Research Series No. 15

BMIS TECHNICAL MANUAL

Volume 1

THE INQUIRY THEORETIC APPROACH
TO CLINICAL PROBLEM-SOLVING
RESEARCH AND APPLICATION

Published By

The Institute for Research on Teaching
252 Erickson Hall
Michigan State University
East Lansing, Michigan 48824

Printed and Distributed
by the
College of Education
Michigan State University

1977

This work is sponsored in part by the Institute for Research on Teaching, College of Education, Michigan State University. The Institute for Research on Teaching is funded primarily by the Teaching Division of the National Institute of Education, United States Department of Health, Education, and Welfare. The opinions expressed in this publication do not necessarily reflect the position, policy, or endorsement of the National Institute of Education. (Contract No. 400-76-0073)
Institute for Research on Teaching

The Institute for Research on Teaching was founded at Michigan State University in 1976 by the National Institute of Education. Following a nationwide competition in 1981, the NIE awarded a second contract to the IRT, extending work through 1984. Funding is also received from other agencies and foundations for individual research projects.

The IRT conducts major research projects aimed at improving classroom teaching, including studies of classroom management strategies, student socialization, the diagnosis and remediation of reading difficulties, and teacher education. IRT researchers are also examining the teaching of specific school subjects such as reading, writing, general mathematics, and science, and are seeking to understand how factors outside the classroom affect teacher decision making.

Researchers from such diverse disciplines as educational psychology, anthropology, sociology, and philosophy cooperate in conducting IRT research. They join forces with public school teachers, who work at the IRT as half-time collaborators in research, helping to design and plan studies, collect data, analyze and interpret results, and disseminate findings.

The IRT publishes research reports, occasional papers, conference proceedings, a newsletter for practitioners, and lists and catalogs of IRT publications. For more information, to receive a list or catalog, and/or to be placed on the IRT mailing list to receive the newsletter, please write to the IRT Editor, Institute for Research on Teaching, 252 Erickson Hall, Michigan State University, East Lansing, Michigan 48824-1034.

Co-Directors: Jere E. Brophy and Andrew C. Porter

Associate Directors: Judith E. Lanier and Richard S. Prawat

Editorial Staff
Editor: Janet Eaton
Assistant Editor: Patricia Nischan
Abstract

This volume discusses BMIS from the point of view of a researcher interested in an overall understanding of the system. It includes the theoretical basis of the system, short descriptions of the system components, and results of applications of the theory and system in medicine and reading.
THE INQUIRY THEORETIC APPROACH TO CLINICAL PROBLEM SOLVING
RESEARCH AND APPLICATION

John F. Vinsonhaler, Ph.D.
Project Coordinator

Christian C. Wagner, M.S.
Systems Analyst

Arthur S. Elstein, Ph.D.
Consultant, Problem Solving Theory

Lee S. Shulman, Ph.D.
Co-Director, Institute for Research on Teaching

Copyright, 1977
Institute for Research on Teaching
Acknowledgements

The BMIS programming and research reported herein has been supported by grants from the National Institutes of Education, the National Institutes of Health, and by other support from the College of Education at MSU, the College of Human Medicine at MSU, and the Computing and Data Processing Center at Wayne State University. Specific contributors include the following:

CLIPIR/BMIS Project Staff

IRT/Management/Office Staff

The College of Education

The WSU Computing and Data Processing Staff

Consultants
BMIS TECHNICAL MANUAL CONTENTS OF VOLUMES

Volume 1.0. The Inquiry Theorectic Approach to Clinical Problem Solving Research and Application.

This volume discusses BMIS from the point of view of a researcher interested in an overall understanding of the system. It includes the theoretical basis of the system, short descriptions of the system components and results of applications of the theory and system in medicine and reading.

Volume 2.0. The Users Manual for the Basic Management Information System.

This volume is designed to help the educator or researcher make practical use of the system. It includes detailed descriptions of the BMIS system at WSU, the relevant command phrases, and many example.


This volume is designed to present the most technical level of documentation currently available on BMIS.
TABLE OF CONTENTS
FOR VOLUME 1.0

1.0 INTRODUCTION

2.0 THE THEORY
   2.1 Introduction
   2.2 The Clinical Encounter
   2.3 The Case
   2.4 The Clinician
   2.5 Clinical Performance Evaluation
   2.6 Clinical Instruction and Training

3.0 THE SYSTEM
   3.1 Introduction
   3.2 The SIMCASE Concept: Case Simulation
      3.2.1 Introduction
      3.2.2 Case Simulation Subsystem (CASS): Using a SIMCASE
      3.2.3 Case Creation Subsystem (CACS): Creating a SIMCASE
   3.3 The SIMCLIN Concept: Clinician Simulation
      3.3.1 Introduction
      3.3.2 CLINician Simulation Subsystem (CLSS): Using a SIMCLIN
      3.3.3 CLINician Creation Subsystem (CLCS): Creating a SIMCLIN
   3.4 The COMPUCEPTOR Concept: Computer Assistance for Clinical Performance and Training.
      3.4.1 Introduction
      3.4.2 PRReceptor Simulation Subsystem (PRSS): Using a Compuceptor
      3.4.3 PRReceptor Creation Subsystem (PRCS): Creating a Compuceptor
   3.5 The BIRS, RAPS, and STDS Subsystems: Other Aids for Clinical Performance and Training.
      3.5.1 Introduction
      3.5.2 BIRS, The Basic Information Retrieval System: A User Oriented Subsystem for Storage and Retrieval.
      3.5.4 STDS, The Set Theoretic Data System: A Programmer Oriented Utility for Efficient Storage and Retrieval.

4.0 APPLICATIONS IN MEDICINE
   4.1 Introduction
   4.2 SIMCASE Examples
   4.3 SIMCLIN Examples
   4.4 A Preliminary Study of SIMCLIN Performances

5.0 APPLICATIONS IN READING
   5.1 Introduction
   5.2 SIMCASE Development for Problems in Reading Diagnoses and Remediation
      Manually Based SIMCASES
      Computer Based SIMCASES
   5.3 Observational Studies of Clinical Problem Solving Behavior
      Procedures
      Results
5.4 SIMCLIN Development for Clinical Problem Solving Behavior Prediction
   Procedures
   Results
5.5 Simulation Studies of Clinical Problem Solving Behavior
   Procedures
   Results
6.0 CONCLUSIONS
7.0 REFERENCES
8.0 GLOSSARY OF TERMS
9.0 INDEX
1.0 Introduction

The purpose of the present volume of the BMIS Technical manual is to provide a synthesis of the many concepts, computer programs, and applications which comprise our past and current research on Clinical Problem Solving.

Specifically, we shall begin our discussion with a summary of the Inquiry Theory of Clinical Decision Making originally developed by Elstein, Shulman, et al. (1976); and formalized by the present authors in the BMIS Programs.

Next we shall examine in limited detail the system of programs which comprise the BMIS. We should consider, for each subsystem, method and examples of usage.

Finally, we shall summarize past and present applications of the system in education for clinical problem solving in medicine and reading.
2.0 THE THEORY

2.1 Introduction

Philologists seem to gain a special insight into the true nature of scholarly or scientific work. For example, Tabers (1973) medical and Websters unabridged both include two old, if not ancient, definitions which describe the subject of our theory and research with startling accuracy. Both definitions concern the same word root. "Clinical -- pertaining to medical practice based upon actual observation and treatment of patients, as distinguished from . . . laboratory studies"; and, "Clinical -- pertaining to the teaching of medicine by examining and treating patients in the presence of students".

Our problem concerns the psychology of clinical problem solving and clinical learning; the at times obscure processes by which clinicians diagnose and treat; and the manner by which they learn this remarkable trade.

The history of our study possibly begins with Nash (1954) and his diagnostic slide rule; a mechanical device for associating signs and symptoms with potential diagnostic categories. Certainly, Nash was one of the first to attempt to formulate a theoretical base for empirical observations of clinical problem solving.

Research on clinical problem solving has been actively pursued over the past two decades under a variety of names. One area of concentration has dealt with computer assisted diagnosis and treatment (Kleinmuntz, 1969). Another area concerns the use of patient simulations for training and evaluation in medical education (McGuire and Solomon, 1969). Our entry into the field was by way of a third approach: the direct study of the behavior of clinical problem solvers and the development of rigorous theory for the prediction of such behavior.

Most fundamentally, for the past five years we have been concerned with the development of a formal theoretic structure which could integrate the many concepts and empirical findings on clinical problem solving. We and our many colleagues have come to refer to this structure as the Inquiry Theory of Clinical Problem Solving.

The purpose of the present discussion is to summarize briefly the most recent version of the Inquiry Theory and the system of computer programs which support research and applications of this theory (The Basic Management Information System, BMIS).

2.2 The Clinical Encounter

Before we begin consideration of the concepts abstracted from the
Inquiry Theory, we must establish the basic definitions and parameters of our theory. The phenomenological base of our research is summarized in the figure below.

FIGURE 1. The Clinical Interaction

The clinical encounter comprises the behavioral domain addressed by the Inquiry Theory. The clinical encounter may be defined as the set of events occurring while a clinician (e.g., a physician, a reading clinician, a teacher, or a trouble shooter) attempts to solve some difficulty of a case (a patient or patient record, a client, a student, or a system), by making two fundamental decisions: The diagnosis (DX: What is the problem?; i.e., the current state of the case); and the treatment (RX: What action can be taken to solve the problem?; i.e., what will get the client/case into a more desirable state).

The result of the encounter is an interaction between clinician and case which occurs over a fixed (finite) period of time. The key behaviors usually observed in research on clinical problem solving are summarized in the figure.
The direction of interaction is indicated by arrows.

The Inquiry Theory, like many other approaches, attempts to predict only recurrent aspects of the clinical interaction—the features observed when several physicians interact with the same case, or a single physician interacts with several cases.

2.3 The Clinical Case

The first principle we must consider concerns the simulation of cases or patients. In order to achieve a scientifically acceptable level of objectivity in the study of the clinical encounter, case simulation is mandatory, since otherwise there is no basis for replication because no "real" case can be presented in exactly the same manner to more than one clinician.

2.3.1 The Inquiry Principle of Case Simulation

As illustrated in the figure below, one of the two basic tenets of the Inquiry Theory is that cases can be effectively simulated through the provision of sets of information on request by the clinician. Thus, for purposes of our theory, a CASE may be defined simply as a collection of:

1. Clinical problems defining the diagnostic and therapeutic task presented by the case;

2. The cue names and values defining the data base available for diagnostic and therapeutic decision making; and

3. The cue names and values defining changes in the data base to be expected from the passage of time and from the different treatments available to the clinician and the passage of time.

We should note that this conception concentrates upon the "intellectual" aspects of the encounter. The affective domain is largely excluded. It is not that we fail to appreciate the impact of emotions in proper clinical care, rather we all too well appreciate the unique advantage of humans in handling affective problems.
2.3.2 Related Research

Some workers in computer simulation of medical patients have emphasized the need for highly realistic natural language interaction between the student and the simulated case (Harless, et al., 1973a and 1973b). Indeed the "CASE" project expended great effort in the preparation of software which supports this interaction. The difficulties with this approach have been the same as those which ultimately crushed so many natural language R and D projects: costs and user acceptance. First, the production of medical cases tends to be extremely costly in terms of both human effort and money. Second, many people see little advantage in chatting with computers, especially when they must use a typewriter-teletype to do so (Friedman, 1973). Medical students especially seem to prefer more directed interactions which require less "interaction" and provide more guidance (DeDombal and Horrocks, et al., 1974).

Our concepts of patient simulation follow the arguments of Friedman
(1973) at Wisconsin and DeDombal, Horrocks, et al., (1974) at Leeds that significant features of clinical problem solving skills can be taught without completely simulating the natural language interaction. Like the Wisconsin system, we minimize required input by students, and maximize the ease of preparing cases not by a clinician but by R and D teams. We have intentionally attempted to extend the Wisconsin approach by eliminating from our simulations all but a few restrictions on the language of the interaction between a case and a clinician.

2.3.3 The SIMCASE Concept

The cognitive characteristics of a case may be viewed as a set of values associated with a list of cues, a set of problems that are reflected in altered cue values, and a set of responses to all possible treatments. Simulating a case can, then, be accomplished by simulating these individual elements. In BMIS, such a simulation is called a SIMCASE (SIMulated CASE). An interaction with a SIMCASE is illustrated below, where the BMIS computer is acting as the source of information about the patient. A SIMCASE is created by inputting cue names and values to the BMIS programs which combines them with "standard" information to generate a case.

```
#RUN WBSW:BMISX*OLDPIO 1=WBSW:FMEDIT 2=WBSW:FMEDDFT
#EXECUTION BEGINS

THE BASIC MANAGEMENT INFORMATION SYSTEM VERSION 1.1
THE INSTITUTE FOR RESEARCH ON TEACHING
MICHIGAN STATE UNIVERSITY
1976

?*cass

ENTERING THE CASE SIMULATION SUBSYSTEM

?*list cases

THE ALLOWED CASES ARE:

#SIMCASE,JAMES WALKER
JAMES WALKER IS A 47 YEAR OLD WHITE MALE WITH A COMMON PROBLEM IN INTERNAL MEDICINE. THIS CASE IS RELATED TO THE MSU FOCAL PROBLEMS FOR PHASE II, TERM 2 STUDENTS.

#SIMCASE,JENNIFER JONES
JENNIFER JONES IS A 51 YEAR OLD WHITE FEMALE WITH A PROBLEM IN INTERNAL MEDICINE. THIS CASE IS RELATED TO THE MSU FOCAL PROBLEMS FOR PHASE II, TERMS 2 STUDENTS.

?*list vocabulary
```
THE ALLOWED KEYWORDS ARE:

ABDOMEN
ADULT INFECTIONOUS DISEASES
ALLERGIES

VITAL SIGNS
WBC

BEGINNING CASE SIMULATION

THE ENCOUNTER WITH THE CASE BEGINS:
THE PATIENT WAS REFERRED BY ANOTHER PHYSICIAN FOR GASTROINTESTINAL PROBLEMS.

ENTER KEYWORD
?general appearance
THE PATIENT IS A MARKEDLY OBESE, MIDDLE AGED, WHITE FEMALE. SHE HAS SMALL VESICULAR PAPULES OVER THE FACE AND UPPER THORAX. THE NAILS ARE THIN AND YELOWISH.

ENTER KEYWORD
?abdomen
STRIKING HEPATOSPLENOMEGALY. VIRTUALLY ALL OF THE LIVER WAS BELOW THE COSTAL MARGIN AND WAS MEASURED ALMOST 18 CM IN THE MID CLAVICULAR LINE AND 15 CM IN THE MID LINE. THE LIVER WAS FIRM. THE SPLEEN WAS 6-8 CM BELOW THE COSTAL MARGIN RATHER MEDIANALLY AND QUITE FIRM. THERE WERE NO DILATED VENOUS COLLATERALS OF THE ABDOMEN WALL.

