing the segment the examiner reclaimed the segment and the tape machine was started. The examiner then asked the subject "what stands about the case after reading that segment?" The subject was then allowed to "think aloud" about the case material with minimal prompts from the examiner. Prompts were given only when the subject did not seem to understand

the task or when a statement was patently unclear and the examiner did not understand to what the subject was referring. Prompts, when given, were as nondirective as possible, with the subject being encouraged to spontaneously generate a "think aloud" narrative.

When the subject indicated that he or she had mentioned all that "stood out" about the case, the examiner asked of the subject "what additional information would you look to next?" Subjects were then allowed to "think aloud" about questions which occurred to them about the case and the manner in which they would like to proceed. The examiner made a judgement at this point as to whether any of the remaining segments addressed the subject's questions. If the subject's questions were in any way addressed by a remaining segment, that segment was selected by the examiner as next in line for the administration procedure. If more than one question was asked by the subject addressing multiple remaining segments, the first question addressing a segment was used to indicate the next segment. If none of the subject's questions addressed a remaining segment, then the next segment in the standard order was selected for the next administration. The standard order administration list

(see appendix two) was used by the examiner as a loose guideline for administration order which could subsequently be modified by the subject's questions. The purpose of this technique was to allow the subject as much flexibility as possible in generating a narrative of his or her diagnostic reasoning spanning from segment to segment. This also gave the subject some perceived influence over the course of the interview and provided an interpersonal setting more similar to a clinical case seminars or evaluation. Despite the variation in order of presentation, however, all subjects received the same eight vignette segments and were not prompted to formulate their diagnostic impressions before exposure to all of the case material available.

After the subject completed the "think aloud" procedure for each segment described above, the tape machine was turned off and the next selected segment was administered. The administration procedure for all of the remaining segments followed this sequence. Most subjects quickly learned the task sequence and began to anticipate the examiner's prompting questions. In these instances, since the subject clearly understood the task, the standard examiner prompts were omitted in the later segments.

(see appendix two) was used by the examiner as a loose guideline for administration order which could subsequently be modified by the subject's questions. The purpose of this technique was to allow the subject as much flexibility as possible in generating a narrative of his or her diagnostic reasoning spanning from segment to segment. This also gave the subject some perceived influence over the course of the interview and provided an interpersonal setting more similar to a clinical case seminars or evaluation. Despite the variation in order of presentation, however, all subjects received the same eight vignette segments and were not prompted to formulate their diagnostic impressions before exposure to all of the case material available.

After the subject completed the "think aloud" procedure for each segment described above, the tape machine was turned off and the next selected segment was administered. The administration procedure for all of the remaining segments followed this sequence. Most subjects quickly learned the task sequence and began to anticipate the examiner's prompting questions. In these instances, since the subject clearly understood the task, the standard examiner prompts were omitted in the later segments.

After the subject indicated that he or she had completed thinking aloud about the last administered segment, the examiner asked:

"You've mentioned a number of different things and may have already given a diagnosis, but now I would like you to put it all together. What diagnosis would you settle on for this case and how would you formulate that diagnosis?"

As in the administration of the case vignette segments, the examiner gave minimal prompts and encouraged the subject to generate a spontaneous narrative without interruption.

Prompts were given only for the purpose of gross clarification. The one exception to this was the examiner prompt, given to subjects listing multiple diagnoses, as to which of the diagnosis being discussed did he or she consider the primary diagnoses.

After the subject indicated that he or she had completed his or her thoughts about diagnosis and formulation, the examiner asked:

"I want you now to recall into the microphone all of the facts you can remember about this case"

The subject was then allowed to speak freely recalling information or adding ideas about the case without interruption from the examiner. When the subject indicated that he or she was finished, the examiner asked the subject to select how confident he or she felt about his or her primary diagnosis from among seven options.

Very confident

Confident

Fairly Confident

Not sure

Fairly uncertain

Uncertain

Very uncertain

This completed the interview and the data collection phase of the procedure. The administration procedure for the second case was identical to that for case one.

Protocol Analysis

Cassette tapes of the interview data were transcribed into typed protocol records through use of a dictaphone machine. Codeable key words and phrases from the transcribed protocols were first highlighted and then coded according to the coding system included in appendix one. The key words and phrases selected for coding were based upon the examiner's judgement of the formal elements of the subject's communication of ideas in the protocol. These judgements were made much as one would use short hand to select key elements of a verbal communication for a summary record. Extraneous, non-case specific, and purely conversational communications from the raw protocols were excluded from consideration in the scoring. The results of this selection procedure were a series of highlighted words, clauses, and sentences containing the examiner's judgement of the ideational building blocks of the subject's communication. This procedure of extracting ideational elements from the raw syntax of the subject's discourse inevitably caused disruption to the logical point by point flow of the subject's reasoning. However, since the coding

system was developed to functionally categorize only the types of statements used by subjects to communicate ideas, rather than the logical content of the ideas themselves, the disruption did not effect the outcome of the quantitative analysis.

The categorization method used for coding protocols was based upon an assessment of the protocol elements theoretical degree of abstraction from the original case material. These assessments involved the examiner's judgement regarding the manner in which the subject made use of the highlighted element as part of his or her communication. While qualitative aspects of the highlighted elements could not be codified, it was necessary for the examiner to make use of the context of the subject's communication for scoring. This included making inferences regarding the subject's intentions based on a comprehensive reading of the entire protocol segment. The purpose of the coding system was to delineate the formal quality of the subject's statements used to communicate ideas while excluding the actual content and logical connection of these statements. The coding system provided a simple tally of categorized statements and key words used by the subject in

"thinking aloud" about the case diagnosis. The validity or level of sophistication of these ideas could not be addressed by this quantitative scoring method. In actual practice, the examiner's knowledge of the coding system categories provided a guideline for highlighting codeable elements from the raw protocol data. Thus, while highlighting and coding were conducted sequentially as a two step protocol analysis process, they were not independent steps. Since highlighting and coding were conducted by the same examiner, this aspect of the process was unavoidable. See appendix one for a complete description of the coding system.

In order to provide a more qualitative analysis of the protocols, one category of the coding system entitled "diagnostic hypotheses" was selected for additional consideration. Each statement coded as diagnostic hypothesis was listed across the protocol segments. The diagnostic hypotheses generated by each subject are listed in appendix 3.

CHAPTER 3

Rational for Coding System

The theoretical basis for the quantitative approach to protocol analysis used in this study is drawn from the classic "expertise approach" theory outlined in the previous section. According to the associationist theory of expertise, the expert is capable of superior task performance due to the acquisition over many years of an accumulated domain specific knowledge base by which problem instances are recognized and "matched" according to prelearned patterns. The representational theory of expertise views the expert's advantage over the novice on domain specific tasks in terms of the availability of a schematic knowledge base which allows for more efficient "guided" search of the specific information necessary for task completion. These views have been operationalized in the language of problem solving research in terms of enhanced recall effects and forward verses backward reasoning strategies.

Both the associationist and representational theories of expertise share a conceptualization of the expert

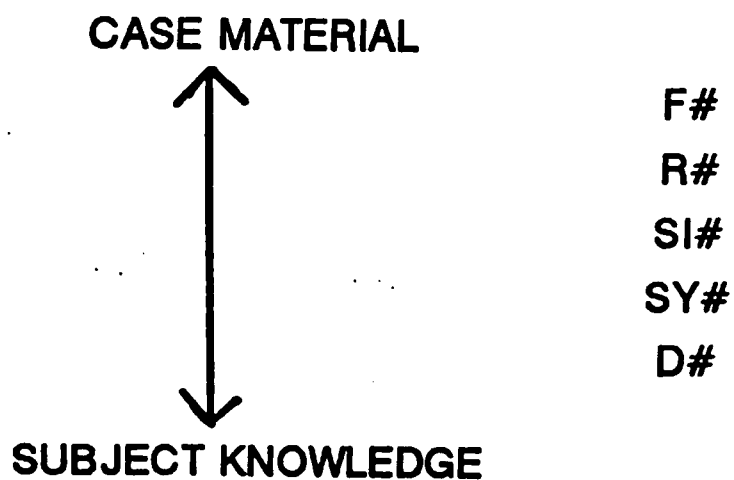
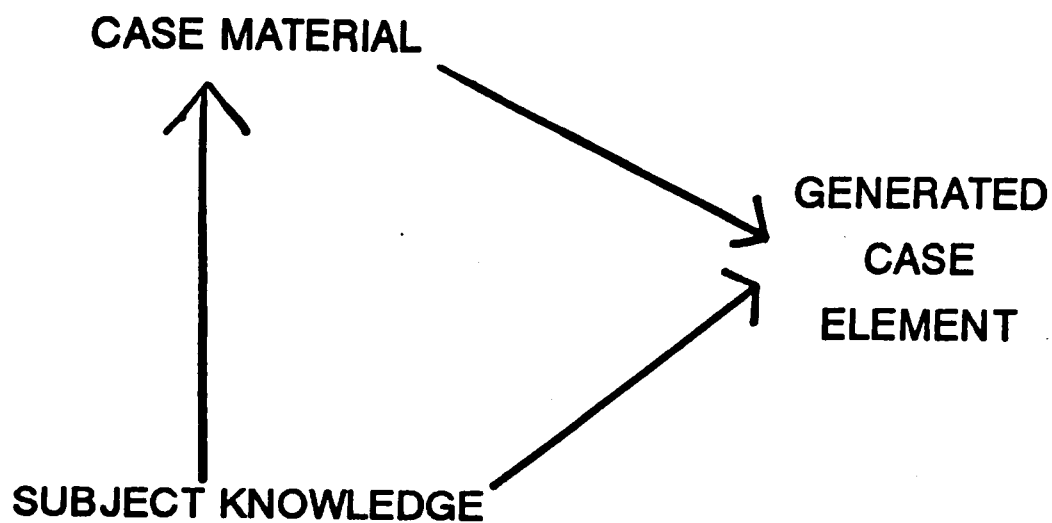