ENTER KEYWORD
?vital signs
THE PATIENT IS 5 FEET 3 INCHES TALL AND WEIGHS 150 POUNDS. HER ORAL TEMPERATURE IS 98.6. PULSE IS 75 AND REGULAR. RESPIRATION RATE IS 22/MINUTE. BLOOD PRESSURE IS 180/100 RIGHT ARM RESTING.

ENTER KEYWORD
?liver biopsy
A LIVER BIOPSY OF THE PATIENT REVEALED A FATTY CHANGE, BRIDGING, ADVANCED CIRRHOSIS, LIMITING PLATE DISRUPTION, FIBROSIS AND FOCAL NECROSIS.

ENTER KEYWORD
?*$diagnose
2.4 The Clinician

The second principle of the Inquiry Theory concerns the major factors which govern the behavior of a clinician in a clinical encounter.

2.4.1 The Inquiry Principle of Clinician Simulation

As illustrated in the figure below, we postulate that the major events in the clinical interaction are determined probabilistically by the CASE and the CLINICIAN’s MEMORY and STRATEGY. The clinician’s memory (Clinical Memory) consists of problem, cue, prescription, and treatment descriptions. It further contains a number of important relations including: (1) a set of relations between cue and problem descriptions that are used to infer the presence of problems in a given case based on the cues already collected; (2) a set of relations between problem and cue descriptions that are used to determine which cues should be collected next to confirm or disconfirm the hypotheses currently under consideration; (3) a set of relations between problems and treatments which are used to evaluate and select treatment plans for a given case and diagnosis; and (4) a set of relations between treatments and prescriptions that are used to define the specifics of case management for a particular case and treatment plan. The clinician’s strategy (Clinical Strategy) consists of a general approach to the basic information gathering and information processing tasks required to make diagnostic and treatment decisions, i.e., a planned sequence of information processing tasks using the clinical memory to make decisions about diagnosis and treatment. Our set of information processing tasks was derived from empirical studies of clinical problem solving with simulated cases (Elstein, et al., 1976). The tasks include:

(1) Cue Acquisition:
This process consists of deciding which cues should be collected during the clinical encounter and the collection of cue values for those selected. For a physician this would include the cues of the medical history, physical examinations and laboratory work-up. Cues are usually selected on one of two bases: first, as a means of confirming or disconfirming one or more competing hypotheses about the case’s problem; or second, according to some information gathering heuristic,
FIGURE 3. The Inquiry Theory Conceptualization of a Clinician

CLINICAL MEMORY
- Set of Problems (Q)
- Set of Cue Values (C)
- Relations (R(C,P))
- Set of Treatments (R)
- Set of Relations (R(P,Rx))

CLINICAL STRATEGY
- Sequence of Information Processing
- Actions, Including:
  - Hypoth. Generation
  - Cue Collection
  - Hypoth. Evaluation
  - Dx Judgment
  - Rx Evaluation

CLINICIAN

CLINICAL INTERACTION

CASE

e.g., according to a work-up "routine" by body systems.

(2) Hypothesis Generation
This process consists of retrieving from memory a number of problem formulations (hypotheses) based on the cues already collected. Hypotheses generated early are used in part to guide the subsequent work-up. The set of problem formulations retrieved is based upon the relations between cues and problems, R(C,P), that are part of the clinician's memory.

(3) Cue Interpretation:
Cues are evaluated in terms of fit to specific hypotheses.

(4) Hypothesis Evaluation and Diagnosis Judgement:
This process consists of the estimation of the likelihood of each hypothesis under consideration. Two separate levels may be distinguished. First, unlikely hypotheses are eliminated from the set
under consideration. Second, one or more of the hypotheses may be
accepted as the diagnosis if the likelihood of those hypotheses reaches
a preset cutoff point. The calculation of the likelihood of any
hypothesis is based upon the relations between problems and cues \( R(P, C) \)
which are stored in the clinician's clinical memory. Several methods
of calculating the likelihood of hypotheses have been successfully used
including: Bayesian probability, multiple regression, and simple
algebraic summation of weights (Slovic and Lichtenstein, 1971; Dawes
and Corrigan, 1974).

(5) Treatment Evaluation:
This process consists of the estimation of the expected gain for each
of the available treatments given the patient's diagnosis. Expected
gain is calculated on the basis of the relations between the problem
included in the diagnosis and available treatment plans, \( R(P, T) \). These
relations may include precautions and contraindications appropriate to
the patient or case.

(6) Prescription Selection:
After the treatment plan is selected, the relations between treatments
and prescriptions \( R(T, Rx) \) are used to write out the specifics of the
case management.

Clinical memories have been developed in a variety of ways but all
approaches seem to be reducible to the same conceptualization, which
postulates a memory storage element (e.g., record or list) for each problem
in the clinician's memory, for each cue (symptom or sign or test) in the
clinician's repertoire, and for each treatment or prescription used by the
clinician. We conceive of problem representations as associating problems
with cue names (i.e., given a set of symptoms, what are the most likely
problems? And given a set of possible problems, what cues should be
collected?); cue representations as relating cue names to cue interpretation
and collection routines; and treatment representations as associating
treatments with problems. The present version of our Inquiry Theory computer
programs also have the capability of altering an overall strategy and
switching to a "sub-strategy" upon the generation of any particular
hypothesis (e.g., a procedure for checking out neoplasms when an unexpected
abdominal mass is discovered while performing a routine abdominal
examination.)

2.4.2 Related Research

The representation of a clinical encounter as an interaction between a
case and a clinician's memory and strategy derives from empirical studies of
clinical problem solving. By conceptualizing CLINICAL MEMORY and STRATEGY
from an information processing perspective, as we have, several features of
this reasoning may be seen to be particular instances of more general
principles of the psychology of thinking. The principles of special
importance to our work are as follows.

Use of a Problem Space:
Problem solving is characterized (1) by the representation of a complex problem by a simplified "problem space" (Newell and Simon, 1972); (2) by the "structuring" of the problem space to determine the information processing activities to be used in solving the problem; and (3) by the accuracy of problem representation and the effectiveness of the actions taken to achieve a goal determining success.

The Hypothetico-Deductive Method:
In clinical problem solving, the vast complexity of the diagnostic problem is accommodated by (1) the early generation of tentative diagnostic hypotheses used by the clinician to structure the "problem space" within manageable limits; and (2) the use of the hypothetico-deductive method to select the cues to be collected in the clinical encounter (Eistean, et al., 1976).

Strategies for Handling Complexity:
Given the limitations on the number of hypotheses and cues that can be retained in "immediate memory," the clinical problem solver tends to use strategies for reducing the complexity of problem space.

(1) Linear Problem Solving: testing one or two hypotheses at a time -- thus reducing the size of the problem space or making it possible to find space for new hypotheses or more adequate reformulations.

(2) Problem Space Reorganization: revising the problem domain by "chunking" or revising the diagnostic concepts. Methods observed include: reformulation of the problem; conversion to a set of problems related to predecessors; beginning with general problem and proceeding to specific; beginning with specific complaints and proceeding to general problems.

Hypotheses Generation and Testing:
Hypothesis generation is characterized by (1) initial problem formations (hypotheses) generated by associations of a single cue to a hypothesis, or a set of cues to a hypotheses, or a hypothesis to a hypotheses; (selective, systematic cue collection is more closely associated with hypothesis evaluation than with generation); and (2) the number of hypotheses tested simultaneously being usually 4-5 with a maximum of 6-7 hypotheses.
Possible memory structures for these problem formulations include lists of cue findings stored under diagnostic labels. One reason why physicians are "inconsistent" across cases may be that the adequacy of these lists for different hypotheses (problems) varies within a single physician as well as between physicians.
According to studies at McMaster University (Neufeld, et al., 1976 and Barrows, et al., 1976), hypothesis generation and testing behavior may be further characterized by early generation of hypotheses (55% of the hypotheses were considered in the first 12% of the encounter, 78% in the first 25% of the encounter) and early elimination of hypotheses.
(There is very little modification of hypothesis pool in last 50% of the encounter, 94% of hypotheses considered in the last 50% of the
encounter are retained).

Cue Interpretation:
The Inquiry and McMaster studies indicate cue interpretation is characterized by
(1) Subjective weighting of cues for given hypotheses. Cues largely seem to be weighted on a three point scale tending to confirm or disconfirm a hypotheses, or as non contributory. The scheme can be shown to be a simplified approximation of Bayes Theorem.
(2) Accuracy of cue interpretation is directly related to accuracy of Dx outcome and independent of thoroughness of cue acquisition.

Cue Collection:
According to the same McMaster studies, (1) most of the history (68%) is elicited in first two quarters of the encounter; (2) most of the physical is elicited in third quarter of the encounter; and (3) early acquisition of findings is documented as follows: 50% of the total findings are elicited in the first 25% of the encounter, 56% of the supercritical and 53% of the critical findings are elicited in the first 25% of the encounter.

2.5 Clinical Performance Evaluation

Thus far we have developed ideas regarding the structure of clinical problems and the information processing of clinicians. Let us now consider some implications of these notions for the practical problem of developing more effective clinicians. To address the problem of providing more effective clinical problem solving, we must first consider the implications of the Inquiry Theory for performance improvement in clinical problem solving.

2.5.1 The Performance Evaluation Corollary

The Inquiry Theory asserts that the clinical interaction, whether simulated or real, is a probabilistic function of the case and the clinician's memory and strategy. From this assumption, we may suggest the following deduction.

If the clinician's portion of the clinical interaction is determined by clinical memory and strategy, it follows that performance evaluation based upon the interaction must be determined by clinical memory and strategy. Hence, the Performance Corollary: Assuming there exists a reliable, valid performance evaluation procedure, the performance value produced by this procedure for a clinician working with a particular case is probabilistically determined by the CASE and the CLINICIAN's MEMORY and STRATEGY. By valid evaluations, we mean evaluations requiring (1) accurate, non-ambiguous case records, (2) non-ambiguous criteria for measuring clinical performance, and (3) clinical preceptors having special competence for the particular case.
2.5.2 Related Research

The Performance Corollary has a certain a priori validity about it. If valid measures of diagnostic and therapeutic performance can be demonstrated, they must be defined upon the clinical interaction since this is the behavioral domain in which the decisions of diagnosis and therapy are made. The question is whether valid measures can be obtained for diagnostic and
therapeutic effectiveness. A number of studies have significantly affected our thinking.

The MSU Inquiry Studies: diagnostic success on some medical problems is not highly correlated with success on other problems, i.e., diagnostic accuracy is case specific (Elstein, et al., 1976); diagnostic performance is related to the amount of experience (during and after medical school) but to little else.

The CBX Studies (Hoffman, 1974): Performance criteria do not readily distinguish among clinicians. Diagnostic success is problem dependent (or at least "knowledge domain" dependent).

The McMaster Findings (Neufeld, et al., 1976; Barrows, et al., 1976): Quantitative studies of clinical performance seem to indicate that only a few criteria are sufficient. The most important criteria for the description of clinical performance are (1) diagnostic commonality; (2) thoroughness; and (3) efficiency. Clinical performance seems to improve as a function of experience under under conditions of performance feedback, when performance is measured by commonality or agreement with "absolute" criteria, e.g., pathologic findings, reaction to treatment, etc.

2.6 Clinical Instruction and Training

Given our brief discussion of clinical problem solving performance evaluation, we may now turn to the more significant problem of how to improve such performance through education and other modes of problem solving support.

The Inquiry Theory has some very specific implications on how to improve clinical performance. We have summarized these implications in what we refer to as the Instructional Corollary.

2.6.1 The Instructional Corollary

If, as has been stated in the performance evaluation corollary, clinical performance is probabilistically determined by the clinician's memory and strategy, it follows that clinical performance may be improvements by improvements in clinical memory and strategy. Hence, the Instruction Corollary: clinical problem solving performance may be improved by (1) providing performance aids to augment clinical memory and strategy; (2) providing training and feedback aids for the acquisition of more accurate and effective clinical memories and strategies; and (3) providing combinations of training and decision aids.

2.6.2 Related Research
A great deal of research seems to offer some degree of support and explication for the Instructional Corollary. To begin with, let us consider studies attempting to improve diagnostic performance by the use of computer aided diagnostic procedures. Two research programs were of prime importance to our work since they indicate the practicality of using computers in the real world of a clinic.

The first set of studies was conducted in the Leeds Medical Center. In their most dramatic study (DeDombal, 1973; DeDombal and Horrocks, 1974), a real-time controlled trial was carried out on a consecutive unselected series of 552 patients. This trial included all patients who came to the Department of Surgery between January 1971 and August 1972 with abdominal pain of less than one week's duration. The pre-operative diagnosis of the most senior clinician who saw each case was accurate in some 87% of the patients studied. Using the same information (and producing its "diagnosis" before the operation), the computer-aided system proved to be accurate in 98% of the cases. In a subsequent study, the Leeds group found that training with the computer aided diagnostic system significantly improved students' clinical
performance as well.

The second set of studies in computer aided diagnostic concerns is the MEDITEL system (Barnes, et al., 1974). The data bank for the system includes 1500 diagnoses in pediatrics. Since many of these have diagnostic names that are categorical, the data bank encompasses nearly all pediatric diagnoses of systemic illness. There are approximately 650 abnormal symptoms, physical findings, and laboratory results in the MEDITEL vocabulary. Of these, 75 are historical, 375 are physical findings (including 60 roentgenographic findings), and 200 are laboratory findings. All of the vocabulary consists of commonly used medical terminology that has been selected for the potential to distinguish one diagnosis from another.

Computer-assisted diagnosis for a given patient is obtained by selecting from the list of abnormal symptoms, physical findings, and laboratory results those noted at the time of the initial history and physical examination. These findings are sent via terminal to a remote time-sharing computer over telephone lines. The computer systematically searches its data bank, comparing the patient's findings with each diagnosis, and selects the diagnoses that should be included in a complete differential diagnosis. These are listed by the teletypewriter terminal.

Perhaps the most impressive evaluation of MEDITEL was by Swender (1974). In this investigation, MEDITEL was evaluated for its ability to consider a correct diagnosis in cases in which the correct diagnosis was not considered in the original differential diagnosis by physicians involved in the patient's care. The computer provided the correct diagnosis in 83% of the 33 cases studied when the information was entered by a physician and in 65% of the cases when the information was provided by a medical secretary.

In summary, research on such computer aids to memory and strategy reaches the following generalizations: Given subjective weights provided by physicians, the accuracy of computer assisted diagnosis is variable, sometimes equalling physicians, but often less effective (Leaper, et al., 1972). Given empirical weights, computers can diagnose more accurately than senior clinicians. However, computer diagnosis has been demonstrated only in limited areas of practice/decision making -- e.g., abdominal surgery. Finally, student clinicians can benefit from interaction with computerized diagnostic mentors.

As a second approach, let us consider attempts to improve diagnostic or therapeutic performance by non-computer external aids to memory and strategy. Many studies indicate the effectiveness of decision aids in the form of "books" of decision making charts -- not entirely similar from the familiar MERCK MANUAL (Honey, 1972). For example, Essex (1975) reported that medical students were able to improve their performance from 70% agreement to 98% agreement with senior clinicians -- if they were provided with "decision flow charts." Further, while the physicians required an average of 13.7 minutes per case -- the students required only 1.9 minutes.
Basically, the results on external memory and strategy aids seem to follow these generalizations: decision making aids to clinical memory and strategy can significantly improve the performance of medical students or para-professionals on specific problems for which aids have been developed (Grimm, Harlan, Estes, and Shimorii, 1975). Under some circumstances, "printed" decision charts can help para-professionals function nearly as effectively as physicians and with greater efficiency.

To conclude these brief research notes, we offer some hypotheses on the nature of clinical learning. Diagnostic learning seems to consist of the following:

1) Development of improved problem descriptions by means of (a) description extension (adding cues to the problem description) and (b) description validation (establishing more valid relations between cues and problems).

2) Development of improved cue descriptions by means of (a) extension (adding cues and cue characteristics) and (b) cue validation (improving the recognition of cue values).

3) Development of "strategy" structures to control the diagnostic process for individual cases.

4) Improvements in memory and strategy require experience with cases under conditions of feedback on performance in Dx and Rx.
3.0 THE SYSTEM

3.1 Introduction

The Basic Management Information System (BMIS) is formed about the structure of the Inquiry Theory discussed in the previous section. Thus, corresponding to each theoretic concept, we find one or more BMIS subsystems. Perhaps the most useful introduction would be a brief review of these subsystems.

CACS -- The CAse Creation Subsystem is designed to assist with the preparation of computer based SIMCASES.

CASS -- The CAse Simulation Subsystem is designed to execute a SIMCASE for performance evaluation or training.

CLCS -- The CLinician Creation Subsystem assists with the preparation of a computer based SIMCLIN.

CLSS -- The CLinician Simulation Subsystem permits the user to execute a SIMCLIN on a simulated or real case.

PRCS -- The PReceptor Creation Subsystem is designed to assist with the creation of a COMPUCEPTOR; e.g., a simulated clinical preceptor to provide CAI support for training in clinical problem solving.

PRSS -- The PReceptor Simulation Subsystem permits the execution of the COMPUCEPTOR for a given case, and clinical environment to provide CAI support for the student clinician.

BIRS -- The Basic Information Retrieval System is a general purpose information handling utility which is used to support the other subsystems and to produce materials for the manually based SIMCASES, non-computer based decision aids, etc.

RAPS -- The Report Analysis Programming System is a general purpose utility used with BIRS to prepare analyses for SIMCASE and SIMCLIN development, etc.

STDS* -- The Set Theoretic Data System is a proprietary program used to perform all needed random access retrieval.

DRILLS* -- The DRILL and Practice Subsystem is designed to provide specific training on the clinical memory required for successful clinical problem solving. DRILL provides for the input of stimulus frames, anticipated responses with feedback, and overall performance evaluation scores.
Except for STDS, all subsystems of BMIS are written in USASI standard FORTRAN. Most of the programs are compatible with IBM 360 and IBM 370 hardware. BMIS presently operates on the AMDAHL/IBM system at Wayne State University using the Michigan Terminal System. To continue our system review let us consider each of these BMIS subsystems in more detail.

3.2 The SIMCASE Concept: Case, Client or Patient Simulation

3.2.1 Introduction

The SIMulated CASE or SIMCASE is a computer generated and/or presented set of materials which is intended to elicit problem solving behavior from students and practicing clinicians. These forms of SIMCASEs supported by BMIS are the The MBSC (Manually Based Simulated Case); the CBSC (Computer Based Simulated Case); and the ASC (Automated Simcase).

The MBSC is designed for manual presentation by experimentors to clinical problem solvers or for use as a self teaching aid by students. The CBSC is designed for computer presentation to clinical problem solvers or students by computer terminals or low cost mini computers. The ASC is designed to permit SIMCLINS (Simulated Clinicians) to directly interact with a SIMCASE without human intervention.

In the following we shall primarily consider the CBSC which is the form of the SIMCASE that appears to have the greatest practical potential. The MBSC and the ASC will be considered later on in discussions of applications.

The basic concept of a SIMCASE is summarized in Figure 6. As Figure 6 shows, the SIMCASE is a physical realization of the Inquiry Theory conceptualization of a case.

Essentially, the SIMCASE concept is that the cognitive characteristics of a case may be viewed as a set of values associated with a list of cues, a set of problems that are reflected in cue values altered from their "normal" state, and a set of responses to all possible treatments and to the passage of time. Simulating a case can, then, be accomplished by simulating these individual elements. In BMIS, such a simulation is called a SIMCASE (SIMulated CASE). An interaction with a SIMCASE is illustrated in the next section where the BMIS computer is acting as the source of information about the case. A SIMCASE is created by inputting cue names and values to the BMIS programs which combines them with "standard" information to generate a case. The creation of a SIMCASE will be described later.

* These programs are presently either not available in BMIS or are available in experimental versions.
3.2.2 The CASE SIMulation Subsystem (CASS)

The CASS Subsystem is designed to allow a student or practicing clinician to interact with a previously prepared SIMCASE. Interaction takes place through a previously defined language hereafter referred to as the CAI language. The CAI language consists of the set of terms, phrases, or sentences to which the SIMCASE has been prepared to respond. The CAI language is entirely the choice of the SIMCASE author. Either complex, natural language or highly controlled acronyms may be used. For each element in the CAI language, a response has been stored by means of a special BMIS data record. When the clinical problem solver inputs the CAI language
element, the data record is output. For example, in the reading SIMCASE "Stephen," the CAI language consists of a set of keywords identifying particular data; thus the keyword "DOL01" causes the SIMCASE to respond with test scores for the Dolch Sight word test.

In general, we have used both "natural" and highly structured CAI languages. Presently, the structured CAI languages are being used in most of our work in order to provide better user acceptance by curtailing problems of non-recognition due to poor user typing skills and a lack of standard vocabularies in reading diagnoses. In any case, CASS provides, at user request, a summary of the CAI vocabulary.

The listing given below illustrates the use of the CASS subsystem in a typical problem. Annotations are provided to explain the interaction. The interaction is preceded by a listing of the CAI vocabulary. It should be noted that provision is made for non-computer based data on the case, through a SIMCASE STUDY GUIDE. The guide includes written materials, test booklets, pictorial materials, and audio tape recordings made by the client.

```plaintext
#$RUN WBSX:BMIS+*OLDPIO 1=WBSX:STEPHENF 2=WBSX:STEPHENFDT
#EXECUTION BEGINS

THE BASIC MANAGEMENT INFORMATION SYSTEM VERSION 1.2
THE INSTITUTE FOR RESEARCH ON TEACHING
MICHIGAN STATE UNIVERSITY
1977

ENTERING THE CASE SIMULATION SUBSYSTEM
YOU ARE GOING NOW TO RECEIVE TWO SETS OF INSTRUCTIONS:
1. INSTRUCTIONS ON TERMINAL USAGE;
2. INSTRUCTIONS ON HOW TO USE A COMPUTER BASED SIMULATED CASE.
AS YOU GO THROUGH THESE INSTRUCTIONS YOU CAN, AFTER EACH INTERACTION
WITH THE COMPUTER, ASK FOR ANY ONE OF THE FOLLOWING PROCEDURES:
1) STOP (END) THE INSTRUCTIONS - BY TYPING "STOP";
2) RESTART THE INSTRUCTIONS - BY TYPING "RESTART";
3) BACKUP TO THE PREVIOUS SECTION OF THE INSTRUCTIONS - BY TYPING "BACKUP".
WHENEVER YOU HAVE TYPED SOMETHING INTO THE COMPUTER AND YOU ARE NOT SURE
WHAT TO DO NEXT (OR YOU DO NOT KNOW WHAT TO TYPE), TYPE IN SOME GARBAGE
AND HIT "RETURN" KEY; THE COMPUTER WILL TELL YOU WHAT TO DO NEXT.
DO YOU NEED ANY INFORMATION ON HOW TO USE YOUR TERMINAL?
?YES
DO YOU NEED INFORMATION ON WHAT A TERMINAL IS?
?YES
YOU ARE USING A COMPUTER TERMINAL. IT IS A COMMUNICATION DEVICE BETWEEN
YOU AND THE COMPUTER. THE COMPUTER COMMUNICATES WITH YOU BY TYPING LINES
OF INFORMATION AS IT IS RIGHT NOW. YOU, IN TURN, CAN SEND MESSAGES,
COMMANDS, ETC., TO THE COMPUTER BY TYPING THE INFORMATION ON THE KEYBOARD
```
AND THEN HITTING THE 'CARRIAGE RETURN' BUTTON. SPECIFICALLY, TO SEND
INFORMATION TO THE COMPUTER, YOU MUST WAIT FOR THE COMPUTER TO PRINT A
PROMPTING CHARACTER (USUALLY A '?' BUT SOMETIMES A ':' OR '#'), AND THEN
TYPE IN ONE LINE OF DATA, ENDING THE LINE BY PUSHING THE 'CARRIAGE RETURN'
BUTTON. TO SUMMARIZE, WAIT FOR PROMPTING, TYPE IN A LINE, PUSH 'CARRIAGE
RETURN'.
IF READY TO CONTINUE TYPE "GO". IF NOT SEE CLIPIR INSTRUCTOR.
?STOP
WHEN READY TO CONTINUE ENTER THE COMMAND 'SENDFILE'
?SENDFILE
DO YOU NEED ANY INFORMATION ON HOW TO USE A COMPUTER BASED SIMULATED CASE?
(COMPUTER BASED SIMULATED CASE = CBSC)
A: NO- SKIP THIS EXPLANATION
B: ONLY A SHORT SUMMARY
C: YES- A FULL DESCRIPTION
?SUMMARY
THE ALLOWED RESPONSES ARE:
A
B
C
STOP
RESTART
BACKUP
TRY AGAIN. ENTER RESPONSE.
?A
WHEN READY TO CONTINUE ENTER THE COMMAND 'SENDFILE'
?SENDFILE
?SENDFILE
?*LIST VOCABULARY
THE ALLOWED KEYWORDS ARE DESCRIBED IN THE FOLLOWING INVENTORY:
#SIMCASE, STEPHEN
#BLOCK,

****************************************************
CASE INFORMATION INVENTORY
****************************************************

STEPHEN

**********

FORMS OF INFORMATION
-----------------------

TEST
TEST
BOOKLET
DIRECTIONS

EXAMINER'S
COMMENTS

AUDIO
RECORDING
OF SESSION

****************************************************
### DEMONSTRATION ITEMS

<table>
<thead>
<tr>
<th>DEMONSTRATION DATA</th>
<th>DEM1</th>
<th>DEM2</th>
<th>DEM3</th>
<th>DEM4</th>
<th>DEM5</th>
<th>DEM6</th>
</tr>
</thead>
<tbody>
<tr>
<td>BACKGROUND INFORMATION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOGRAPHICAL DATA</td>
<td>BKG2</td>
<td></td>
<td></td>
<td>BKG6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHYSICAL/HEALTH</td>
<td>BKG8</td>
<td></td>
<td></td>
<td>BKG6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HOME/FAMILY</td>
<td>BKG14</td>
<td></td>
<td></td>
<td>BKG6</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLASSROOM INFORMATION</td>
<td>BKG20</td>
<td></td>
<td></td>
<td>BKG6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### STANDARDIZED ASSESSMENT

| ACHIEVEMENT TEST-INDIVIDUAL (PEABODY) | | | | | | |
| READING RECOGNITION | PEAl | PEa2 | PEa3 | PEa5 | PEa6 |
| READING COMPREHENSION | PEa7 | PEa8 | PEa9 | PEa5 | PEa6 |
| SPELLING | PEa13 | PEa14 | PEa15 | PEa5 | PEa6 |
| GENERAL INFORMATION | PEa19 | PEa21 | | PEa5 | PEa6 |
| ACHIEVEMENT TEST-GROUP (GATES-MACGINNITIE) | | | | | | |
| VOCABULARY | GMG1 | GMG2 | GMG3 | | GMG6 |
| COMPREHENSION | GMG7 | GMG8 | GMG9 | | GMG6 |
| SPEED AND ACCURACY | GMG13 | | GMG15 | | GMG6 |
| DIAGNOSTIC TEST OF READING DIFFICULTY-(GATES-MCKILLOP) | | | | | | |
| RECOGNIZING AND BLENDING COMMON WORD PARTS | GMK1 | GMK2 | GMK3 | GMK4 | GMK5 | GMK6 |
| DIAGNOSTIC TEST OF READING DIFFICULTY (DURRELL) | | | | | | |
| ORAL READING | DUR1 | DUR2 | DUR3 | DUR4 | DUR5 | DUR6 |
| SILENT READING | DUR7 | DUR8 | DUR9 | DUR11 | DUR6 |
| LISTENING | DUR13 | DUR14 | DUR15 | DUR17 | DUR6 |
| COMPREHENSION | | | | | | |
| WORD RECOGNITION AND ANALYSIS | DUR19 | DUR20 | DUR21 | DUR22 | DUR23 | DUR6 |
| VISUAL MEMORY OF WORDS-PRIMARY | DUR25 | DUR26 | DUR27 | DUR29 | DUR6 |
| HEARING SOUNDS IN WORDS | DUR31 | DUR32 | DUR33 | DUR35 | DUR6 |
| SOUNDS OF LETTERS | DUR37 | DUR38 | DUR39 | | DUR6 |
| GRADED WORD LIST | | | | | | |
BEGINNING CASE SIMULATION
THE ENCOUNTER WITH THE CASE BEGINS:

OBSERVATIONS/COMMENTS:
ACCORDING TO THE TEACHER, STEPHEN IS HAVING DIFFICULTIES WITH
READING. PROGRESS IN ALL OTHER AREAS APPEARS SATISFACTORY EXCEPT
THOSE REQUIRING READING. STEPHEN'S TEACHER AND HIS PARENTS ARE
CONCERNED AND HAVE REQUESTED A READING CONSULTATION.

TEST DATA:
GENERALLY, SCHOOL TEST RESULTS INDICATE HE IS NOT PERFORMING AT
HIS PROJECTED LEVEL IN MOST READING RELATED ACTIVITIES.
BEFORE YOU CONTINUE ON LOOK AT STEPHEN'S PICTURE IN THE STUDY GUIDE
AND LISTEN TO INITIAL CONTACT TAPE.