Figure 1
Abstraction of Coding Catagories

engaging domain specific tasks with the advantage of a storehouse of knowledge. This knowledge base is organized so as to allow optimum use of information for completing the task. Experts operate on a cognitive level characterized by reasoning which is primarily knowledge driven. Non-experts operate at a cognitive level where reasoning is driven by the concrete problem features. The associationist and representational theories of expertise share a basic assumption: that task performance is a function of an interaction between 1) the individual's acquired knowledge base and 2) the specific context of the task at hand. Pure knowledge driven (forward) reasoning relies almost exclusively on the subject's knowledge base to direct information search in the narrowest and most optimum fashion. Pure backward reasoning, on the other hand, is akin to an almost random search for pragmatic combinations linking task completion with given information. Despite the apparent opposition of these strategies, however, most complex multi step tasks do not lend themselves to pure forward or pure backward reasoning strategies. The subject's ability to flexibly combine both forms of reasoning to match tasks demands, allows experts to solve

novel or unexpected problems.¹ In this sense, forward and backward reasoning may be understood as a continuum between predominantly knowledge driven information search and problem solving, and predominantly context driven information search and problem solving.

The protocol analysis approach developed for this study assumes that subject performance on any particular task is a function of the individual's abstract knowledge base in relation to the specific context of the case material. This relation is understood as based along a continuum between knowledge driven and context driven reasoning. It was hypothesized that subjects with more experience in the domain of clinical diagnosis will exhibit reasoning driven by clinical knowledge to a greater extent than the reasoning of subjects with less experience. In contrast, the reasoning of subjects with less clinical diagnostic experience will be driven by case features to a greater extent than the reasoning of subjects with more experience.

Figure 1 illustrates the categories used in the coding system in terms of relative degree of abstraction from the

¹This is similar to the perspective presented by Hammond (1980). See footnote #16 in the introduction.

concrete case material. The theoretical assumption underlying this illustration is that the subject's "spoken thoughts" about the case represent direct information regarding cognitive processes.² This assumption was built into the methodology of showing subjects relatively brief vignettes of information interspersed with "think aloud" interview periods. During the procedure subjects were engaged in either reading information or "thinking aloud," with little opportunity to "think silently." It is reasonable to assume that the "thought aloud" protocol information represents an unedited picture of the subject's cognitive process while formulating a diagnosis for the case.

The coding system in this study is based on the theoretical idea that the relative balance between a subject's abstract knowledge-based reasoning, and his or her concrete feature-based reasoning, can be assessed through analysis of the subject's statement usage in the "think aloud" protocols. Statements which are reported directly

²This view on the validity of protocol data as a direct result of cognitive processes is championed by Ericsson and Simon in their book 'Protocol Analysis' (1984). Their view is well documented and generally accepted in the literature.

from the original case material correspond to a balance of predominantly feature-based reasoning. These statements reflect the subject's use of a more concrete search driven reasoning strategy. In contrast, statements which are task focused diagnostic ideas about the case material correspond to a balance of predominantly knowledge-based reasoning. These statements reflect the subject's use of a more abstract reasoning strategy guided by the subject's acquired clinical knowledge. Symptom hypotheses, speculative ideas, and reference statements each represent a different relative degree of abstraction from the case material. These statements each reflect a varying balance of knowledge-driven verses search driven reasoning along a theoretical continuum (see figure one). Speculative idea statements, for example, are considered more dependent upon abstract knowledge based reasoning than reference statements because the former type of statement involves the subject moving beyond the scope of the case material in an attempt to explain or theorize. Symptom hypotheses represent a further degree of dependence upon abstract knowledge based reasoning than speculative idea statements because they are more specifically focused to the task of generating a diagnosis.

The coding system is designed to indirectly assess reasoning strategies by categorizing statement types and examining frequencies across protocol segments. While this procedure cannot assess differences in the depth, richness, and sophistication of subjects' reasoning, it has the advantage of being a quantifiable method of examining cognitive patterns. Since the actual coded data consists of more or less "natural" speech elements listed in order of generation by the subject, the content of the subject's line of reasoning is recoverable within the coding system.

Variables

The variables generated from the protocol analysis for quantitative comparison are listed in table 1.

Hypothesized predictions

- 1) The experienced group will generate significantly greater numbers of diagnostic hypotheses for variables 17 (D#SUBTOT), 18 (D#FORM), 19 (D#TOT), and 20 (D COMBINED).
- 2) The novice group will generate significantly greater

Table 1

Variables

- 1) F#SUBTOT: The total number of features for case (#), excluding formulation and recall segments.
- 2) F#RECALL: The total number of features for case (#) coded from the recall segment.
- 3) F#TOT: The total number of features for case (#).
- 4) F COMBINED: The total number of features for both cases combined.
- 5) R#SUBTOT: The total number of references for case (#), excluding formulation and recall segments.
- 6) R#RECALL: The total number of references for case (#) coded from the recall segment.
- 7) R#TOT: The total number of references for case (#).
- 8) R COMBINED: The total number of references for both cases combined.
- 9) SI#SUBTOT: The total number of speculative ideas for case (#), excluding formulation and recall segments.
- 10) SI#FORM: The total number of speculative ideas for case (#) coded from the formulation segment.
- 11) SI#TOT: The total number of speculative ideas for case (#).
- 12) SI COMBINED: The total number of speculative ideas for both cases combined.
- 13) SY#SUBTOT: The total number of symptom hypotheses for case (#), excluding formulation and recall segments.
- 14) SY#FORM: The total number of symptom hypotheses for case (#) coded from the formulation segment.
- 15) SY#TOT: The total number of symptom hypotheses for case(#).
- 16) SY COMBINED: The total number of symptom hypotheses for both cases combined.
- 17) D#SUBTOT: The total number of diagnostic hypotheses for case (#), excluding those coded from the formulation and recall segments.
- 18) D#FORM: The total number of diagnostic hypotheses for case (#) coded from the formulation segment.
- 19) D#TOT: The total number of diagnostic hypotheses for case (#), including all segments.
- 20) D COMBINED: The total number of diagnostic hypotheses for both cases combined.

numbers of features for variables 2 (F#RECALL), 3 (F#TOT), and 4 (F COMBINED).

3) The experienced group will generate significantly greater numbers of symptom hypotheses for variables 13 (SY#SUBTOT), 14 (SY#FORM), 15 (SY#TOT), and 16 (SY COMBINED).

4) Of the total of variable 3 (F#TOT) added to variable 7 (R#TOT), the experienced group will generate a significantly greater proportion of reference codings. This is also predicted for the total of variable 4 (F COMBINED) added to variable 8 (R COMBINED).

The rational for these predictions is based on findings from previous research, the representational theory of expertise presented above, and the theoretical assumptions outlined in the figure one. These predictions are relatively straight forward.

Prediction #1 is based on the idea that experts reason primarily in a forward direction, using their superior acquired knowledge base to navigate through the case information. As a result, more of the statements generated in the protocols will be on the abstract, knowledge driven end of the continuum. Novice subjects will not generate as many abstract, knowledge driven statements because they are

limited by their dependence on backward reasoning, or a weak problem solving approach.

Prediction #2 is based on Patel and Groen's (1987) finding that novice medical diagnosticians demonstrated greater recall for case information than did expert subjects. Patel and Groen's explanation for this in terms of expert's greater selectivity of attention also applies to this study.

The rational for prediction #3 is identical to that of prediction #1.

The rational for prediction #4 is similar to that of prediction #1. An additional assumption is that references are essentially equivalent to features as case elements, but are reported by the subject with a greater degree of cognitive abstraction. Therefore, the proportion of references to the combined total of case elements described (F+R) in the protocol should be greater for experienced subjects. As in prediction #1, this finding will support the conclusion that experts utilize a more abstract reasoning strategy than do novices.

CHAPTER 4

Results

The variables generated from the protocol analysis were analyzed using T-test statistics to test significant differences between means. The T-tests for study predictions described above were computed one-tailed. Additional exploratory comparison were conducted for variables 11 (SI#TOT), and 12 (SI COMBINED). Since no predictions were posited for these variables, these analyses were computed using two-tailed T-tests. These results are presented in Tables 2, 3, and 4. Interrater reliability of the protocol scoring system was assessed using Cohen's kappa ($k=0.76$). The reliability of each individual variable was examined by computing the percentage of agreement between the two coders: F = 87%, R = 74%, SI = 78%, SY = 67%, D = 85%.

Table 2

Case 1:

Variable	Exp. Mean/SD	Novice Mean/SD	t	DF	p
F RECALL	34.22 14.84	50.50 25.19	-1.736	14.8	0.0515
F TOTAL	53.44 23.34	72.70 35.52	-1.409	15.7	0.089
R TOT/R+F	0.49 0.12	0.40 0.11	1.704	16	0.054
SI TOTAL	86.11 36.65	65.40 41.13	1.161	17	0.262 [^]
SY SUBTOT	24.00 12.44	21.50 5.86	0.551	11.1	0.297
SY FORM	9.33 2.83	8.30 2.89	0.211	16.8	0.418
SY TOTAL	55.33 16.03	35.20 8.82	3.34	12.1	0.003*
D SUBTOT	5.44 5.03	5.10 4.23	0.161	15.8	0.437
D FORM	5.89 2.62	5.30 2.58	0.492	16.7	0.315
D TOTAL	45.00 45.55	10.90 4.91	2.234	8.2	0.0275*

[^]= Two Tailed T-test

*= Significant @=0.05

Table 3

Case 2:

Variable	Exp. Mean/SD	Novice Mean/SD	t	DF	p
F RECALL	52.00 40.66	62.67 26.77	-0.657	13.8	0.261
F TOTAL	76.11 63.30	90.00 33.79	-0.581	12.2	0.286
R TOT/R+F	0.50 0.17	0.37 0.07	2.029	10.8	0.034*
SI TOTAL	82.56 29.51	54.11 32.02	1.96	15.9	0.068^
SY SUBTOT	34.67 9.49	19.40 5.68	4.198	12.8	0.0005*
SY FORM	5.33 2.29	5.20 2.28	0.093	12.3	0.928
SY TOTAL	39.22 16.78	31.78 7.24	1.222	10.9	0.123
D SUBTOT	20.33 7.98	2.40 1.58	6.623	8.6	0.000*
D FORM	5.33 3.04	3.70 2.06	1.356	13.9	0.197
D TOTAL	11.67 8.54	6.78 3.63	1.58	10.8	0.0715

^= Two tailed T-test

*= Significant @=0.05

Table 4

Cases Combined:

Variable	Exp. Mean/SD	Novice Mean/SD	t	DF	p
F COMB	129.56 83.15	164.89 69.00	-0.981	15.5	0.171
RC/FC+RC	0.49 0.13	0.38 0.08	2.179	13	0.024*
SI COMB	168.67 59.44	123.11 72.03	1.463	15.4	0.163 [^]
SY COMB	94.56 28.36	65.67 11.19	2.842	10.4	0.0085*
D COMB	56.67 44.10	17.89 7.42	2.602	8.5	0.015*

[^]= Two Tailed T-test

*= Significant @=0.05

Prediction #1, that the experienced group will generate significantly greater numbers of diagnostic hypotheses, is partially supported by these results. For case one, D-TOTAL was significant at $\alpha=0.05$, while D-SUBTOT and D-FORM did not approach significance. For case two, D-TOTAL approached significance at $t=0.07$, and D-SUBTOT was significant at $\alpha=0.05$. D-FORM did not approach significance for either case. D-COMBINED was significant at $\alpha=0.05$. Since the N for this study was relatively low, the failure to reach significance for case one D-SUBTOT and case two D-TOTAL, was perhaps due to lack of statistical power. There does not appear to be a measurable effect for D-FORM in either case.

Prediction #2, that the novice group will generate significantly greater numbers of features, is not supported by these results. Although novices did generate greater numbers of features, these differences were not statistically significant. F-RECALL for case one approached significance at $t=0.0515$. If an effect is present, as Patel and Groen (1987) found in their study, the current study did not have the statistical power to differentiate this effect between groups.

Prediction #3, that the experienced group will generate

significantly greater numbers of symptom hypotheses, is partially supported by these results. For case one, SY-TOTAL was significant at $\alpha=0.05$, while SY-SUBTOT and SY-FORM did not approach significance. For case two, SY-TOTAL was not significant, but SY-SUBTOT was significant at $\alpha=0.05$. SY-FORM was not significant for either case. SY-COMBINED was significant at $\alpha=0.05$. As with prediction #1, the small N size of this study limited the statistical power of these results and may have masked the effect.

Prediction #4, that the experienced group will generate a significantly greater proportion of references from the total of case elements is mostly supported by these results. For case one, R TOT/R+F approached significance at $t=0.054$. Case two R TOT/R+F was significant at $\alpha=0.05$. R TOT/R+F COMBINED was also significant at $\alpha=0.05$.

CHAPTER 5

Discussion

As was discussed in chapter 2, the subjects used within the expert group for this study were defined as clinical psychologists with a minimum of three years postdoctoral experience. This cutoff was chosen for pragmatic reasons. It is clear that the experienced clinicians utilized for this study are not necessarily equivalent to clinical experts defined through peer nomination, prestigious awards, and widespread recognition. However, within the domain of expertise research it is common practice to utilize experience as an operational criterion for the definition of expertise. For the purposes of maintaining consistency in hypothesis generation and comparison with other relevant studies, this convention is maintained in the current study. It should be understood, however, that expertise is a rich and varied concept with many lexical nuances and theoretical implications beyond the simple description of subjects as having met a minimum cutoff for level of experience. Interpretation of study results should therefore be tempered with recognition of the limitations imposed by the current

operational definition of expertise.

The principle finding of this study is that experienced and novice clinicians can be differentiated from each other based upon predictions derived from expertise theory. Specifically, the experienced clinicians make use of their superior knowledge base to remain more task focused. This is indicated by experienced group producing significantly greater numbers of both symptom hypotheses and diagnostic hypotheses throughout the protocols.

One note of interest in these results is that while there were significant effects for case one SY TOTAL and D TOTAL, and for case two SY SUBTOT and D SUBTOT, SY FORM and D FORM were far from significance on both cases. While this may be due in part to the fewer number of coded statements in the formulation segment, it also suggests that experienced clinicians generate more abstract case representations across the entire "think aloud" protocols. However, when subjects are asked explicitly to formulate a case diagnosis, both experienced and novice clinicians utilize symptom and diagnostic language. However, during the process of "thinking aloud" about the case, experienced clinicians remain more task focused in terms of using

symptom and diagnostic language more frequently than novices.

It is unclear why differences were found between the experienced and novice groups in the amounts of D and SY SUBTOT for case two, while for case one differences were found for D and SY TOTAL. Although this may be a statistical artifact, it seems to suggest that experienced clinicians generated more SY and D statements in the RECALL segment of case one, but not case two. In fact, since D and SY TOTAL for case two were not significant, it appears that novices may have generated more D and SY statements in case two RECALL. It is likely that if this difference represents an actual effect, then it is the particular character of the cases (see appendix two) which accounts for this anomaly.

The lack of support for prediction two in this study is probably a function of the deliberate vagueness of the cases. In Patel and Groen's (1987) study, there was a single "correct" diagnosis for the case which had been predetermined by the "expert's expert." The case used by Patel and Groen(1987) was written with clear and non-contradictory sign and symptom information such that experts could key upon only relevant information. The present

study, in contrast, did not utilize this methodology for case development. Cases had no predetermined "correct" responses and there were many potentially plausible diagnoses applicable for each case (see appendix three). As a result, while experts probably do selectively key upon more relevant information than do novices, in the current study, this effect may have been masked by the variety of potentially relevant diagnostic options.

The finding that of the total case elements generated in the protocols (R+F), experienced clinicians generated a significantly larger number of references than novices is interesting. Since subjects are generating case elements throughout the protocols, the greater use of reference statements by experienced clinicians suggests that the initial case information is being encoded differently across groups. This finding supports the notion that experts are "pre-processing" information while in the process of forming their initial task representations. "Pre-processing" is a concept which supports the representational theory of expertise in that it demonstrates the subjects' automatic altering of information to fit evolving cognitive task representations.¹ ~~Novices are less likely to "pre-process"~~

information in this way because they do not form cognitive representations as quickly, as consistently, or with as much complexity.

The purpose of the present study was to demonstrate the potential for methodology to differentiate experts from novices on a relatively ecologically valid task. According to Ericsson and Smith (1991), defining tasks which differentiate experts from novices in this manner is the first step in the "expertise approach." A secondary purpose of this study was to begin a theoretically informed analysis of the processes making for expert performance in this domain. Some attempt was also made to theorize about the nature of expertise and to link this concept with other theoretical accounts in the literature.

The primary limitations of this study are both methodological and practical. One methodological weakness of this research is that, while the open-ended nature of the subjects' task was a deliberate aspect of the design, there are not other standardized research designs to which to

¹This automatic cognitive processing is analogous to the perceptual "pre-processing" which occurs in the apparatus of visual system.

compare this approach. A study examining expert-novice ability to diagnose "accurately" using the DSM system, for example, would inform the results of the present study by providing a better understanding of differences in basic factual diagnostic knowledge between these groups.

A practical design problem with this study is the small number of subjects. This limited both statistical power and the generalizeability of these results. Another problem was that the criteria used for selection of experts and novices were simplified for practical reasons. Chase and Simon (1973), for example, suggest that ten years of experience in the context of adequate training is necessary for the development of expertise. Other researchers emphasize that experience alone is a necessary, but not sufficient condition for the development of expertise. In the current study, however, the availability of "expert" subjects was limited. Ultimately, the decision to define expert subject as clinicians with three years of post doctorate experience was arbitrary. Also, the fact that all the experienced subjects were male was due to local availability of subjects and may have influenced the results.

Another difficulty with this research is that it was

impossible to code the protocols blind to group membership. Although this was in part due to unavailability of additional coders, it was also evident from the protocol content as to which group subjects belonged. While this aspect of the study methodology reduced the objectivity of coding, it also allowed the coder to use the interpersonal context in making coding decisions.