ENTER KEYWORD
?BKG2

DATA:
NAME - STEPHEN
AGE - 8 YEARS 6 MONTHS
ADDRESS - BURCHAM DR.
  E. LANSING, MICH.
SCHOOL - MARBLE
DISTRICT - EAST LANSING
GRADE - 3RD

ENTER KEYWORD
?BKG8

DATA:
HEIGHT - AVERAGE IN FIRST AND THIRD GRADES
WEIGHT - AVERAGE IN FIRST AND THIRD GRADES
VISION SCREENING - NORMAL IN FIRST AND THIRD GRADES
HEARING SCREENING - NORMAL IN FIRST AND THIRD GRADES
GENERAL HEALTH - AVERAGE, NO PHYSICAL DISORDERS
ORAL LANGUAGE PRODUCTION - NO SPEECH THERAPY
NO DEFICIENCY

LANGUAGE DEVELOPMENT - NORMAL

EMOTIONAL DIFFICULTIES - TEACHER REPORTS:
'TENDS TO BE AGGRESSIVE VERBALLY AND
SOMETIMES PHYSICALLY IN PEER RELA-
TIONSHPES. SHOWS IMPATIENCE WITH
READING RELATED TASKS.'

ENTER KEYWORD

?BKG20

DATA: CLASSROOM BEHAVIOR: TEACHER REPORTS THAT HE ATTRACTS ATTENTION
TO SELF INAPPROPRIATELY BUT DOES NOT RESIST
CONTROL STATEMENTS.

GENERALLY COOPERATIVE IN NON-READ-
RELATED SUBJECTS, E.G., MATHMATICS.
DISPLAYS IMPATIENCE WITH READING RELATED
TASKS.

PEER RELATIONSHIPS: TEACHER REPORTS THAT SUBJECT TENDS TO BE
AGGRESSIVE VERBALLY AND SOMETIMES PHYSI-
CALLY. HOWEVER THE SUBJECT IS NOT ISOLATED
BY HIS PEERS.

ATTENDANCE PATTERNS - NORMAL.
ATTENDING BEHAVIOR - GOOD.
REFERRED TO COUNSELOR - NO
REFERRED TO SOCIAL WORKER - NO
REFERRED TO READING TEACHER - NO
REFERRED TO READING CLINIC - BY REQUEST OF PARENTS AND TEACHER IN
OCTOBER OF THIRD GRADE YEAR

DATE OF INTERVIEW AND TESTING - NOVEMBER OF THIRD GRADE YEAR

READING SERIES USED IN CLASS - GINN 360

ORGANIZATIONAL SYSTEM - THREE GROUPS WITH INDIVIDUAL
PHONICS INSTRUCTION.

READING GROUP - LEVEL 2-1

ENTER KEYWORD

?DUR3

SEE: STEPHEN SIMCASE STUDY GUIDE
DURRELL ANALYSIS OF READING DIFFICULTY: ORAL READING
(ITEM DUR3)

ENTER KEYWORD

?DUR4

SEE: STEPHEN SIMCASE STUDY GUIDE
DURRELL ORAL READING AUDIO RECORDING
(CASSETTE DUR4)

ENTER KEYWORD

?WISC2

EXAMINER'S COMMENTS: (THIS TEST WAS ADMINISTERED BY THE SCHOOL
PSYCHOLOGIST APPROXIMATELY 9 MONTHS AGO. HIS
COMMENTS AND CONCLUSIONS FOLLOW.)

- ABOVE AVERAGE INTELLIGENCE WITH EXTREMELY
  GOOD PRACTICAL KNOWLEDGE AND SOCIAL
JUDGMENT,
- REFLECTS STRENGTH IN MAKING USE OF INFORMATION IN HIS ENVIRONMENT.
- PERFORMED SLIGHTLY LOWER ON TASKS WHICH REQUIRED CONCENTRATION OR WERE TIMED.
- MAY INDICATE THAT HE NEEDS MORE TIME TO PROCESS INFORMATION AND/OR THAT HIS CONCENTRATION IS HINDERED WHEN HE FEELS PRESSURED TO PERFORM.

ENTER KEYWORD
?DOL3
SEE: STEPHEN SIMCASE STUDY GUIDE
DOLCH WORD LIST ANNOTATED TEST BOOKLET
(ITEM DOL3)

ENTER KEYWORD
?DOL6
ENTER KEYWORD
?*DIAGNOSE
ENTER YOUR DIAGNOSES (END LIST WITH BLANK LINE)
?SIGHT WORD PROBLEMS
? POTENTIAL ADEQUATE
?PSYCHOLOGICAL MAKEUP NOT PROBLEMATIC
?
END OF CASE SIMULATION
?*$FINISH
#EXECUTION TERMINATED

3.2.3 The Case Creation Subsystem (CACS)

The CASECRE subsystem is designed to support the preparation of manually based and computer based SIMCASEs. For both types of SIMCASE development three major types of data input are required.

First, a CAI vocabulary must be defined which includes all recognized CUE NAMES. A cue, it will be recalled, is the informational element central to the clinical interaction, i.e., that elemental "chunk" of information provided by the case to each request for information from the clinician.

Second, a CAI data record must be input. The data record specifies the information or "cue value" to be presented upon receipt of each CUENAME.

Third, additional data which cannot be stored in the computer must be collected and referenced by the CAI vocabulary. These data include actual samples of the client's or patient's physical output, e.g., test booklets, laboratory test results, oral reading recorded on audio tape, etc. This process of case creation is summarized in Figure 7.

To complete our introduction to the CASECRE subsystem, let us briefly review the BMIS CAI data record used to create SIMCASES. As shown below the CAI data record includes "fields" for CUE NAMES and CUE VALUES. A field is
FIGURE 7. THE RESUME OF THE INPUT AND OUTPUT ASSOCIATED WITH CASE SIMULATION

CAI INSTRUCTIONAL TEAM

CAI Vocabulary
Names of clinical signs and symptoms

CHIEF Complaint
Dolch list Test
Durrell Oral Test

'Smith Brown
'Chief complaint
10-year old 4th grade boy reporting
difficulty in textbooks assignments
'Dolch list

SIMCASE DATA RECORDS
Simulated clients
data records stored
with labels from
CAI Vocabulary

BMIS

STUDENT CLINICIANS

*SCASS
What case would you like to see?
Eric Brown
Past Standardized Tests

BMIS

SIMCASES
analogous to a paragraph or subsection devoted to a particular purpose. Thus the REFER field contains the CAI vocabulary entry for the particular cue in question and the BLOCK field contains the cue value to be output.

The example CAI data record given below should clarify the task of creating a SIMCASE. The example includes explanatory annotations. A complete discussion of the use of BMIS to prepare SIMCASE's is given in Volume 2.0 of the present manual.

*§ABSTRACT
#RTYPE,CAI
#SIMCASE,STEPHEN
#TEST INFORMATION, WECHSLER INTELLIGENCE SCALE FOR CHILDREN (WISC) - VERBAL SCALE
#REFER, WISC7
#BLOCK,
SCORE: 115 IQ
  VERBAL TESTS       SCALED SCORE
  INFORMATION        11
  COMPREHENSION      12
  ARITHMETIC         8
  SIMILARITIES       11
  VOCABULARY         11
  (DIGIT SPAN)       9
  SUM OF VERBAL TESTS 62

HA=1.00
*§ABSTRACT
#RTYPE,CAI
#SIMCASE,STEPHEN
#TEST INFORMATION, SLOSSON ORAL READING TEST (SORT)
#REFER, SORT2
#BLOCK,
EXAMINER'S COMMENTS: ERRORS INCLUDED 'FELLED' FOR 'FIELD', 'EVERY' FOR 'VERY' AND 'THERE' FOR 'THREE'.

HE MADE NO ATTEMPT TO READ SOME OF THE WORDS.

VISD CONF SORT EXCM IAPP=1.00
EXCM NU=1.00
/L2/
*§ABSTRACT
#RTYPE,CAI
#SIMCASE,STEPHEN
#TEST INFORMATION, DURRELL ANALYSIS OF READING DIFFICULTY:
  SOUNDS OF LETTERS-ANNOTATED TEST BOOKLET
#REFER, DUR39
#BLOCK,
SEE: STEPHEN SIMCASE STUDY GUIDE
      DURRELL ANALYSIS OF READING DIFFICULTY: SOUNDS OF LETTERS
3.3 The SIMCLIN Concept: Clinician Simulation

3.3.1 Introduction

The SIMulated CLINician or SIMCLIN is a computer based clinician which will attempt to solve real or simulated cases using a previously "programmed" clinical strategy and memory. The SIMCLIN may serve many uses, e.g., to study clinical problem solving theories, to aid with the diagnosis and treatment of real cases, to assist students with complex clinical problem solving problems, etc.

As may be seen from Figure 8, the SIMCLIN is a physical realization, implemented on a computer, of the Inquiry Theory conceptualization of a clinician.

The Inquiry Theory of clinical problem solving states that the cognitive characteristics of a clinician are determined by the interaction of his clinical memory of problems, cues, treatments, prescriptions and relationships among them, and his clinical strategy for the sequencing of a set of information processing tasks. Simulating a clinician can then be accomplished by simulating a memory and strategy and their interaction. In BMIS, such a simulation is called a SIMCLIN (SIMulated CLINician). A SIMCLIN is produced by inputting a clinical memory and one or more strategies. Strategies can be linked to problems so that a SIMCLIN can pursue an "overall" approach while investigating special problems by means of sub strategies linked to particular findings, e.g., an "abnormal" abdominal mass.

In the next sections of the present review of the SIMCLIN concept we will examine a sample interaction between a SIMCLIN and a researcher with a representative case of reading difficulty. In the subsequent sections we shall review the data files required to generate such SIMCLINS.

3.3.2 The CLINician SIMulation Subsystem (CLSS)

The CLSS Subsystem is intended to permit researchers to simulate theories of clinical problem solving, by allowing previously defined SIMCLIN's to interact with real or simulated case.

In the present version of BMIS, interaction is limited. SIMCLINS permitted only cue requests or questions which may be answered from a fixed set of alternatives (including: "Yes", "No", and "Unavailable"). This restriction will be eliminated in subsequent versions of BMIS.

We had intended to eliminate this limitation earlier. Our failure to do
so came about as a result of a hidden benefit of this mode of interaction: limiting responses required our "clinician models" to objectively state the criteria for interpreting cue values. Thus, simulations seem to be quite successful, in spite of—or perhaps because of—this restriction. In any case, the next versions of BMIS will allow the input of cue values rather than yes, no, or unavailable responses. For example, in place of the inquiry "Is the student's grade level at 3.0 or lower?" the system will ask "What is the student's grade level?" The printout below, illustrates the interaction of a SIMCLIN with a researcher who is studying the SIMCLINS reaction to a case record. As shown the mode of interaction leaves cue interpretation to the researcher. The SIMCLIN merely questions the
researcher about the contents of the clients file.

 helicopters

 helicopters

 helicopters

 helicopters

 helicopters
DOES THE PATIENT HAVE HYPERTENSIVE RETINOPATHY?
?yes
THE SIGNS AND SYMPTOMS SEEM TO INDICATE THAT THE CLIENT'S PROBLEM(S) INCLUDE: HYPERTENSION

HAS THE PATIENT LOST HIS APPETITE?
?no
HAS THE PATIENT'S TEMPERATURE BEEN HIGH?
?no
DOES THE PATIENT'S LIVER BIOPSY SHOW FIBROSIS?
?yes
DOES THE PATIENT'S LIVER BIOPSY SHOW FATTY CHANGE?
?yes
DOES THE PATIENT'S LIVER BIOPSY SHOW ADVANCED CIRRHOSIS?
?yes
DOES THE PATIENT'S LIVER BIOPSY SHOW BRIDGING?
?yes
DOES THE PATIENT'S LIVER BIOPSY SHOW LIMITING PLATE DISRUPTION?
?yes

THE SIGNS AND SYMPTOMS SEEM TO INDICATE THAT THE CLIENT'S PROBLEM(S) ARE:
PRIMARY HYPERTENSION
PORTAL CIRRHOSIS OF THE LIVER

END OF CLINICAL SIMULATION
?$finish
#EXECUTION TERMINATED

3.3.3 The CLINician CREation Subsystem (CLINCRE)

The CLINCRE subsystem is intended to aid with the preparation of SIMCLIN's -- both manually and computer based. In any case, the preparation of SIMCLINS requires the development of two major types of data input: the CLINICAL MEMORY data records and the CLINICAL STRATEGY data records. As illustrated Figure 9, these data records control the clinical interaction of SIMCASES.

To complete our brief overview of the CLINCRE subsystem let us examine an annotated listing of the data records required for the SIMCLIN's reviewed in previous sections of this report.

*ABSTRACT
#RTYPE,PDR
#SIMCLIN,STEPHEN
#MEMORY,STEPHEN
#PROBLEM,EOR RW SW LO
#DXFNAME
ELEMENTS OF READING
FIGURE 9. SUMMARY OF THE INPUT/OUTPUT ASSOCIATED WITH THE PRESENT MANUALLY GENERATED SIMDOC PRECEPTOR.

DECISION THEORY OR APPROACH (7):
A Program Written in the DOCS Simulation Language:

DIAGNOSE
PDR SEARCH
HISTORY CUE SEARCH
PHYSICAL CUE SEARCH
DIAGNOSIS JUDGMENT
END
TREATMENT SELECTION
END
PRESCRIPTION SELECTION
END

CLINICAL MEMORY
Required by cases (M):

A Set of DOCS Data Records

PDR - Problem Data Records
(associating cue values with problems)

CDR - Cue Data Records
(associating cue values with cue value collection query)

TDR - Treatment Data Records
(associating treatment plans with specific prescriptions)

CASE or PATIENT RECORD

PRINCIPAL COMPLAINT
REQUEST FOR CUES
CUE VALUES
DIAGNOSIS
TREATMENT

DOCS MODULE of BMIS.
RECOGNIZING WORDS
SIGHT WORDS
LOW
#DXDISCR
RECOGNIZING SIGHT WORDS INADEQUATE
#CTYPE, NORMAL
#ACTION,
#INTERVIEW CUES
EF ST GL RCD LE 3=9.00
#OBSERVATION CUES,
EOR RW SW BASIC DOL TS LOW=9.00
EOR RW SW BASIC DOL TB LOW=9.00
EOR RW SW VISD DOL TB IADQ=3.00
EOR RW DUR SLT TS IADQ=6.00
EOR RW SW DUR REC EXCM VL=9.00
#LAB CUES,
#DATE,UPDATE=770219,IN=770218
*$ABSTRACT
#RTYPE, CDR
#SIMCLIN, STEPHEN
#MEMORY, STEPHEN
#INQUIRY,
****
BASED UPON

SCHOOL RECORDS AND TEACHER REPORTS

IS THE STUDENT

ONE YEAR OR MORE BELOW PLACEMENT IF IN 1-3 GRADES, OR
TWO YEARS OR MORE BELOW PLACEMENT IF IN 4TH+ GRADES?
OR IS HE PLACED IN A LOW LEVEL READING GROUP?