In spite of the weaknesses outlined above, the current methodology demonstrates that the "expertise approach" may be applied to domains of interest involving open-ended and ecologically valid tasks. In addition, this research demonstrates that differences may be shown between experienced and novice clinicians based on predictions generated from the expertise literature. This methodology is a preliminary attempt to define the type of clinical task which may be used to examine processes leading to superior expert performance. Future research in this area may further refine and standardize this approach, as well as develop means of operationalizing the complexity of clinicians' cognitive representations. Ultimately it is hoped that such attempts may bridge the ecological validity of purely phenomenological description with the systematic

operationalization of quantitative research methods.

REFERENCES

- Anzai, Y. (1991). Learning and use of representations for physics expertise. In K. A. Ericsson and J. Smith (Ed.), *Toward a general theory of expertise: prospects and limits* (pp. 64-92). Cambridge, MA: MIT Press.
- Anzai, Y., & Yokoyama, T. (1984). Internal models in physics problem solving. *Cognition and Instruction*, 1, 397-450.
- Barsalou, L. W. (1987). The instability of grade structure: implications for the nature of concepts. In U. Neisser (Ed.), *Concepts and conceptual development: ecological and intellectual factors in categorization* (pp. 101-140). New York: Cambridge University Press.
- Camerer, C. F., & Johnson, E. J. (1991). The process-performance paradox in expert judgement: how can experts know so much and predict so badly? In K.A. Ericsson and J. Smith (Ed.), *Toward a general theory of expertise: prospects and limits* (pp. 195-217). Cambridge, MA: MIT Press.
- Charness, N. (1981). Aging and skilled problem solving. *Journal of experimental psychology: General*, 110, 21-38.
- Charness, N. (1989). Expertise in chess and bridge. In D. Klahr & K. Kotovsky (Ed.), *Complex information processing: The impact of Herbert A. Simon* (pp. 183-208). Hillsdale, NJ: Erlbaum.
- Charness, N. (1991). Expertise in chess: the balance between knowledge and search. In K. A. Ericsson and J. Smith (Ed.), *Toward a general theory of expertise: prospects and limits* (pp. 39-63). Cambridge, MA: MIT Press.
- Chase W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215-281). New York: Academic Press.
- Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5, 121-152.
- de Groot, A. (1978) *Thought and choice in chess*. The Hague: Mouton. (Original work published in 1946.)
- Dreyfus, H. L. & Dreyfus, S. E. (1986). *Mind over machine: the power of human intuition and expertise in the era of the computer*. New York: The Free Press.
- Elstein, A. S., Shulman, L. S., & Sprafka, S. A. (1978). *Medical problem solving: An analysis of clinical reasoning*. Cambridge, MA: Harvard University Press.

- Elstein, A. S., & Bordage, G. (1988). Psychology of clinical reasoning. In J. Dowie & A. Elstein (Ed.) *Professional judgement: a reader in clinical decision making*. (pp. 109-129). New York: Cambridge University Press.
- Ericsson, K. A., & Smith, J. (1991). Prospects and limits of the empirical study of expertise: an introduction. In K. A. Ericsson and J. Smith (Ed.), *Toward a general theory of expertise: prospects and limits* (pp. 1-38). Cambridge, MA: MIT Press.
- Garb, H. N. (1989). Clinical judgement, clinical training, and professional experience. *Psychological Bulletin*, 105, 387-396.
- Hamm, R. M. Clinical intuition and clinical analysis: expertise and the cognitive continuum. In J. Dowie & A. Elstein (Ed.) *Professional judgement: a reader in clinical decision making*. (pp. 78-108). New York: Cambridge University Press.
- Hammond, K. R., McClelland, G. H., & Mumpower, J. (1980). *Human judgement and decision making*. New York: Hemisphere.
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgement under uncertainty: Heuristic and biases*. New York: Cambridge University Press.
- Kaplan, H. I., & Sadock, B. J. (Eds.). (1989). *Comprehensive textbook of psychiatry*. Baltimore: Williams and Wilkins.
- Keil, F. C. (1987). Conceptual development and category structure. In U. Neisser (Ed), *Concepts and conceptual development: ecological and intellectual factors in categorization* (pp. 201-233). New York: Cambridge University Press.
- Lakeoff, G. (1987). Cognitive models and prototype theory. In U. Neisser (Ed), *Concepts and conceptual development: ecological and intellectual factors in categorization* (pp. 63-100). New York: Cambridge University Press.
- Larkin, J. H. (1977). *Problem solving in physics*. Technical report, Group in Science and Mathematics Education, University of California, Berkeley.
- Larkin, J. H., McDermott, J., Simon, D.P., & Simon, H. A. (1980). Models of competence in solving physics problems. *Cognitive Science*, 4, 317-345.

- Medin, D. L., Wattenmaker, W. D. (1987). Category cohesiveness, theories, and cognitive archeology. In U. Neisser (Ed), *Concepts and conceptual development: ecological and intellectual factors in categorization* (pp. 25-62). New York: Cambridge University Press.
- Meehl, P. E. (1954). *Clinical verses statistical prediction: A theoretical analysis and a review of the evidence*. Minneapolis: University of Minnesota Press.
- Miller, G. A. (1956). The magic number seven, plus or minus two. *Psychological Review*, 63, 81-97.
- Neisser, U. (1987). From perception to conceptual structure. In U. Neisser (Ed), *Concepts and conceptual development: ecological and intellectual factors in categorization* (pp. 11-24). New York: Cambridge University Press.
- Patel, V. L., & Groen, G. J. (1986). Knowledge-based solution strategies in medical reasoning. *Cognitive Science*, 10, 91-116.
- Patel, V. L., & Groen, G. J. (1991). The general and specific nature of medical expertise: a critical look. In K. A. Ericsson and J. Smith (Ed.), *Toward a general theory of expertise: prospects and limits* (pp. 93-125). Cambridge, MA: MIT Press.
- Rosch, E. (1973). On the internal structure of perceptual and semantic categories. In T. Moore (Ed.) *Cognitive development and the acquisition of language*. New York: Academic Press.
- Sarbin, T. R. (1944). The logic of prediction in psychology. *Psychological Review*, 51, 210-228.
- Scardamalia, M., & Bereiter, C. (1991). Literate expertise. In K. A. Ericsson and J. Smith (Ed.), *Toward a general theory of expertise: prospects and limits* (pp. 172-194). Cambridge, MA: MIT Press.
- Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.) *Children's thinking: What develops?* (pp. 325-348). Hillsdale, NJ: Erlbaum.
- Spitzer, R. L., Gibbon, M., Skodol, A. E., Williams, J. B. W., & First, M. B. (Eds.). (1989). *DSM-III-R casebook: a learning companion to the Diagnostic and statistical manual of mental disorders (third edition, revised)*. Washington, DC: American Psychiatric Press.

APPENDICES

APPENDIX 1
Protocol Coding System

Although scoring protocols using this system requires a clear understanding of criteria for coding the various categories, employing the actual procedure is fairly simple. Each segment of the protocol is scored independently of other segments and the entire context of the segment is used to make coding decisions. This semantic based method is in contrast to propositional based protocol analysis where a line by line coding is conducted using only the propositions contained within specific grammatical units. The coder utilizing this semantic system must consider the protocol segment as an overall context, as well as working on a line by line basis. The coder must also have a working knowledge of the original case material in order to recognize features and references used by the subject while "thinking aloud." Because of the moderate complexity of the subject matter, coders should have a some basic knowledge of psychopathology categories and terminology. In order for results to be considered reliable, coders must undergo training in this system, including the completion of practice protocols.

As a general rule, coding should be conservative, with unclear instances placed in the more concrete of the categories under consideration (eg. features are more concrete than references).

Coding Categories and Notations

F#	Features
R#	References
SI#	Speculative Idea
Sy#	Symptom hypothesis
D#	Diagnostic hypothesis
()	Repetition Coding
q	Question notation
-	Negation notation

Coding Criteria

Features: F#

Features are the basic "unrefined" elements of the protocol. A feature is scored for each instance of factual information from the case material mentioned by the subject. The subject need not have used the exact wording of the case material, but the information must be clearly translated from the original and must be explicitly stated. A feature is scored for each element of information translated from the case material.

#1 Example: "So he quietly goes in the phone booth, takes off his clothes, paints himself red and gesticulates and then moves across the street, doesn't try to avoid arrest, then bites the cop and then settles down."

This example is scored for eleven separate features (underlined) taken from the case material. Note that verb-subject constructions such as "paints himself" and "bites the cop" are scored as single features, while descriptive

adjectives such as "red" are scored as a separate feature. Features may be embedded within speculative ideas, each element being scored separately as an independent code. Features which are identified by the subject as a symptom hypothesis are scored as symptoms and the feature coding is placed in brackets following the Sy# notation (eg Sy3[F1]).

References: R#

References are similar to features, but represent the next stage of abstraction from the case material.

References are scored for instances in which the subject refers to a specific case element (feature), but does not explicitly restate the information. References may be descriptive or concretely interpretive, but do not involve theoretical speculation beyond the case elements.

References are distinguished from speculative ideas in that the latter involve the implicit or explicit use of theory to "expand" beyond the original material to include speculation regarding intent or diagnostic meaning. The coder should be careful to include in reference scoring only those instances referring to a specific or concrete set of feature elements.