^Y
#CNAME, EF ST GL READ RCD LOW
#CTYPE,
#CTYPE,
#RELIABILITY, 3
#DATE, IN=770208
*$ABSTRACT
#RTYPE, DSDR
#CLINICIAN, GS1
#STRATEGY, GS1
#SIMCLIN, GS1
#DSL,
DIAGNOSE
THRESHOLD FOR DIAGNOSIS IS 0.60
SELECT CUES WITH METHOD 1
EVALUATE HYPOTHESES WITH METHOD 6
SELECT HYPOTHESES WITH METHOD 1
EVALUATE DIAGNOSIS WITH METHOD 4
THRESHOLD FOR TERMINATING DIAGNOSIS IS 0.45
INITIAL CONTACT : TYPE 2
SELECT DIAGNOSIS WITH METHOD 2
EVALUATE CUES WITH METHOD 3
RETAIN 8 HYPOTHESES
INITIAL CONTACT
LABEL 100
PDR SEARCH
HYPOTHESES EVALUATION
RETAIN 10 CUES
IO CUE SEARCH
DIAGNOSIS JUDGMENT
IF NTRY LT 6 GOTO 100
REFER
END
TREAT
TREATMENT EVALUATION
TREATMENT SELECTION
END
PRESCRIBE
PRESCRIPTION SELECTION
END
#DATE, IN=770319


3.4.1 Introduction

The COMPUCEPTOR concept relates to the Inquiry theory corollaries of performance and instruction; i.e. that performance in clinical problem solving can be improved by providing external aids to decision and training.

In our judgement, the Performance and Instruction Corollaries held many important implications for methods to aid clinicians in their training and problem solving. The COMPUCEPTORship concept is our present attempt to provide computer support for the application of the performance evaluation and Instruction Corollaries to diagnosis and therapy training and problem solving. Essentially, the concept is to provide computer support for all aspects of training and performance for clinical problem solving — including computer aided diagnosis, the production of decision aids for clinical memory and strategy, SIMCASE aids for training, and complex CAI aids permitting students to interact with SIMCASE problems under the direction of computer preceptors (COMPUCEPTORS) which reflect the "best" decision making strategies and memories for the given case — based upon expert opinion.
The COMPUCEPTOR is an amalgamation of a SIMCASE and a SIMCLIN designed to assist clinicians with problem solving or learning to problem solve. The COMPUCEPTOR provides the basis of a COMPUCEPTORship.

The COMPUCEPTORship is an organized system of computer and non-computer based materials designed to aid clinicians in solving and/or learning to solve a prescribed set of clinical problems.

Figure 10 summarizes our notion of a COMPUCEPTORship. As shown, one important element in the COMPUCEPTORship is a computer based preceptor which will aid the student by providing "preceptor" assistance on simulated or real cases in the form of feedback on the value of cues collected, the appropriateness of the diagnostic hypotheses earned and the expected gain for proposed treatments. Thus, the COMPUCEPTOR is intended to fulfill the role of a "senior resident" making the equivalent of "grand rounds".

We anticipate three factors in the potential value of such computer and non-computer based COMPUCEPTORS in reading. First, there are even fewer "senior clinicians" in reading than in medicine. (The average resident in Internal medicine will "work-up" at least 5000 cases during his or her training. The average reading specialist/clinician will -- at most -- work up 5 cases). Second, the costs involved in preceptor training (internship clinic or residency) are much too high in medicine and are quite impossible in reading. By implementing COMPUCEPTORships on mini computers and even non-computer materials, we expect to provide such training for less than $1.00 per student hour. Third, it would appear that the systematic examination of clinical problem solving has, in itself, a value in making much less subjective the process of clinical problem solving in reading -- and therefore may significantly improve our knowledge of this field of human behavior.

3.4.2 The PRreceptor Simulation Subsystem (COMPSIM)

The PRECEPTOR SIMulation Subsystem is intended to allow the user to create a Computer Aided preceptor by combining together three basic elements: (1) A SIMCASE or Set of SIMCASES, i.e., a previously prepared set of one or more clinical cases. (2) A SIMCLIN or set of SIMCLINS, i.e., a previously prepared set of approaches to solving the problems of the SIMCASES, utilizing a method judged to be "valid" by a set of senior clinicians, or by empirical criteria. (3) SIMPRE* i.e., a strategy for providing instruction or assistance to learning using the SIMCASEs and SIMCLINS.

*This concept was suggested by the work of George Bordage, M.D. We are pleased to acknowledge his contribution.
Presently, we are only beginning to understand and provide for reasonably complex Simulated Preceptor strategies. The printout given below summarizes a very simple SIMPRE which merely provides a critique of the clinical problem solver's diagnosis.

$\text{RUN WBSW:BMISX+*OLDFIO 1=FMEDIFT 2=FMEDDFT}$
EXECUTION BEGINS

THE BASIC MANAGEMENT INFORMATION SYSTEM VERSION 1.1
THE INSTITUTE FOR RESEARCH ON TEACHING
MICHIGAN STATE UNIVERSITY
1976

?*$prss
ENTERING THE PRECEPTOR SIMULATION SUBSYSTEM
?*$simcase
?*$name james walker
?*$go

BEGINNING PRECEPTOR SIMULATION

THE ENCOUNTER WITH THE CASE BEGINS:
A 47 YEAR OLD WHITE MALE ELECTRICIAN WAS BROUGHT TO THE HOSPITAL
IN A STATE OF ACUTE WITHDRAWAL. THE HISTORY AS FOLLOWS WAS OBTAINED FROM
RELIABLE FAMILY MEMBERS:
THE PATIENT HAD BEEN DRINKING APPROXIMATELY ONE QUART OF 86 PROOF
BOURBON EVERY DAY FOR 35 DAYS BEFORE HIS ADMISSION. DURING THIS TIME,
HIS FOOD INTAKE HAD CONSISTED OF SUBSTANTIAL DAILY AMOUNTS OF EGGS, MILK,
BREAD AND POTATOES. HE HAD TAKEN TWO OR THREE MULTIVITAMIN CAPSULES DAILY.
HE HAD HAD NO VOMITING OR DIARRHEA.

ENTER KEYWORD
?past medical history
SALIENT FEATURES OF THE PAST HISTORY INCLUDE TWENTY
YEARS OF ALCOHOLISM WITH FREQUENT HOSPITALIZATION FOR ACUTE
INTOXICATION AND WITHDRAWAL WITHOUT OVERT MALNUTRITION.

ENTER KEYWORD
?*$compuceptor hypotheses with 2 cues

ENTER YOUR HYPOTHESES (END LIST WITH BLANK LINE)
?none

BASED ON ALL CUES WHICH SHOULD BE COLLECTED FOR EACH
HYPOTHESES, THE LIKELIHOODS OF THE HYPOTHESES UNDER
CONSIDERATION ARE:

<table>
<thead>
<tr>
<th>RELEVANCE</th>
<th>POSSIBLE DIAGNOSIS</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0392</td>
<td>CHRONIC HEPATITIS</td>
</tr>
<tr>
<td>0.1860</td>
<td>PORTAL CIRRHOSIS OF THE LIVER</td>
</tr>
<tr>
<td>0.2647</td>
<td>ACUTE ALCOHOLISM</td>
</tr>
</tbody>
</table>

THE NEXT MOST IMPORTANT CUES TO COLLECT ARE:
ENLARGED LIVER
LIVER BIOPSY

ENTER KEYWORD
Liver biopsy

LIVER BIOPSY SHOWS FATTY CHANGE, ADVANCED CIRRHOSIS,
BRIDGING, LIMITING PLATE DISRUPTION, FIBROSIS, AND
FOCAL NECROSIS.

ENTER KEYWORD
**$diagnose

ENTER YOUR DIAGNOSES (END LIST WITH BLANK LINE)
Acute alcoholism
Cirrhosis of the liver

THE COMPUCEPTOR DIAGNOSIS IS;
ACUTE ALCOHOLISM, CIRRHOSIS OF THE LIVER, ANEMIA

YOUR PERCENT AGREEMENT IS 66%

THE DIAGNOSIS YOU FAILED TO AGREE WITH THE COMPUCEPTOR WAS ANEMIA.

END OF COMPUCEPTOR SIMULATION
**$finish
$EXECUTION TERMINATED

3.4.3 The PRReceptor Creation Subsystem (PRCS)

The input required for the creation of a COMPUCEPTOR is summarized in Figure 11. As shown, the input required for SIMCASE and SIMCLIN creation is necessary. In addition, commands must be given to select a SIMulated PRReceptor (SIMPRE) strategy. These commands specify the nature of the instructional strategy including:
(1) Control of Interaction, i.e., whether an active or passive role will be allotted to the compucer, i.e., will the COMPUCEPTOR be allowed to interrupt the student or will the student receive instruction only upon demand.
(2) Types of feedback, i.e., whether evaluative responses will be offered to cue requests, to hypothesis generations or both.
(3) Complexity of feedback, i.e., whether only manual numerical evaluations will be provided or more complex messages, offered on the basis of students' responses.

Presently, only the least sophisticated types of SIMPRE can be created with the PRCS subsystem. More complex strategies will become available within the next year. The data input sample given below illustrates the preparation required for our present COMPUCEPTORS.
FIGURE 11. FUNCTIONING OF THE COMPUCEPTOR INSTRUCTIONAL SYSTEM

INSTRUCTIONAL TEAMS

SIMPLIFIED INSTRUCTIONAL STRATEGY

CAI VOCABULARY

CASE DATA RECORD FOR SIMCASE

DIAGNOSTIC & TREATMENT DATA RECORDS FOR SIMCLIN

SIMCASE

BASIC INFORMATION MANAGEMENT SYSTEM (BMIS)

SIMCLIN

BASIC INFORMATION MANAGEMENT SYSTEM (BMIS)

CAI VOCABULARY

INTERACTION

(SIMCEPTOR) INTERACTION

STUDENT DECISION MAKERS
4.0 APPLICATIONS IN MEDICINE

4.1 Introduction

Our work in medical education has primarily been the formalization of the Inquiry Theory in the form of the BMIS programs. However, we have also conducted many simulated experiments in the process of validating the system's capability for predicting or "postdicting" clinicians' behavior.

We have limited our reports of these experiments for two reasons: First, many of the experiments were conducted on versions of the system which became obsolescent almost immediately after the experiment. Second, most of the experiments were conducted with clinical memories of less than 30 medical problems.

We now feel constrained to make some limited report of our findings for several reasons. First, some of our results have been replicated repeatedly within and outside of medicine. Second, the set of problems used in recent studies constitutes a limited -- but representative -- problem domain. Thus, the sample includes both highly similar and highly dissimilar problems, all but one of which occur with fair frequency in the practice of primary care (e.g., see Marsland, Wood, and Mayo, 1976). Third, a medical colleague of ours (Bordage, 1976) has developed a database including about 250 medical problems which constitute over 90 percent of family practice problems according to the Virginia Study (Marsland, et al., 1976); and is now helping us implement this data bank on BMIS. Thus, we should soon be able to attempt a replication of these results with a large sample of problems. However, until this replication is performed, the following results should only be tentatively generalized to predictions involving a very limited number of problems.

4.2 The Dr. ACADM vs Dr. CLINIC SIMCLIN Studies

The saga of Doctor A and Doctor C began about three years ago. A very early version of BMIS was evaluated and demonstrated by contrasting two stereotypic clinical strategies using the identical clinical memory. The strategies were suggested by an Internist in our medical school as two of the most obvious approaches which different physicians take, or the same physicians take on different cases.

Our physicians' words were certainly prophetic, since Doctor ACADM and Doctor CLINIC have been appearing and reappearing in medicine, reading, trouble shooting and instructional design for the past three years.

The original doctors A and C experiment follows, presented in much the same manner as it was at the American Association of Medical Colleges annual SMC conference in 1975.
We used two case records of real patients for the experiment. Case 1 was a patient record for a case of mononucleosis with the classic symptoms of weakness, fever, and lymphatic infection. Case 2 was a patient record for a case of viral hepatitis with the usual symptoms of fatigue, weakness, fever, and enlarged liver. The chief complaints of both problems were identical. The approaches to problem solving simulated in our experiment were intended to reflect two stereotypes of medical practice, the "academic" versus the "clinically" oriented physician. Dr. ACDM placed strong emphasis upon careful, precise diagnosis prior to treatment and stresses the value of laboratory studies. Dr. CLINIC placed his emphasis on good but efficient medical services. He stressed the need for rapid identification of the most likely problem with minimum information processing. It should be emphasized that a single physician uses both of these strategies depending on the circumstances. For example, in serious problems a clinically oriented physician may well switch to a more conservative academic approach.

The figure below illustrates and explains the strategies used to represent CLINIC and ACDM. The two strategies are represented as Decision Simulation Language (DSL --- the BMIS programming language) programs.

The results for ACDM versus CLINIC performance are shown in FIGURE 13. As indicated, the quantitative prediction for these cases is that the academic approach requires about twice as many cues and costs about four times as much to reach correct diagnoses as the CLINIC approach.

These findings were later confirmed in more extensive studies using variants of CLINIC and ACDM. In general, the results seem to conform to the evidence supplied by physicians, that highly conservative approaches tend to be more expensive for straightforward medical cases. It is very interesting that the A and C diagnosis strategies reappeared so very strongly in reading clinicians which are variously described as people "who like to give tests" versus people "who like to get right to the students' problem". We will consider this in a subsequent section.

4.3 The Diagnostic Complexity SIMCLIN Studies

The complexity studies dealt with the question of the factors involved in making a medical problem "difficult" to diagnose. Two factors were considered for examination: the number of problems in the case and the reliability of the data base provided by the case. Reliability was defined in two ways: (1) the reliability of symptoms, signs and laboratory findings which were judged by clinicians as relevant to the diagnosis; and (2) the reliability of the same types of cues which were judged by clinicians to be irrelevant to the diagnosis.