General speculation relating groups of features according to the theoretical understanding incorporated by the subject should not be scored as a reference. This distinction can be subtle. In more difficult instances the coder's decision should be based on an estimate of the subject's intended communication derived from the overall context of the segment. As a general rule, if the subject is confining him or herself to describing the case material, then reference coding is applicable. If the subject is attempting to explain, categorize, or illuminate some aspect of the case material, then the element should not be coded as a reference.

#2 Example: "Well, his erratic, irrational actions in the face of, uh, it doesn't seem like he was demonstrating or protesting or whatever, so it was a kind of solitary, bizarre action."

In this example the first and last underlined segments are examples of references. Note that, unlike feature coding, individual descriptor adjectives such as "solitary" are not coded separately. The middle segment is not coded as a

reference because it involves speculation regarding intent (use of a personal theory to go beyond the case material). The first and last segments are concretely interpretive of the case material, but are less concrete than simple feature descriptions. Like features, references may be embedded in speculative ideas and are coded separately. References may be identified by the subject as symptom hypotheses, in which case they are coded as symptoms with the reference notation included in brackets (eg Sy3[R2]). The third statement in example #2 (bizarre behavior) is an example of a reference also used as a symptom hypothesis.

Speculative ideas: SI#

Speculative ideas, symptom hypotheses, and diagnostic hypotheses differ from features and references in that they reflect the subject's active manipulation of the case material in accordance with his or her clinical base of knowledge. These three categories differ from each other, however, in degree of specificity related to the task of providing a diagnosis. Given the limited nature of the case information provided to subjects, all three categories

involve some degree of speculation. In general each type of coding is used for instances where the subject is clearly using personal or general psychological theory to go beyond the given case material. It is the specificity and type of theory used which differentiates the coding of these three categories. Such coding is sometimes difficult because the theories used by subjects are often implicit and may not be verbally expressed within the protocol, or even consciously recognized by the subject. As with the reference coding, consideration of the overall context of the subject's communication within the segment is often necessary to differentiate categories.

Speculative ideas are a general category coded for instances where the subject clearly speculates beyond a descriptive level with the case material, but does not hypothesize about symptoms or diagnostic ideas. Speculative ideas often involve the subject pondering the patient's intentionality or purposefulness with regard to behavior. This category includes elaborated theoretical explanations or inferences taken from the case material by the subject in an effort to illuminate or give clinical meaning to features or references.

Symptom Hypotheses SY#:

Symptom hypotheses are coded for instances in which the subject explicitly or implicitly identifies a clinical aspect of the case as related to a diagnostic category. A symptom hypothesis is coded when: 1) there is a clear link to a diagnostic hypothesis stated within the segment, 2) when the subject identifies an aspect of the clinical situation in diagnostic language which clearly implies a link to a potential diagnostic hypothesis, or 3) when the subject speculates about the specific etiology of a psychopathological condition evident in the protocol. In coding symptom hypotheses the coder is attending to the subject's use of language to make judgements regarding the subject's intended communication. If the idea is being used in connection with a diagnostic line of reasoning and involves diagnostic language, then Sy# coding is appropriate. If the idea is part of a more open ended speculation unrelated to specific categorical diagnostic thinking, then SI# is coded. As a general rule, symptom hypotheses are more specific to case aspects and more descriptive than speculative ideas. Symptom hypothesis

coding often applies to separately coded features or references when the subject modifies these with adjectives identifying their use as symptom hypotheses.

#3 Example: "She was sexually molested by her surrogate father."

The first statement is an example of an Sy#[R#] coding. The R# coding is based on the subject's reference to an incident of a 14 year old girl being fondled in the original case material (case two). The Sy# coding is based on the modification to "molested," which indicates that the subject is using this aspect of the case as an etiological factor in considering a possible psychopathological condition. The last statement in this example is coded as a feature.

Diagnostic Hypotheses D#:

Diagnostic hypotheses are fairly straight forward and more easily recognized than other coding categories. Since the instructions to the subject are to diagnose the patient presented in the case, obvious attempts by the subject to do

this will usually be scoreable as diagnostic hypotheses. Diagnostic hypotheses are scored for instances in which the subject presents a clear cut diagnostic label or overarching clinical concept accounting for the presence and/or absence of symptoms or features in the case material. These are distinguished from symptom hypotheses and speculative ideas in that diagnostic hypotheses are intended by the subject to integrate all of the relevant case information into a single, or few, accepted psychological or psychiatric conceptualizations. Often these conceptualizations will be presented in the form of "disease entities" or syndrome disorders. Sometimes a subject will present a diagnostic term (eg. "paranoid") to explain some portion of the data, without presenting an overarching hypothesis (eg. "he's just acting paranoid"). These instances are not scored as diagnostic hypotheses and the scorer's judgement is required to determine whether a given speculation is intended by the subject as an attempt at diagnosing the case.

#4 Example: "What would stand out first of all would be the assumption that he is psychotic, and then to look at what ,uhm, whether there's

drug involved (sic), whether this is some
kind of induced ...(sic) or whether this is a
break, a psychotic break in a 17 year old."

In this example, the first, third, and fourth underlined statements are coded as diagnostic hypotheses. The second statement is a symptom hypothesis, but is not integrated into a diagnosis hypothesis until the third statement (ie a drug induced psychosis). The last statement is coded as a separate diagnostic hypothesis from the first statement because the context of the segment indicated clearly that the subject was considering a schizophrenic break as a possible diagnostic hypothesis. The first statement in contrast is explaining mental status information.

Question Notation:

This notation is not actually part of the coding system. It is noted on the raw protocols for clarification purposes and does not effect score totals. Question notation is used to indicate instances when the subject speculate about diagnostic hypotheses or symptom hypotheses

which are not addressed by the case information. These usually occur near the end of the protocol segment in response to the examiner prompt: "what additional information would you look to next..." The language used by subjects is usually in the form of "wondering" about symptoms or diagnoses which may be addressed in upcoming case segments. Note that these symptom and diagnostic hypotheses are coded as such, regardless of the question notation.

Negation Notation:

Like question notation, negation notation is not considered in the quantitative coded totals. Occasionally subjects will indicate the absence of particular symptoms, or will rule out a particular diagnosis. These instances are coded as symptom and diagnostic hypotheses as usual. A negative sign is placed before the standard notation to indicate a negation. As with question notation, the symptom and diagnostic hypotheses are counted in the quantitative totals, ignoring the (-) notation.

Symptom Relevance:

Symptom relevance notation is added to F#, R#, and SI# codes for instances where these elements are judged by the examiner to have some relation to DSM symptoms, but are not specifically linked to the subject's diagnostic hypotheses (see example #5). This coding is more liberal than for symptom hypotheses. Often subjects will use symptom relevant ideas along with symptom and diagnostic hypotheses within a line of diagnostic reasoning. Symptom relevance is noted by underlining the F#, R#, or SI# code.

Repetition coding:

Subjects often repeat themselves in verbal protocols. Identical information is recoded and placed in parentheses. Repetitions are not included in coded totals.

Coding Examples:

#5 Example: "Uh, the withdrawn (sic), I wonder if he's having, would be having hallucinations

as he was talking to this person, for example, because he didn't seem to know to answer this thing about do you hear voices that other people don't seem to hear."

In this example the first statement is a symptom relevant reference. This statement is not coded as a symptom because it is descriptive of verbatim material from the case and is not identified (either in this excerpt or in the larger segment context) in relation to a diagnostic idea. A key factor in this distinction is that "withdrawn" could, or could not be, considered a pathological indicator depending on context. The second statement in contrast is clearly a symptom speculation linked to an implicit diagnostic hypothesis (psychosis). Note that this statement would also be coded as a question. The last statement is coded as a feature because it is taken literally from the case text.

#6 Example: "Well, uhm, a couple of things and one is the general sense all the way through the material of someone who is disorganized in

their thinking and certainly has some rather
complicated, uh, thoughts, ideas about
himself and about art or whatever, so there's
some kind of ideation, whether that becomes
bigger in terms of paranoid thinking or
whatever, but that would be certainly the
direction I would think about..."

In this example the first underlined segment is a symptom hypothesis, the second is a symptom relevant speculative idea, the third is a reference, and the fourth is a symptom hypothesis. The subject's use of the term "disorganized thinking" in the first statement can only be in reference to psychopathology (including intoxication). The second statement is suggestive of the subject considering possible thought disorder, but is not specifically indicated. "Complicated thoughts" are not definitively linked to a diagnostic idea. The third statement is simply a descriptive reference to a feature of the case material. The fourth statement is close to suggesting a diagnostic hypothesis (a paranoid disorder), but limits itself to describing pathological thinking and so is coded as a

symptom hypothesis.

Notation

Coding for the categories outlined above is performed by writing the notation for each instance along the right hand margin next to the line of text containing the highlighted element. Multiple notations are separated by a comma. Each instance of a coding category is numbered sequentially, starting with one, within each segment.

Scoring Procedure

The procedure for scoring involves four stages:

- 1) Reading
- 2) Underlining
- 3) Coding
- 4) Reviewing and checking

1) **Reading:** The coder begins each segment by reading the entire segment. While reading the coder should pay

special attention to the overall context of the points expressed by the subject. By reading the entire segment the coder becomes familiar with the structure of the subject's "thinking aloud" with regard to the particular case material.