The SIMCLIN's used in the experiment were all systematic variations on Dr. CLINIC. The results constitute a prediction of what physicians behaving
FIGURE 12. THE ANATOMY OF TWO SIMDOCS

DSL PROGRAM FOR DR. ACADM

1. DIAGNOSE
   - STANDARD
   - THRESHOLD FOR CUES IS 0.0
   - THRESHOLD FOR DIAGNOSIS IS 0.90
2. THRESHOLD FOR HYPOTHESES IS 0.10
   - RETAIN 5 HYPOTHESES
   - RETAIN 10 CUES
   - EVALUATE HYPOTHESES WITH METHOD 6
3. FOR ROUTINE SEARCH
4. HISTORY CUE SEARCH
5. LABEL 100
   - FOR SEARCH
6. HISTORY CUE SEARCH
7. PHYSICAL CUE SEARCH
8. HYPOTHESES EVALUATION
9. LABORATORY CUE SEARCH
10. DIAGNOSIS JUDGMENT
11. IF NTRY LT 5 GOTO 100
12. END
13. TREAT
14. STANDARD
15. TREATMENT EVALUATION
16. TREATMENT SELECTION
17. PRESCRIBE
18. PRESCRIPTION SELECTION

DSL PROGRAM FOR DR. CLINIC

1. DIAGNOSE
   - STANDARD
   - THRESHOLD FOR CUES IS 0.20
   - THRESHOLD FOR DIAGNOSIS IS 0.70
2. THRESHOLD FOR HYPOTHESES IS 0.10
   - RETAIN 2 HYPOTHESES
   - RETAIN 10 CUES
   - EVALUATE HYPOTHESES WITH METHOD 6
3. LABEL 100
   - FOR SEARCH
4. HP CUE SEARCH
5. HYPOTHESES EVALUATION
6. DIAGNOSIS JUDGMENT
7. LABORATORY CUE SEARCH
8. DIAGNOSIS JUDGMENT
9. IF NTRY LT 4 GOTO 100
10. END
11. TREAT
12. STANDARD
13. TREATMENT EVALUATION
14. TREATMENT SELECTION
15. PRESCRIBE
16. PRESCRIPTION SELECTION
17. END

X ACADM begins by collecting a set of routine cues. CLINIC does not.

3 Both establish the beginning of regular diagnosis with "LABEL"

4 ACADM collects history & physical, evaluates hypothesis but does not try to diagnose. CLINIC collects history & physical at same time and tries to diagnose.

5 Both doctors collect lab cues. Both attempt a diagnosis.

6 If diagnosis is possible, both skip to treatment. Otherwise ACADM tries 5 and CLINIC 4 times more before referring patient to a diagnostician.

7 Both physicians attempt to obtain an acceptable problem management procedure and write out prescriptions if one is acceptable.

According to the Inquiry Theory would do if faced with the problem of diagnosing cases varying in complexity. Specific procedures and findings are given below. Again, these predictions are limited to relatively small clinical memories.

4.3.1 Procedure
4.3.2 Results

4.4 The SIMCASE Studies

For purposes of completeness, we should mention the medical compuceptor work. As a side project of the BMIS activities we developed in 1975-1976 a series of programs to permit the simulation of patients and later cases. Research in the field has not been actively pursued due to our commitments to the Institute for Research on Teaching. However, we intend to move back into this field when time permits.

4.5 Summary and Discussion

Our studies of the application of BMIS in medicine have yielded some interesting but very tentative results.

To begin with, many of the empirical findings observed in studies of physicians (Elstein, et al., 1976) seem to be reproducible with our simulated clinicians. Some of these findings (such as the relative efficiency of clinic oriented versus medical school oriented strategies) even seem to
generalize into other fields of clinical problem solving.

Second, our studies have yielded predictions on the accuracy of diagnosis which seem "reasonable" in terms of clinical experience. Thus, the probability of a correct, complete, diagnosis varies inversely with the number of problems in the case and the accuracy or reliability of the database associated with the case.

Third, those SIMCLINs which made diagnostic decisions on the basis of the aggregation of findings involving many (often redundant) cues: (1) retained accurate diagnostic performance rates in spite of error in the data; (2) tended to become less efficient in the presence of error; (3) tended to become less efficient when errors involved cues irrelevant to the diagnosis than when errors involved cues that were relevant. The mechanism for this effect seems to be the generation of greater numbers of incorrect hypotheses in the case of error in irrelevant cues.

Especially interesting to us was the finding that an hypothesis testing model which aggregates data for decision making seems to be highly robust with respect to errors in the database and in information processing. Accurate diagnoses can be achieved even with relatively large errors in the data, such as might be introduced by inaccurate recall of past medical history by a patient or laboratory errors, or with suboptimal decision making by the physician due to fatigue or distraction.

As the bottom line of this discussion, we should like to present an interesting research problem. In 1976, one of our SIMCLIN experiments explored the effect of reducing the size of a SIMCLIN's memory to approximate that of a human. As the size of memory was reduced, the SIMCLINS studied switched from a parallel to a linear strategy for hypothesis testing; i.e., only one or two hypotheses were considered simultaneously, but the complete set was eventually examined and the correct diagnosis obtained. This result depended upon the separation of the diagnostic decision from the decision to continue or halt the diagnostic process. Otherwise, no "pre-programming" was involved.

We have not been able to replicate these results with other SIMCLINS. The key factor seems to be the existence of an order of precedence among the problem representations; i.e., the problems must be ordered with respect to overall likelihood of occurrence (i.e., the main weight associated with the problem data record). As yet we have lacked the resources to actively pursue the question of such strategy selection decisions by SIMCLINS.
5.0 APPLICATIONS IN READING

5.1 Introduction

Our work on the observation and simulation of clinical problem solving in reading began in 1975 when professor Lois Bader helped us develop a demonstration of the use of BMIS in reading. This application included both a SIMCASE and a SIMCLIN study. This system was made a part of the IRT proposal which was funded in April 1977. From April until October 1977 a pilot study was run on the feasibility of and the procedures for conducting Observational and Simulation studies in reading.

The results of the pilot study confirmed the feasibility of using methods developed in medicine by Elstein and Shulman et al. (1976), Borrows, et al. (1976), and others -- to study clinical judgment in reading. Also confirmed was the idea of computer simulation of clinical problem solving in reading. The pilot study indicated that substantial "tightening" of procedure would be required. The overall results were very promising and the Clinical Information Process In Reading (CLIPIR) project was officially initiated by the IRT in November 1976. The present section provides a summary of the work of the CLIPIR project during the pilot study and the past months of the operational project. The figure below summarizes the current organization of the project designed for the production of SIMCASEs, SIMCLINs, and COMPUCEPTORS in reading.

5.2 Simcase Development for Problems in Reading Diagnosis and Remediation

5.2.1 Introduction

The development of SIMCASES has become a fairly well regulated process since November 1976 when we first implemented the present procedures. Also such production has become rather complex and somewhat technical. Therefore, let us begin with some basic definitions of the various types of SIMCASEs generated.

The SIMCASE is a collection of materials which simulate a case of reading difficulty with a substantial amount of fidelity. These materials include taped interviews, test scores, test booklets, and taped selections of the child reading. The appearance of a SIMCASE changes depending upon it's purpose and presentation. The distinctions currently used include the purposes of instruction and performance testing and the presentation by computer and manually.

The Instructional SIMCASE is a SIMCASE designed for aiding student clinicians. Included are "tutorial" comments on the case materials prepared by a senior reading clinician, including suggested cue interpretations, diagnoses, etc.
FIGURE 14. CLIPIR AND CLOSELY RELATED PROJECTS

1. SIMCASE Development Team (SIT) #1
   Objectives: Develop and validate 2 manually based SIMCASEs. Perform validation studies. Complete documentation.
   
   **Staff:**
   - Professor George Sherman, Coordinator - responsible for the planning and validating of the SIMCASEs (MSSC's)
   - Anne D. Varshney, Team Leader - responsible for the production of the SIMCASEs (MSSC's)

2. CONCEPTIZER Development/Data Analysis Team (CDTA Team)
   Objectives: Develop and validate 2 CONCEPTIZERs for the 4 MSSC's. Perform validation studies. Complete documentation.
   
   **Staff:**
   - Chris Wagner, Coordinator - responsible for planning for production of CONCEPTIZERs
   - Susan Cull, Team Leader - responsible for production of CONCEPTIZERs

The Performance SIMCASE is a SIMCASE designed to permit a student or
practicing clinician "practice" his or her diagnostic and remediation skills.

The MBSC (Manually Based SIMCASE) is a SIMCASE designed to be presented to a clinician by another individual using a "cue-inventory" for the case. The clinician may or may not be provided with an inventory.

The CBSC (Computer Based SIMCASE) is a SIMCASE designed to be presented by a terminal or mini computer.

The SIMCASE replication is a SIMCASE generated by altering enough cues which are irrelevant to the diagnosis to cause the case to be unrecognizable as diagnostically identical to the generating case.

In the past our SIMCASEs have primarily been used to study the problem solving techniques of reading clinicians. Thus the interaction of human clinicians with SIMCASEs has yielded our basic data for the computer simulation studies. The next major step in our project will be the use of SIMCASES and SIMCLINS to form COMPUCEPTORships for clinical training. With this brief overview let us now consider the development of the various types of SIMCASES.

5.2.2 SIMCASE Development Procedures

Basically, the procedure developed for SIMCASE development is a modified version of the usual techniques used in the design and preparation of instructional modules. Thus we have employed a modification of the systems approach to instruction development (e.g., see Davis, Alexander, and Yelon). Perhaps the simplest method of conveying our procedures is by means of an annotated diagram. Such a diagram is given in Figure 15. The figure is read from left to right. The upper portion of the figure describes graphically each step in the method. The lower portion provides a narrative of the process.

5.2.3 The SIMCASEs in Reading

Perhaps the most important aspects of our SIMCASEs are their contents and their characteristics. Let us first meet those simulated clients which have thus far been prepared by CLIPIR staff.

5.2.3.1 The Present CLIPIR cases

As may be seen from the above descriptions, each of these cases was based upon a real child with a set of real reading difficulties. Further, each case included problems judged important by the SIMCASE Selection and Evaluation Team.

We cannot, at present, discuss these cases in great detail since information about them might compound the results of experiments planned for 1977. Later, when more cases and their replications are available we shall
publish further on our "children". We can, however, at least summarize for
the reader the contents of the SIMCASEs.

5.2.3.2 The CASE Information

The data provided for each SIMCASE varies. However, all cases include
most standard types of information requested by clinicians and reading
specialists. The contents include: biographical data, physical health,
family data, school records, achievement tests, diagnostic reading tests,
cognitive ability tests, etc. Figure 16 shows the Case Inventory for
Stephen. As shown, for each basic cue (e.g., the Durrell Oral Reading)
several forms of data are provided including test scores, comments, test
booklets, audio recordings, test directions, and tutorials. (Test Directions
and Tutorials are only provided for Instructional SIMCASEs). The acronyms in
the body of the inventory are used by the clinician to request information,
when the Inventory is available. For example a request for "DUR21" would
yield the Durrell word Recognition Word Analysis Test Booklet for Stephen.

The case is introduced by presenting the "Initial contact cues"
including a sketch of the child; a brief introduction to why he was brought
to the clinic and a recorded initial interview.

The cases are designed for presentation either as a stimulus situation
to elicit clinical problem solving behavior or as a teaching aid to be used
by one or more students, as a focus for their learning (the medical
education's concept of a "Focal Problem"), as an example for the instructor,
or as a training-testing experience for students.

5.2.3.3 Comments

As a note in passing, in our studies we have primarily used the
condition in which a clinician is provided with a case inventory. We
recognize that this severely limits the generalization of findings. However,
early pilot studies indicated that until we learned more about diagnostic
problem solving in reading—a more structured situation would be required
which provided the clinician with an aid to memory and a definition of the
task. This opinion was recommended originally by the Inquiry Project staff
which suggested the use of the case inventory to partially eliminate the
memory factor in problem solving.

In any case, ALL OF THE FINDINGS DISCUSSED BELOW MAY BE GENERALIZED ONLY
TO SITUATIONS WHERE THE CLINICIAN KNOWS THE ENTIRE SET OF AVAILABLE
INFORMATION, AND WHERE LITTLE COST OF TIME IS ASSOCIATED WITH THE COLLECTION
OF INFORMATION. Thus, one should be aware that our results are probably
biased in the direction of excessive cue collection.

5.2.4 Applying SIMCASEs in Reading Diagnosis and Remediation

We are only beginning to understand the potential for SIMCASE
FIGURE 16. CASE INFORMATION INVENTORY - STEPHEN

<table>
<thead>
<tr>
<th>TYPES OF INFORMATION</th>
<th>Test Scores</th>
<th>Examiner's Comments</th>
<th>Test Booklet</th>
<th>Audio Recording of Session</th>
<th>Test Directions</th>
</tr>
</thead>
<tbody>
<tr>
<td>BACKGROUNDER INFORMATION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BIOGRAPHICAL DATA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PHYSICAL/HEALTH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SOCIAL/FAMILY</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CLASSROOM INFORMATION</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| STANDARDIZED ASSESSMENT                      |             |                     |              |                             |                 |
| ACHIEVEMENT TEST—INDIVIDUAL (PEABODY)        | PEA1        | PEA2                | PEA3         | PEA5                        |                 |
| Reading Recognition                          | PEA7        | PEA8                | PEA9         | PEA5                        |                 |
| Reading Comprehension                        | PEA10       | PEA14               | PEA15        | PEA5                        |                 |
| Spelling                                     | PEA16       | PEA17               | PEA18        | PEA5                        |                 |
| General Information                          | PEA19       |                     |              |                             |                 |

| ACHIEVEMENT TEST—GROUP (GATES-MACGINTIE)     |             |                     |              |                             |                 |
| Vocabulary                                   | GM21        | GM22                | GM23         | GM25                        |                 |
| Comprehension                                | GM26        | GM27                | GM28         | GM29                        |                 |
| Speed and Accuracy                           | GM21        | GM22                | GM23         | GM25                        |                 |

| DIAGNOSTIC TEST OF READING DIFFICULTY— (GATES-KINGLIP) |             |                     |              |                             |                 |
| Recognizing and Blending Common Word Parts   | GM31        | GM32                | GM33         | GM34                        |                 |

| DIAGNOSTIC TEST OF READING DIFFICULTY—(DURBELL) |             |                     |              |                             |                 |
| Oral Reading                                 | DUR1        | DUR2                | DUR3         | DUR4                        | DUR5            |
| Silent Reading                               | DUR6        | DUR7                | DUR8         | DUR9                        | DUR11           |
| Listening                                    | DUR13       | DUR14               | DUR15        | DUR16                       | DUR17           |
| Comprehension                                | DUR19       | DUR20               | DUR21        | DUR22                       | DUR23           |
| Word Recognition and Word Analysis           | DUR25       | DUR26               | DUR27        | DUR28                       | DUR29           |
| Visual Memory of Words-Primary                | DUR31       | DUR32               | DUR33        | DUR34                       | DUR35           |
| Hearing Sounds in Words                      | DUR37       | DUR38               | DUR39        |                             |                 |

| GRADED WORD LIST (SLOSSON ORAL READING TEST) |             |                     |              |                             |                 |
| (SORT1)                                      |             |                     |              |                             |                 |

| COGNITIVE ABILITY (WISC)                     |             |                     |              |                             |                 |
| Full Scale                                   | WISC1       | WISC2               | WISC3        | WISC4                       | WISC5           |
| Verbal Scale                                  | WISC7       | WISC8               | WISC9        | WISC10                      |                 |
| Performance Scale                             | WISC13      | WISC14              |             |                             |                 |

| NON-STANDARDIZED CSDD-SEPBT                  |             |                     |              |                             |                 |
| BASIC SIGHT VOCABULARY (BOLCH)               | DOL1        | DOL2                | DOL3         | DOL4                        | DOL5            |

| INDIVIDUAL READING                          |             |                     |              |                             |                 |
| SKILL ANALYSIS (SOUND-SYSGOL ASSOC.)         |             |                     |              |                             |                 |

Applications. Within the limits imposed by our concentration on the intellectual-rational aspects of clinical problem solving, three main usages
are most important to our thinking: (1) as aids to the study of problem solving behavior; (2) as aids to the training of problem solvers; and (3) as aids to the evaluation of problem solvers' performance. Within the foreseeable future, our work will concentrate on the first two applications. The problem of performance evaluation will require many years of preliminary work.