2) **Underlining:** After reading the entire segment, the subject rereads through the segment, stopping to highlight scoreable instances according to the scoring categories. At this stage the coder need not make detailed distinctions regarding categories, but need only identify scoreable aspects of the text.

3) **Coding:** After reading and underlining the text, the coder will have some familiarity with segment and may tackle the task of coding the categories. Most of the coding will be straight forward. More difficult elements require extra time and effort. Repetition coding is also completed during this stage.

4) **Reviewing and checking:** In the last stage the coder should review each coding to double check accuracy.

This completes the coding for the particular segment
and the coder moves to stage one with the next segment.

APPENDIX 2

Case 1

Demographics and presenting problem

Ronald is a 17 year old single white male. He was seen for consultation in a medical hospital psychiatric ward after being brought to the unit by the police. Ronald had been arrested for causing a public disturbance after he had gone to a busy downtown intersection, found a phone booth, removed his clothing, streaked his body with red paint he had brought with him, and run out into the intersection gesticulating wildly. After the police arrived he attempted to avoid capture by moving across the street, but he did not attempt to flee and he continued his gesticulations, waving to passing motorists and dodging traffic. When finally cornered he submitted passively to arrest and then proceeded to bite the arresting police officer. The police said that he had a "crazed" look in his eyes when he was arrested, but that he calmed down quickly in the patrol car.

Interview observations and subject's story

In the consultation interview, Ronald had been allowed to clean up, but traces of red paint were still evident in his hair. He wore a hospital smock to the interview. The (female) interviewer's impression was that Ronald was an attractive and immature young man who appeared significantly younger than his stated age of seventeen. He seemed very shy and had trouble holding eye contact with the examiner. When he did make eye contact he smiled as though embarrassed and quickly looked away. His eyes had a "glassy" look to them, but he was otherwise alert and socially responsive. His speech was coherent, but his use of language had an odd disconnected quality which gave the examiner a sense that he was either being evasive or had trouble communicating his thoughts. Ronald's voice was also quiet and he had a tendency to mumble.

When asked about the incident which brought him to the hospital, Ronald explained that he was protesting a local museum's announcement that the price of admission was to be raised to pay for a new exhibit. This exhibit included a collection of ancient Roman military artifacts, a subject in which Ronald evidently has a great interest. He digressed at some length about Roman artifacts in general and related a litany of loosely related facts about ancient Rome. Ronald repeatedly expressed his opinion that such exhibits ought to be free to the public. When asked why this in particular upset him so, he did not seem to understand the question and proceeded to rail about how people should not tolerate such injustice.

When asked about the events leading to his arrest, Ronald seemed uncertain. He remembered reading the night before about the new exhibit and that increased fees were to be charged. Ronald lay awake thinking about this all night. He reread the article the next morning and impulsively ran out to the shed to get the paint. Ronald was confused about what happened next and could not relate a clear account of events between leaving his house and arriving at the hospital. When asked how he had felt when he read that the museum was charging fees, Ronald repeated his tirade about how unjust this was and never really answered the examiner's question.

Ronald has no previous history of arrest and has never met with a mental health professional in the past. When asked about "hearing voices that other people don't appear to hear" he did not seem to understand the question, stating "I hear voices when people talk." When pressed for a clearer answer, Ronald became anxious and distracted, stating only that "voices come from people talking." Ronald at this point seemed upset and he became withdrawn and unresponsive. It was decided to terminate the interview to avoid agitating him further.

Physical Exam

Ronald was given a physical exam when admitted to the psychiatric ward. He initially complained of nausea due to a bit of paint he accidentally ingested, but this was not deemed serious. No physical abnormalities were discovered, but it was noted in the exam that Ronald is under weight for his age with underdeveloped musculature. Cardiac and respiratory functions were normal. Blood counts showed a slight anemia, but were otherwise normal. Drug screen was positive for THC.

When asked about eating, Ronald related that his mom has always called him a "picky" eater and that he doesn't like red meat. He complained that his mother is always preparing foods he doesn't like. His favorite foods are "Jiffy peanut butter and Wonder bread, McDonald's chicken McNuggets, and Ore-ida french fries." He apparently lives almost exclusively on these foods. He answered questions about his sleeping habits by digressing into a diatribe about how in a republic like ancient Rome, one could eat, sleep, or come and go as one pleased. He then launched into another description of Roman artifacts and had to be interrupted. Ronald had evidently slept very little in the week preceding the incident at the intersection, and his sleeping habits in general were noted as erratic.

Family History

Ronald is from a middle class family residing in the suburbs of a large metropolitan area. His father is a professional with a successful private practice and good income. His mother works half time in real estate sales and is active in the community. She started working when Ronald was two years old. Ronald's father is described as "a solid personality" with somewhat rigid values. He apparently has difficulty showing his feelings. His relationship with his son has been distant and strained since Ronald hit adolescence. In contrast, Ronald's mother is considered warm and vivacious with many friends. She has always paid special attention to Ronald, doting on him and keeping all the neighbors abreast of her "baby's" development. Ronald has three older sisters, all of whom have left home for college and have gone on to successful professional careers. His sisters are married with children and, at family gatherings, Ronald reports that he feels uncomfortable around these children. While growing up, Ronald played with his sisters often, but now has little contact with them other than on holidays. He reports an incident when he was about age seven where at a neighbor's house his sisters and their friends stripped him to his underwear and left him to walk several blocks home by himself. Ronald admitted that he didn't like this, but was quick to defend his sisters actions as "just a joke." When asked how he had felt about having this done to him, Ronald looked puzzled and repeated that it was "just a joke."

Interpersonal History

Ronald has always been shy and considered as a loner by his peers. As a child he spent much of his time entertaining himself and had only a couple of friends with whom he played infrequently. Ronald liked watching television and had a host of favorite programs which he never missed. He liked action shows with a fast pace. He tended to get upset if he was forced to miss any of his favorite programs, and at these times he was difficult to sooth or comfort, other than by allowing him to watch his show. Ronald still likes to watch TV and many of his favorite programs are the same as when he was a child.

When not watching TV, Ronald's time by himself was mostly spent reading or working out his "games." Ronald described these with some pride, explaining long sequences of actions he had developed as a kid, each of which had to occur in the proper order. He did not explain any meaning which these "games" held for him and there was no sense that these were particularly fun for Ronald. The examiner could not get him to talk about why these "games" were important or what significance they held for Ronald. The examiner's best guess was that Ronald assumed everybody naturally spent time in this manner.

With other children, Ronald was usually very passive and did not commonly initiate social play. He did seem to like repetitive social games, like "ring around the rosy," but other children quickly tired of playing these with him. Ronald never much liked the "pretend" games initiated by other children and he seemed uncomfortable and awkward when compelled into imaginative play situations. The exception to this was when he was young, Ronald was often enlisted by his sisters to play the make believe role of "the baby," a part which only required that he sit passively and be doted upon.

For the most part, Ronald was very quiet and considered odd by other children. As he began his adolescent years he became even more quiet and shy, actively avoiding most social groups and only interacting with a few friends. He was unpopular in school, but was generally protected by his older sisters, who were socialites of a sort. His teachers thought him naive and immature, although he presented few behavior problems. When his sisters went off to college, Ronald got along badly at school and was often the butt of practical jokes. Ronald has never gone on a date or attended a school social activity. He currently has no friends and spends most of his time watching television or rereading his books about Roman artifacts.

Educational History

Ronald had learned to read before he started school, a fact which his mother attributes to public television programs. In school he advanced quickly in reading and showed a good ability to memorize material quickly and accurately. His grades were generally mediocre, however, and he did especially poorly in mathematics. In high school he became interested in Roman artifacts and spent much time in the social studies rooms after school studying artifacts and archeological techniques. Ronald decided at this point that he wanted to pursue archeology as a career and began telling people he would someday become a famous archeologist. His father strongly discouraged Ronald's interest stating "if you want to starve, be an archeologist." Ronald's social studies teachers were supportive of his interest in archeology, but were uncertain as to how much encouragement to give him because Ronald seemed to have little aptitude for history, and little abstract understanding of how history may serve as a context for the pursuit of archeological discoveries. Ronald, in fact, had great trouble in all courses, such as history, which required him to express written ideas. Ronald insists that the ability to write essays is unimportant to his goal of becoming an archeologist. He does not understand why others make this an issue. Ronald had graduated from high school eight months before the incident in the intersection.

Following graduation, Ronald continued to live at home and was to have applied for college admission at a local university. When he discovered that he would have to write an essay describing himself and his interests, Ronald put off the application and allowed the deadline to pass. At his father's insistence he applied for and got a job at a yogurt shop. This job lasted two months before Ronald was fired or laid off for unspecified reasons. Since that time Ronald has lived at home and not looked for work or pursued his college application. He still reports that he will become a famous archeologist and will not discuss alternative options for his future.

A and D History

Ronald had a friend when he was about age ten who pressured him into huffing paint in his parent's garage. This occurred on a few occasions, but ended abruptly when the boy and his family moved away. As an adolescent Ronald's sister, who was away at college, "got him high" on marijuana at a party she had arranged for him to attend. This was repeated on two other occasions while the sister was visiting on holiday. Ronald asked for and received a small amount of marijuana from his sister which he kept for several years. He eventually smoked this by himself a few days prior to the incident at the intersection.

Ronald's parents have a relaxed attitude about drinking alcohol and he was given small amounts on special occasions as a child. When he was fifteen his sisters were visiting on holiday and he "raided" the liquor cabinet, drinking a mixture of several spirits. This made him very ill. He had not consumed alcohol since that episode until approximately one month prior to the incident of his arrest. Recently he has taken "a beer or two" from his father's refrigerator.

Medical and Developmental History

Ronald's mother reports that his birth involved a long and very difficult delivery. He reportedly had episodes of cyanosis in the first forty eight hours. The suffering his birth caused his mother has been repeatedly related to Ronald by his mother throughout his life. His language and motor development were somewhat delayed in comparison to his sisters, but his parents did not consider this delay significant. He was a quiet, self contained baby and child, but presented no obvious problems. As a child he was poorly coordinated compared to peers and he disliked athletic activities. Ronald had no outstanding medical problems during childhood other than occasional bouts of anemia and the usual childhood diseases.

APPENDIX 2

Case 2

Demographics and Presenting Problem

Sally is a 35 year old divorced Hispanic woman seeking psychiatric consultation at a university counselling center. She is worried that she might be going insane. For the past two years she has experienced increasingly frequent episodes of feeling "outside" herself, as if her body is dead and without sensation. These episodes are accompanied by disorientation, or "foggy thinking," in which she sees herself as if from the outside, like a "zombie on autopilot." During the episodes Sally can identify no feelings other than that of "deadness" in her body centered around her chest, and a kind of giddiness which reminds her being anesthetized.

Interview Observations and Subject's Story

Sally had called ahead to the center for an appointment and arrived on time. She was well groomed and dressed in a white blouse and long professional skirt. During the interview Sally seemed embarrassed to talk about herself and nervous about consulting a mental health professional. She was, however, socially appropriate and articulate, making good eye contact and expressing a sense of humor. Early in the interview she asked whether the interview was confidential and seemed very concerned about this. When asked if she could be more specific about her concerns, she said only that she was anxious about people knowing about her problem. The (male) interviewer's impression of Sally was of a reasonably intelligent and attractive, if somewhat uptight, woman who was trying in the interview to make the best of an uncomfortable situation.

What initiated Sally seeking help at this time is that the frequency of the episodes has increased from about once a month to about once a week. She is unsure of the duration of the "attacks," but guesses that they usually last from half an hour to possibly several hours. Recently Sally confided in a female friend who advised her to seek help at the counseling center. It took her a week to work up the courage to make an appointment.

In elaborating her story, Sally talked of her fear of having a "breakdown" like that of cousin several years before. She was not sure about exactly what had happened with her cousin, but it apparently had been a great embarrassment to the family. What upsets her most about the episodes is her sense of losing control over herself. Sally describes the onset of the episodes like a slow wave which washes over her, usually when she is out in public and feeling anxious. Sometimes, however, the episodes occur when she is by herself at home or when she is driving. She is especially concerned that she will wreck her car and die, or burn her house down as a result of the disorientation. When she feels the onset of an episode, Sally fights the disorientation by shaking her head and speaking "No!" to herself. This sometimes clears her head temporarily, but the sensation of deadness and giddy disorientation usually returns in short order. Sally describes the experience of episodes as like a "split consciousness" with one part of her giddy and "wild," while another part is observing and feeling panicked and adrift. The feeling of deadness in her body is prominent over both sides of the "split," and is very unpleasant and frightening. During the episodes she has "wild" thoughts of hitting or stabbing her chest, or of slowly cutting her genitals with a razor blade. These thoughts both excite and horrify her and she is frightened that if she "breaks down" she might hurt herself. Sally expressed with a smile that it would be very embarrassing to have to explain these injuries to the emergency room physician, much less to her mother if it came to that.

Family and Social History

Sally was born and raised in a large Catholic Hispanic family in a mid sized southwestern city. She has two older brothers, two younger brothers, and a younger sister, all of whom were raised in the same home. Her father was a farm laborer who abandoned the family when Sally was age five. Two years later another man who worked as a mechanic moved into the household and remained for seven years before he too moved on. This man provided for the family, but was emotionally distant with the children and occasionally would hit Sally's mother when the children were present. At age fourteen Sally remembers being alone with her surrogate father who massaged her breasts for a time and then reached down her pants to touch her genitals. This incident was interrupted when one of her siblings returned home unexpectedly and was never repeated. This man left the family several weeks later and Sally has not seen him since. At the time Sally was only surprised by this advance, but later felt great shame and has never before told anyone of the incident.

Sally's mother is a woman who expresses very traditional values and has strong ties to a large extended family. The mother supported the family by collecting welfare checks, as well as by working part time as a cleaning lady for wealthy Anglos on the other side of town. Sally feels that her mother is a "good woman," but also described her as moody and "hard to read" emotionally. Sally believes that her mother feels inferior and has never really been happy, but that her mother would never admit this to her children.

At an early age Sally was expected to take responsibilities for many chores around the house, especially caring for the younger children while her mother worked. Sally remembers resenting her older brothers who were free to do as they pleased while she had to baby sit. One of her older brothers is now in prison for assault, while the other is unemployed and supported by extended family members. Sally did not want to talk about these two siblings, but was happy to tell the examiner about her little sister who is a junior at a local community college. Her two other brothers are both employed as construction workers and have families of their own.

Educational and Employment History

Sally remembers that she liked school very much as a child, but that she had trouble because she often didn't have time to do her homework due to the chores at home. She sometimes had to miss school in order to watch her younger siblings while her mother earned extra money cleaning houses. She especially liked mathematics and English classes, although she doesn't consider her abilities in these areas particularly strong. Her teachers saw her as hard working, but stubborn, and she occasionally got into trouble for getting too "wild." Sally remembers grade school as a place where during recess she could be herself, running around, playing sports, and flirting with boys. In high school she became interested in education as a profession and for a time had the idea that she would go to college and become a teacher. After graduating from high school, however, she got a job as a secretary at which she worked successfully for eight years. At age twenty six she decided to quit her job and take out student loans to go to a community college for a business degree. Her grades were barely passing and after two semesters she dropped out feeling very disappointed in herself. Her mother told her that it was silly for her to want more education and that she should find a man and have children. Sally remembers thinking at the time that she must be a great embarrassment to her family. She went back to work as a secretary, but could find no satisfaction in this, becoming very depressed and unhappy with her life. At one point she became so depressed that she thought of suicide frequently, but she never acted on these thoughts and did not seek professional help. At age thirty she decided to "get it together" and she forced herself to begin saving money with the goal of returning to school. Two years ago she dropped down to half time at work and resumed her course work in business. Her grades have been good and she is approaching the point where she can graduate. While in school she has also "allowed" herself to take two education courses, but she considers pursuing teaching as an unrealistic dream and as just her "wild" side coming out. Sally blushed as she said this. Sally's "attacks" of disorientation have begun to interfere with her school work and she has doubts that, if things continue as they have been, she will be able to continue in school.

Interpersonal History

Sally has always presented herself to people as somewhat conservative and old fashioned, but she acknowledges that there is a more adventurous side to her which comes out when she least expects. She feels that this has gotten her into trouble in the past and she tries hard to keep this side "bottled up tight." This adventurous side is more likely to emerge when she has been "partying" and she is always sorry afterward.

Sally remembers her first experience of sexual intercourse which occurred when she was sixteen with a boy whom she knew after a party where they both had been drinking. Although she liked the boy and had known him for some time, she was so ashamed that this had occurred that she refused to see him after that. There were a few other similar incidents during and after high school in which she allowed a guy to pick her up at a party, but she always felt terrible afterwards and would force herself to "abstain" from any contact with males for a period of time.

When she was twenty two she married an older Hispanic man whom she had met through her older brother. This man worked off and on as a mechanic for a local auto dealership. While she was "in love" with him, he tended to drink heavily and occasionally would beat Sally severely. After two unhappy years she moved out and ended the marriage. Sally remembers that this was terribly difficult for her family and that she was especially ashamed of how this looked to her uncles and cousins who had not known her husband well. Sally's mother objected to the divorce and seems to have blamed Sally for the failure of the marriage.

After the divorce Sally completely "abstained" from dating men for four years, vowing to herself "never again." She "broke down" in her vow when she began attending college and had a boyfriend for a time. This relationship ended when he graduated and moved to another city. It was this breakup which she feels precipitated her depression and thoughts of suicide. After a year of "wallowing in my sorrow" Sally decided to "get it together" and renewed her vow to avoid dating men. After two years she "broke down" again and dated a man she had met shortly after returning to school. She liked this man very much and had "allowed" herself fantasies that they might marry. The relationship ended after a few months, however, when the man followed his career to another city.

For the past two years Sally has felt very lonely, a "weakness" which especially troubles her because in the past she has always felt able to get on well by herself. In spite of her loneliness, Sally is convinced more than ever that relationships with men can mean only trouble and that she doesn't "deserve to be happy." While she acknowledges that part of her would like a happy marriage and to have children, she "knows this will never happen." She is therefore determined to concentrate only on her career and on "getting it together" for herself emotionally.