Whatever the application, SIMCASEs are used in one of two modes, i.e., Manually Based or Computer Based, as illustrated in Figures 17 and 18, below.

**FIGURE 17. USING THE MANUALLY BASED SIMCASE**

As shown in the figure, control of the interaction is given to the clinician after the initial contact (chief complaint) has been presented.
Requests for cues yield such cues in the form of printed materials (school records, test scores, etc.); graphic materials (a sketch of the child, test booklets, sample writing); and audio records (samples of speech, oral reading). The Case Presenter serves as an interface or communications channel between the MBSC and the clinician or student. In most of our work, the case presenter has been an experimenter studying clinical problem solvers' behavior. Other configurations are possible. For example, Prof G. Sherman has suggested that students in Diagnostic Classes could provide instruction to each other by alternatively acting as clinician and case presenter.

**FIGURE 18. USING THE COMPUTER BASED SIMCASE**

As shown in the figure, the Computer Based SIMCASE (CBSC) functions identically to the Manually Based version except that a computer terminal or
mini-computer is used as the case presenter. Printed materials are handled via a terminal, graphics and auditory cues are referenced by the terminal for retrieval from the SIMCASE Study Guide by the clinician or student. The major objective of the CBSC is greater economy of applications. Thus, given the present low cost of mini-computer systems (about $8500 for the CLASSIC) we might expect to be able to deliver applications at a fraction of the cost of using humans for the somewhat boring task of case presentation.

Given this brief review of applications, let us now turn to the uses of SIMCASEs which are presently most important in our work.

5.3 Observational Studies of Clinical Problem Solving Behavior

5.3.1 Introduction

In all of our work -- but most especially in that dealing with applications -- we are attempting to be quite careful and conservative as befits those who call themselves 'scientists'. In our judgment, far too many "research validated" innovations have been implemented and discredited in the past two decades within education. All of our staff in the CLIPIR and BMIS projects are encouraged to put into daily practice the working scientific principles of empiricism, objectivity, and measurement. As a result we are very slowly beginning to build an understanding of clinical problem solving in reading. Our work is in no manner ready for implementation. Thus our findings function mainly as sources of hypotheses upon which to base future research. findings are mostly sources of hypothesis which will be pursued by However, given enough time we are confident that a sound theory of clinical problem solving can be developed. The major source of this optimism is our past research on the observational study of clinical problem solvers in reading. Portions of this research are summarized below.

5.3.2 Methods

In general, the methods applied in our observational studies have been adopted from those methods used in studies of medical problem solving, (e.g., in Elstein, et al. 1976). These methods have been refined over the past months through our pilot CLIPIR study and subsequent tests of procedures. The following are the results of our methodological studies. Hopefully, these procedures will be further improved during the next period of the project.

5.3.2.1 Subjects.

Defining a coherent population of subjects and drawing random samples has proved quite impossible. The new federal regulations on "informed
"consent" requires such a frightening array of disclosures and cautions that we can only depend upon volunteer subjects, thus eliminating any "random" sampling. Further, we have found great difficulty in identifying populations of clinicians from which to sample. There appears to be some degree of confusion as to what constitutes the required training for a reading clinician, i.e., someone qualified to diagnose and remediate reading problems. The population parameters we presently are using are those classifications set up by the International Reading Association for Reading Teachers, Reading Specialists, Reading Diagnosticians and Supervisors. (e.g., see Ekwall, 1976).

In particular, the subjects used in the CLIPTR pilot study are described in detail by Gardner (1977) as including: (1) college teachers of reading and reading diagnosis; (2) reading specialists at the district level; (3) reading diagnosticians from clinics; and (4) reading teachers recommended as having special competence in diagnosis and remediation.

Sampling has been on the basis of supervisor recommendation. More systematic sampling will be attempted in our planned studies for the spring and summer of 1977. If possible we hope to identify a population of volunteers and randomly sample from this group. In any case we are collecting extensive information on our subjects so as to be able to provide a complete description of them.

With respect to numbers of subjects, our pilot study provided five subjects per SIMCASE. Subsequent studies are increasing this very limited base. For example, Stephens' (1977) study is using thirty subjects per case. We recognize that a much larger sample would be desirable, but economics presently will not allow such samples. Effectively, the TRT has suggested that we examine fewer subjects in greater detail. Thus far, this appears to have been a sound decision.

5.3.2.2. Procedures

The procedures presently followed in our observational studies are summarized in the figure below. The observational study may be divided into three stages involving three individuals.

The individuals are as follows. The Experimentor (E) is responsible for instructing the other individuals and for the presentation of the SIMCASE (whether Manual or Computer Based) and the management of the study. The Clinical Observer (CO) is an experienced, trained reading clinician who is responsible for preparing an observational session protocol and for debriefing the subject, using the protocol. The Clinical Subject (CS) is the true subject of the Observational study. The CS is chosen from some population of interest in the study (e.g., senior clinicians, reading specialists, teachers, etc). Only the CS actually interacts with the SIMCASE during the observational Session.
The three stages in the Observational Study are the Observational Session, the Debriefing Session, and the Cue-by-Hypothesis Matrix Analysis Session.

**FIGURE 19. PROCEDURES FOR THE OBSERVATIONAL STUDIES**

Briefly, these stages are as follows. In the Observational Session the CS interacts with the SIMCASE under the guidance of the E. The CO, behind a one-way mirror, observes the interaction and prepares a protocol listing time, cue requested, and comments by the CS on his decision making. The session is recorded on audio tape (retained for our archives) and video tape (for debriefing). The TV camera is focused not on the CS but on whatever the CS is examining. The session ends with the CS preparing a written Diagnosis and Remediation Plan.
In the Debriefing Session, the CO and CS attempt to recreate the CS's thinking by studying the Video record, the protocol, and via probing by the CO. A Debriefing Protocol is prepared by the E and CO, indicating cues collected, and comments by the CS on why each cue was collected (e.g., to test a "hunch" or hypothesis; because the cue was "interesting", etc.)

Finally, in the Cue-by-Hypothesis Matrix Analysis Session, the CS is excused and the CO and E prepare a cue-by-hypothesis matrix summary of the session, as is illustrated in Figure 20, below.

As shown in the Figure, the cue-by-hypothesis matrix provides a preliminary analysis of the data. Across the top of the matrix are given cue names and times when collected. Down the side are the "problems" or "strengths" (presumed states of the SIMCASE "child" with respect to reading) which were diagnosed by the CS during her or his interaction with the SIMCASE. Entries in the matrix indicate the CS's judgment with respect to each hypothetical problem or strength on the basis of the corresponding cue (e.g., was the hypothesis confirmed, disconfirmed, etc.).

In case of disagreements in judgments between the E and CO during the preparation of cue-by-hypothesis matrices, the judgment is referred to a third individual who is a senior clinician associated with CLIPIR.

Totally, each experimental session requires about three hours of real time or about 9 person hours. Obviously, rather expensive research. A more complete discussion of procedures is given elsewhere (CLIPIR staff, 1977).

5.3.2.3 Data Analysis

Presently, the cue by hypothesis matrix is used to generate a number of states describing the clinical interaction. All of these were adopted from work in medical education, especially work by Norman (1977) and Elstein, et al., (1976). The major statistics are as follows.

The commonality Index of the diagnosis is a measure of the agreement of a diagnosis with the diagnoses of a set of critical clinicians. If Pi denotes the proportion of the critical physicians mentioning the jth diagnostic category and i denotes the diagnoses of a particular clinician, then

$$\text{CI} = \frac{\sum_{i=0}^{m} P_i}{\sum_{i=0}^{m} \max_{j} P_i}$$

; i.e. the sum of the proportions for the diagnosis divided by the sum of the same number of most agreed upon diagnostic categories

Thoroughness is the proportion of relevant cues collected. Efficiency is given by thoroughness divided by time in minutes to diagnosis or number of relevant cues divided by time in minutes.
Number of cues collected and number of diagnoses generated for the first, second, third, and fourth quarter are also presently obtained. Further discussion of all these statistics is given in Norman (1977), Gardner (1977), and Stephens (1977).

Perhaps the greatest problem in measuring diagnostic performance is in deciding upon the diagnostic categories to be used to combine data across subjects. The problem is that there is a very limited standard vocabulary for describing reading problems and treatments. Hence it is sometimes difficult to determine when clinicians are using different names for the same concept or the same name for different concepts.

Another important problem is in distinguishing between diagnostic "problems" and "strengths". The statement of "sight words" in a diagnoses may mean (1) a deficiency or (2) a strength in recognizing sight words.

At present, we employ several major devices to obtain vocabulary control for the analysis of protocol content.

First, the protocols are left in the "language" of the clinician until after the cue by hypothesis matrix analysis. In this manner errors in our interpretation can at least be confirmed by later investigators.

Second, the Clinical Subject is required to clearly define the diagnostic categories listed in her or his written diagnosis. This procedure has been of great assistance in identifying synonyms for common concepts (e.g., "trouble with look-alike words" translates into "poor visual discrimination of word with similar letter configurations", etc.)

Third, we have established a Taxonomy of Reading Factors (TRF). The TRF is both a dictionary of factors which are judged by "experts" to affect reading and a linguistic nexus which links together the concepts underlying the clinician's verbal expressions in preparing SIMCASES, in reacting to SIMCASES, and in building SIMCLINs.

We shall consider the TRF in more detail later. Presently, we need merely note that the TRF is used as an authority list indicating which standard phrases are substituted for the clinical subjects description of his or her diagnostic categories.

5.3.3 Results

The results for our pilot observational study of clinical problem solving are presently being analyzed. They will be reported in detail by Gardner (1977). The preliminary findings may be worth consideration. The reader should recall that these results are based upon only ten highly selected senior clinicians and two SIMCASEs. Generalizations are hazardous.

5.3.3.1 The Performance of Reading Clinicians
If we apply the measures normally used in medical education to judge the performance of reading clinicians in our pilot study, the results are encouraging. As shown in Figure 21, the average diagnosis commonality index for one SIMCASE is .89. Generally for both SIMCASES the commonality indexes ranged from .33 to .95, with the great majority above .80. In absolute numbers, considering all diagnosis categories of strengths and weaknesses for the SIMCASE Stephen, 80% of the clinicians agreed upon seven out of the total of 18 diagnostic categories and 60% agreed upon 10 out of the total of 18 diagnostic categories.

Medical studies of diagnostic commonality have yielded significantly lower agreement levels; e.g., Diagnostic Commonalities averaging .43 (Barrows, et al. 1976). The difference is probably due to the fact that the more precise diagnostic categorization vocabulary in medicine permits a greater range of disagreement (e.g., a diagnostic category of "ulcer" would count as a non-match if the criterial clinicians used the category "duodenal" or "peptic" ulcer).

In any case, reading clinicians (at least highly trained reading clinicians) seem to agree upon the diagnostic categories applied to SIMCASES. The agreement in fact is higher than that observed in medicine.

Examining the figure further we may note that our clinicians (just as medical clinicians) tend to differ in thoroughness and efficiency, but such differences have no obvious relationship to diagnostic commonality — although the inverse relationship shown in the figure is indeed intriguing. Interpretations should await further confirmation.

5.3.3.2 Other Behavior of Reading Clinicians: Early versus Continuing Hypothesis Generation.

The preliminary data from the pilot study offers one more interesting result. Over the past several years, cumulative evidence has suggested two obvious clinical strategies: (1) the "academic" or information gathering approach, wherein the clinician collects a large number of cues and then attempts to interpret the results (very much like a clinical psychologist "testing" a client); and (2) the "clinical" or hypothesis testing approach, wherein the clinician uses early generation of hypotheses to guide his selection of cues and presumably make diagnosis more efficient. The results of our pilot study indicates that such individual strategies are also employed in reading. For example, Figure 22 below shows a graph indicating the proportions of hypotheses generated during the clinical encounter of five clinicians with SIMCASE Donald. As shown three of the clinicians generated almost all of their hypotheses in the first half of the session; but two (OS-6 and OS-8) generated most of their hypotheses in the second half.

Similar results were found for the SIMCASE Stephen except that more of
the clinicians behaved like subject OS-8; i.e., generated more hypotheses late in the encounter. It is difficult to say whether these preliminary results will be replicated. The use of the Cue Inventory and a generous time limit of one hour may have encouraged this tendency of some clinicians to test extensively without hypothesizing. Time and a good deal more work may tell us the answer.
FIGURE 22. NUMBER OF HYPOTHESES GENERATED BY CLINICIANS IN SUCCESSIVE QUARTERS OF THE CLINICAL INTERACTION FOR THE SIMCASE DONALD

PROPORTION
6
5
4
OF TOTAL
3
2
1
HYPOTHESES
4
5
6
PROPORTION
OS-7
OS-3
OS-1
OS-6
OS-8

GENERATED
SUCCESSIVE QUARTERS OF THE OBSERVATIONAL SESSIONS
5.4 SIMCLIN Development for Reading Diagnosis and Remediation

5.4.1 Introduction

The task of creating an effective Simulated Clinician (hereafter SIMCLIN), is an art which is not well understood. Many techniques have been employed, ranging from questionnaires to actual behavioral observations.

During the past several years we have by trial and error developed a method similar to that employed in the Observational Study, but using clinicians who are trained participants in the project and a procedure which requires a much more exhaustive set of introspections than are expected from our Observational Study clinicians. The methods and results we shall describe are, at best, tentative, and may be modified extensively in subsequent research. Presently they are the best we have. In the future, we may be able to develop empirically based SIMCLINs in reading, as have been developed in medicine at the Leeds Medical Center. Such SIMCLINs utilize empirical data on the association of cues with problems and reactions to treatment in cases observed in an operative clinic. Presently, we must continue to depend upon the best introspective judgment of our most senior clinicians.