Medical and Developmental History

Sally was born by natural child birth and the delivery was of relatively short duration. There were no early complications and her mother considered her an "easy" baby in terms of feeding and establishing consistent rhythms. She was breast fed and weaned at age two, when her younger brother was born. According to Sally, her mother reports that as a two year old Sally became very active and "wild," refusing to obey and getting into things while mother attended to the new born. Sally was told that she once talked of killing the baby and was so generally active and disruptive that her mother at one point even considered seeking professional help. This period of "rebellion," as Sally puts it, lasted until she began school at age five. She feels badly about having put her mother through this, but stated that this is "the price for childhood." She quickly corrected this to say "the price for having children." After beginning school Sally's behavior at home ceased to be disruptive and she was soon "allowed" to begin helping mother with chores. It was not until several years later when she realized other children did not spend all of their free time baby sitting that she remembers resenting the chores. Sally was physically healthy as a child, suffering the usual childhood diseases, but always recovering quickly without much fuss.

When she was sixteen Sally was terrified that she might be pregnant following a one night contact with a boy she knew after a party. This anxiety lasted for two weeks, during which she experienced somatic symptoms of nausea, before she finally visited a clinic for pregnancy testing. This same pattern of anxiety and somatic nausea confirming her worst fears was repeated after each "one night stand" until she finally began taking birth control pills at age twenty. The fact that she used and has continued to use birth control is very embarrassing to Sally and she greatly fears that her mother will "bust" her with the pills and "disown me." She justifies her use of the pills morally by telling herself that they make her period easier.

At age twenty seven Sally had a "scare" when the results of a routine pap smear indicated the need for a biopsy. Before learning the results she had convinced herself that she had cancer and would die, "alone and childless." It was during this period that she thinks that her first "attack" of disorientation and feeling like she was outside her body occurred. Although the results eventually indicated that there was no difficulty with her cervix, Sally is still haunted by, as she calls it, the "irrational thought" that she cannot have children. This incident occurred around the same time that a romantic relationship of over a year ended and it was soon afterward that she suffered her "depression." Sally's current health is good. She exercises regularly and is conscious of what she eats. Sally has no health complaints other than her "attacks" and that she feels that she needs to loose 15 pounds.

Physical Exam

Several weeks prior making an appointment at the clinic, Sally went to see a family physician. She did not want to reveal her fear that she might be going insane, so she complained only of feeling dizzy and weak at times. She was asked if she might be pregnant and she remembers that this made her angry, like she wasn't being taken seriously and that the doctor thought she was a "floozy." A physical exam revealed no notable abnormalities. Her coordination and reflexes were good and there was no evidence of neuropathy. A routine EEG showed no abnormalities and Sally was sent home with a prescription for Xanax and advised to return if the dizziness persisted. She filled the prescription but has kept the pills for a time when she might "need" them. Sally currently takes birth control pills and is given routine exams in connection with receiving these pills.

A and D History

Sally's mother does not drink and has never allowed the men in her life to have alcohol in the home. As such, Sally was not exposed to alcohol until she was about fourteen when she began to drink socially with friends from school. When she was sixteen she was "turned on" to marijuana by her older brother. She smoked this occasionally throughout high school. The man she married at age twenty two was a heavy user of cocaine, marijuana, and alcohol, and was involved in small time dealing of these substances. During the two years of their marriage, Sally occasionally used cocaine, LSD, marijuana, and alcohol. She and her husband sometimes used cocaine during sex, with him licking this off various parts of her body before intercourse. Sally once smoked opium with her husband before having sex and she remembers the feeling of that experience as like that of her current "attacks" of disorientation and "deadness" in her body.

It was with her husband that Sally first began to have fantasies of cutting her breasts and genitals. Her husband on one occasion used a sharp knife to shave her genital area prior to sexual intercourse. During this period Sally began to "play" with razor blades, running the blade along her thighs where "no one could see" if she actually did cut herself. She would do this while alone and naked on her bed. Sally insists that she only did this a few times when things with her husband were so bad that she needed "the release." Since her divorce she has resisted the impulse to "play" with razor blades in this way.

Sally describes her experiences with her husband in a subdued voice and claims that she can find no feelings in herself regarding the things which happened. This is strange to her because she "should" be embarrassed or feel bad about these things. Sally describes the memories and feelings as if these things happened to somebody she had read about, rather than being reflections of her own experiences. She is also puzzled about that period of her life because she has trouble "keeping things straight" in her memory and she wonders if she has "lost" some of her memories.

After the divorce, Sally's drug and alcohol use continued at a moderate level until she decided to go to college. She had friends with whom she "partied," but would occasionally get high by herself. When she began college, Sally decided to abstain from using "hard" drugs (cocaine and LSD), but continued occasional use of marijuana and alcohol. Sally currently uses marijuana about once a week and occasionally drinks socially. She feels that marijuana helps her to cope and function when she is feeling especially lonely or upset. The frequency of her "episodes" of disorientation is greater during periods when, for a lack of supply, she cannot obtain marijuana. Sally associates this with those being times of greater stress. Sally is very concerned that her mother will discover that she smokes pot, but has no real desire to stop and does not consider her drug use to be a problem.

APPENDIX 3
Diagnostic Hypotheses

Table 5

Case One Experienced									
Diagnosis	Exp 1	Exp 2	Exp 3	Exp 4	Exp 5	Exp 6	Exp 7	Exp 8	Exp 9
Psychosis	X	P		XF	X	XF	XF	P	XN
Drug Psych	X	XNF	X	XN	XN		X	XN	XNF
Schizophren	P	X	X	P	XNF	P	X	XF	
Bipolar		X		X	XN		X	XNF	
Autism		X	X				P	X	
Develope	X	XF					X		
Organic	XN	XN	XF	X	XF	XNF	XF	XNF	XN
Intoxication	XN		P	X	XN	XNF	XN	X	
Abuse/Dep		XNF					X		
PD NOS		XF							XF
Borderline		X							
Schizotypal		XF			P		XF	XF	P
Schizoid				XF			XF	XF	
Avoident									
Narcissistic									
Antisocial									
OC Per. Dis									
Malingering	X	X							X
Learning		X							
PTSD									XNF
Neurotic									
Adjustment	X						XN		
Affective									
Dissociative									
Total	8	13	5	7	7	4	12	9	7

P= Primary Diagnosis X= Mentioned in case N= Negated F= Mentioned in Formulation

Table 6

Case One Novices

Diagnosis	Nov1	Nov2	Nov3	Nov4	Nov5	Nov6	Nov7	Nov8	Nov9	Nov10
Psychosis	X		XF	XF	XF			XF	XNF	XF
Drug Psych		XNF			XN					
Schizophren			P	P	XF		P	XF		XF
Bipolar				XNF						
Autism						XF		XF		
Develope										
Organic				XF	XNF	X	X	XNF	X	XNF
Intoxication		X				X			X	XNF
Abuse/Dep	XNF		XN		XNF	XN				XN
PD NOS	P				XF					
Borderline			XF							
Schizotypal		P			XF	P			XF	P
Schizoid					XF				XF	X
Avoident									P	
Narcissistic					XF					
Antisocial				X						
OC Per. Dis	XNF				P					
PTSD										
Malingering	X	XNF						X		X
Learning								XF	XF	
PTSD										
Neurotic									XF	
Adjustment								P		
Affective				XNF						
Dissociative			X			X				
Total	5	4	5	6	10	6	2	7	8	9

Table 7

Case Two Experienced										
Diagnosis	Exp1	Exp2	Exp3	Exp4	Exp5	Exp6	Exp7	Exp8	Exp9	
Psychosis				XN			XN		XNF	
Major Dep			XNF	XF	XF		XNF		XF	
Schizophren										
Bipolar										
Dysthymia			P				XF		XF	
Cyclothymia										
Organic									X	
Deperson									X	
Abuse/Dep		XF		XF			XN	X		
PD NOS	XF	P		XF			XF		XNF	
Borderline		XF	X	XNF	P	P	X	P		
Histrionic		X					XNF		P	
Narcissitic					XNF					
OC Per. Dis				P						
Anxiety				X				XF	XF	
Panic/Agora	P	XF					XN		XNF	
MPD		X	X							
PTSD							P		XF	
Dissociative	XF						XF	X	X	
Neurotic				XF					XF	
Conversion						XF				
Total	3	6	4	8	3	2	10	4	12	

Table 8

Case Two Novices										
Diagnosis	Nov1	Nov2	Nov3	Nov4	Nov5	Nov6	Nov7	Nov8	Nov9	Nov10
Psychosis			XF					XNF		X
Major Dep		XF	XF	X	XF			P	XF	XNF
Schizophren									X	
Bipolar				X						
Dysthymia		XF			XF					XF
Cyclothymia										
Organic							X			X
Deperson				P						
Abuse/Dep				XF						
PD NOS		P			P		XF			
Borderline		XF	XF	XF		P			XF	P
Histrionic		XF		XF	XF		XF		P	
Narcissitic								X		
OC Per. Dis		XF								
Anxiety							X	XF		
Panic/Agora							X			
MPD			P	X						
PTSD	P									
Dissociative				XF	XF					
Neurotic										
Conversion		XF		X			P			
Identity							XNF			
Total	1	7	4	9	5	1	7	4	4	5

VITA

William Deal was born in Los Alamos, NM, the son of a physicist. He completed high school and began college at New Mexico State University in 1979. After graduating with a B.A. in psychology and a B.A. in philosophy in 1985 he spent a semester in London, England studying British language philosophy. Returning to the United States William moved to San Antonio, TX where he took pre-med, psychology, and philosophy courses at Trinity University. In 1988 he moved to Knoxville, TN to attend graduate school for a doctorate in psychology. William attended internship beginning July 1994 at New York Hospital, Cornell Medical Center in white Plains, NY.