5.4.2 Defining a SIMCLIN

To begin our discussion, let us review the information structures which define a Simulated Clinician in the BMIS programs.

A SIMULATED CLINICIAN is a set of information structures which define an automaton. This automaton, when executed by the Basic Management Information System (BMIS), will interact with real or simulated cases of reading difficulty with the objective of determining a diagnoses (the problems, or action required states characterizing the case) and a treatment plan (specifying the actions intended to place the case in a more non-action required state, i.e., "eliminate" the problems).

A SIMCLIN is composed of two major types of information structures: a clinical memory and a clinical strategy. The development of a SIMCLIN requires the creation of both these structures.

5.4.2.1 Defining a Clinical Memory

For SIMCLINs designed primarily for diagnoses three major steps are involved in creating a Clinical Memory.

5.4.2.1.1 Defining a "problem" or "State Theory"

Clinical memories are defined mainly in terms of Problems (states
requiring actions). Hence some basic definition of the action/non-action requiring (i.e., desirable/undesirable) states must be given. For example, the undesirable states in medicine are usually defined as pathological states, described in books on clinical medicine and medical pathology. (out of context). (Dolch test scores). (Low)'.s In the AMA study about 4000 "pathologic" states were identified. It is useful to note that most of these states are quite rare. For example, the Virginia study found that only 315 medical "problems" (including non-pathologic states) covered 98% of the cases observed in family practice (Marsland, et.al. 1976).

The basic definition of the different states used in reading has been reading factors. Reading factors are an observable characteristic of the case which according to behavioral scientists and practitioners may affect the individual's ability to read.

The basic document used in our research is the Taxonomy of Reading Factors (TRF). The TRF is a computer based data file which includes a record for every factor thus far uncovered by our research and is updated almost weekly. (Published versions are issued once per quarter). Examples of entries in our TRF are given in Figure 23. As may be seen from the figure, the TRF contains records including the name of a reading factor, e.g., Recognizing words, Sight words. Basic (the ability to recognize instantly certain lists of high frequency words). Entries are included at various levels of specificity. For example the record for recognizing words is at a less specific level than Recognizing words, Sight Words.

The TRF developed in CLIPIR has been created from an extremely liberal point of view. Any factor was added which seemed to hold promise in aiding with Diagnosis or Treatment, regardless of the theoretic orientation of its proponents. Thus, the TRF freely mixes phonics factors with language experience and psycho linguistic factors, etc. This approach has been widely recommended by our most experienced clinicians, all of whom seem very catholic, indeed.

To summarize, the Taxonomy of Reading Factors or TRF provides what we may call the primitive elements of our vocabulary for defining clinical memory. Thus, the clinical memory is an information structure which may be generated as follows.

Let f::= Any factor (state name) in the TRF, e.g., "Recognizing words, Sight Words"

Let V::= Any value associated with a factor reflecting the state of the case; e.g., "low", "high".

Let M::= Any observable measurement procedure relevant to the TRF; e.g., "DOLCH TEST"

Let C::= Any conditions upon values; e.g., "SCORE", "BOOKLET", "DIRECTED"
Then we may define Diagnostic Clinical Memory as including representations of Problems and cues defined as follows:

Problem: \( P_x := <F, C, V> \), a problem is an ordered set of factor(s) condition(s) and value(s); e.g., "\((\text{Recognizing Words}, \text{Sight Words}), \text{out of context}, \text{low})\)"

cue: \( C_x := <F, C, M, V> \), a cue is an ordered set of factor(s) condition(s) measurement devices), and value(s); e.g. "\((\text{Recognizing Words}, \text{Sight Words}), \text{out of context}, \text{Dolch list}, \text{low})\)"

Once such \( D_x \) and \( C_x \) concepts are well defined in terms of even a rudimentary "state-theory" remarkably effective SIMCLIN memories may be defined with relative ease by creating Problem and cue data records.

Presently, the fundamental source of our TRF has been Ed Ekwall's text (Ekwall, 1976) and our reading clinicians. Other sources include the NIE taxonomy, and advanced graduate students in the Reading Section at MSU.

4.2.1.2. Defining the Problem Data Records (PDRs)
As noted previously, the PDR is a means for associating a problem (Px) with a set of cue names. As shown in Figure 24, the PDR is composed of fields (computer equivalent of paragraphs) which indicate the name of the problem, and provide the names of related cues with weights or probabilities reflecting the reliability and relevance of each cue. Irrelevant cues (with zero weights) do not appear in the PDR.

**FIGURE 24. SAMPLE PDR FOR THE PROBLEM OF SIGHT WORD DIFFICIENCY**

```plaintext
*ABSTRACT
#TYPE,PDR
#SIMCLIN,STEPHEN
#MEMORY,STEPHEN
#PROBLEM,EOR RW SW LO
#DXFNAME
ELEMENTS OF READING
  RECOGNIZING WORDS
  SIGHT WORDS
LOW
#DXDISCR
RECOGNIZING SIGHT WORDS INADEQUATE
#CTYPE, NORMAL
#ACTION,
#INTERVIEW CUES
EF ST GL RCD LE 3=9.00
#OBSERVATION CUES,
EOR RW SW BASIC DOL TS LOW=9.00
EOR RW SW BASIC DOL TB LOW=9.00
EOR RW SW VISD DOL TB IADQ=3.00
EOR RW DUR SLT TS IADQ=6.00
EOR RW SW DUR REC EXCM VL=9.00
#LAB CUES,
#DATE,UPDATE=770219,IN=770218
```

5.1.2.1.3 Defining the Cue Data Records (CDRs)

The CDR is a memory structure which represents a cue (Cx). The CDR permits the association of a cue name (given in a PDR) with a cue
interpretation. Cue interpretations are presently highly formalized questions instructing the individual interacting with the SIMCLIN on what specifics to look for in the case records. A sample, CDR is given below for the cue "<(Recognizing Words, Sight Words), (out of context), (Dolch test scores), (Low)>".

FIGURE 25. SAMPLE CDR FOR THE CUE LOW DOLCH LIST PERFORMANCE

$ABSTRACT

ITYPE, CDR
SIMCLIN, STEPHEN
MEMORY, STEPHEN
INQUIRY,
*******
ADMINISTER OR EXAMINE THE RESULTS FOR
THE BASIC DOLCH SIGHT WORDS LIST

ID THE STUDENT

SCORE BELOW 70% AT HIS GRADE PLACEMENT IF HE IS PLACED IN GRADES 1 THROUGH 4, OR SCORE BELOW 80% ON THE HIGHEST LEVEL TEST IF HE IS PLACED ABOVE THE FOURTH GRADE?

Y

NAME, EOR RW SW BASIC DOL TB LOW
CTYPE,
CSTYPE,
RELIABILITY, 3
DATE, IN=770208

5.4.2.1.4 Comments

Perhaps the most important aspect of the BMIS approach to simulation lies in the notion of pre-designed generality. For example, the conversion of clinical memory from medicine to reading, required only the input of new data records based upon a new "state theory." Thus, the SIMCLIN memory
structure is State-Theory independent, i.e., a clinical memory can be created in any field for any particular approach to diagnosis and remediation without any program modification of BMIS. The principle requirement is that there must exist a moderately coherent theory describing the various states, indicators of states, and methods for moving from one state to another.

With respect to the latter, i.e., treatment actions, BMIS requires data records for describing treatments and associations with problems (Treatment Data Records, TDRs) and data records for associating treatments with output prescriptions (Prescription Data Records, RXDRs). None of these information structures have been much considered in the initial studies described below.

5.4.2.2 Defining a Clinical Strategy

A clinical strategy specifies the order in which the basic information processing tasks (e.g., cue collection and interpretation, hypothesis generation, evaluation, diagnosis judgment, etc) will be performed on the Clinical Memory to yield decisions on diagnosis and remediation.

The creation of a SIMCLIN strategy requires two major steps: The Preparation of a Clinical Strategy and the Preparation of a Decision Simulation Language Data Record (DSDR) containing a program written in the BMIS Decision Simulation Language (DSL) which captures the clinical strategy.

5.4.2.2.1 A representative Clinical Strategy

Figure 26 below shows a directed graph or "flow chart" for a rather commonly proposed strategy for reading clinicians. The figure is self-explanatory.

5.4.2.2.2 Preparing a DSDR for a Clinical Strategy

Although the decision strategy summarized in the previous section may appear to be rigid, the strategy can in fact be represented by many different SIMCLINS. One of such is summarized in Figure 27.

As may be seen from the DSDR, the strategy consists of three sub-strategies. First, an attempt is made to determine if the subject does in fact have a reading problem. An heuristic Data Record is used to initiate the collection of relevant cues. Next a "Diagnosis Judgment" command is used to determine if a diagnosis of "no reading problems" is appropriate. If not, another HDR search considers factors relating to home, school, and physical problems. Finally, the strategy attempts to locate the specific problems in the reading act.

5.4.2.2.3 Comments

As noted previously with respect to clinical memory, the SIMCLIN clinical strategy is state-theory independent within a wide margin of generality. In addition, the clinical strategy is memory independent, i.e.,
FIGURE 26. A SAMPLE CLINICAL STRATEGY IN READING DIAGNOSIS

START

COLLECT INITIAL CONTACT CUES, INCLUDING PRINCIPLE COMPLAINT

TEST TO DETERMINE IF CASE HAS A READING PROBLEM

COLLECT CUES RELEVANT TO READING POTENTIAL & READING ACHIEVEMENT

DO CUES SHOW POTENTIAL SIGNIFICANTLY GREATER THAN ACHIEVEMENT?

YES

EXIT: NO PROBLEM

NO

DX: READING PROBLEMS

EXAMINE HOME, SCHOOL, PHYSICAL RECORDS OF CAUSES OF READING PROBLEMS

ELEMENT SHOWS WEAKNESS PROBABLE CAUSE?

NO

YES

DX: POSSIBLE CAUSES

EXAMINE ELEMENTS OF READING TEST FOR READING PROBLEMS GENERAL TO SPECIFIC

SELECT ELEMENT; COLLECT CUES, EVALUATE AS DX

DX: ADD PROBLEM TO PROBLEM LIST FOR DX

DO CUES SHOW ELEMENT IS A PROBLEM?

YES

NO

TEST FOR CONTINUATION: ARE ALL DX TESTS COMPLETE?

YES

STOP

NO
many different strategies can be used with a single memory. This point will be illustrated in later sections.

5.4.3 Preparing a SIMCLIN

The specific task involved in creating a SIMCLIN memory and strategy is summarized in the figure below. As shown in Figure 28 three steps are involved.

In the initial interaction, the clinician examines each of the physical cues available in a SINCASE and comments upon them indicating (1) hypotheses suggested and (2) the psychological cues (PSY-cues) present within a physical cue (PHY-cues), e.g., the Dolch word test booklet might include "psychological" cues like: missing words having similar appearances; missing words with the suffix "ing"; etc. Members of the Compuceptor Development.
Teams (CDT) monitor PSY cues and hypotheses and compare names with PDR, and CDR and TDR lists for previously generated cues and hypotheses. New and revised PDRs and CDRs are added to the computer listings at the case of the Initial Interaction session.

In the second interaction session, problem-cue weights are generated for each PDR and Inquiry records are generated for each CDR. Presently, we have found that a two-step method for cue weight assignment is most effective. First, for each PSY-cue, the Clinician assigns a reliability weight of "1", "2", or "3" on the basis of his or her judgment of the "reliability" of the cue. By reliability is meant the degree of dependability of the cue; i.e., would the cue be expected to appear in the same value occasionally, much of the time, or most of the time. Second for each PSY-cue problem combination, the clinician assigns a validity weight of "1", "2", or "3", i.e., the validity of the cue for the identification of the problem described by a given PDR, where 3=very important, 2=important, 1=some importance.

The actual weight assigned is the product of the reliability and validity weights. Thus, The relevance weight for the ith PSY-cue (CDR) given the jth problem (PDR) is given by

\[ R_{ij} = R_{wi} \times V_{wj} \]

where

\( R_{wi} = \) reliability of the ith cue and
\( V_{wj} = \) validity of the ith cue for the jth problem.

This method of weighting causes the relevance of a cue to increase most rapidly as both reliability and validity increase, and more slowly as one or the other increases.

In general, the validity weight is associated with the importance of the cue in identifying a reading factor of significance. The reliability is associated with the methods of measuring the factor used by the cue, e.g., the particular test used.

In the final session, the clinician re-examines the entire set of PDRs, CDRs, and TRF entries from the point of view of the case, adjusting weights, adding new PDRs and CDRs where appropriate. Finally, the clinician suggests a general strategy which should be used in diagnosing the SIMCASE. These elements are translated by the CDT staff into the data records required by the BMIS programs.

5.4.4 Using a SIMCLIN

To complete our discussion, let us review the principal methods of using the SIMCLIN as suggested earlier. The methods of greatest importance in our work are summarized in Figure 29. The figure is self-explanatory: the
numbers are used to identify various types of uses and related explanations are given at the bottom of the figure.

5.5. SIMCLIN Studies of Clinical Problem Solving in Reading

5.5.1 Introduction

The Simulation studies in reading have not been actively pursued in reading since SIMCASE developments are pre-requisite to serious simulations in reading. However, we do have some preliminary studies which deserve comment.

5.5.2 CLINIC and ACADM in Reading, a historical note.

Since the present document is partially designed to serve as an archive for the CLIPIR BMTS projects, we should like to insert here a note on the history of our SIMCLINs. In the Fall of 1975, Lois Bader, her students*, and present CLIPIR staff developed two demonstrational SIMCLINs for the IRT proposal. Upon the suggestion of the reading clinicians, the ACADM and CLINIC strategies developed in medicine were applied to clinical memories in reading diagnosis.

The study involved four cases of reading problems; "word calling" (i.e., word analysis skills, extremely poor comprehension) "sight word problems" (i.e., poor recognition of high frequency words), "blending problems" (inability to perform basic phonetic tasks), and "affix recognition" (the inability to identify important word parts). Each case was presented identically. The results of running rather primitive clinical reading memories and the medical strategies on the four cases are summarized in Figure 30.

As may be seen from the figure all of our expectations were grandly confirmed. The SIMCLINs produced nearly perfect diagnoses; CLINIC cost less and requested fewer cues than ACADM. It was a study to delight almost any scientist. Interestingly, this study seems to have set the pattern for further work. One of our continuing concerns has been why SIMCLINs seem so consistently effective and predictable often given rather poor data preparation.

In any case, our initial reading simulation study deserves at least this brief note in passing.

5.5.3 Preliminary Simulation Studies of Clinical Problem Solving in Reading

*Participants included Lynne Hofmeyer, Elaine Stephens, Jay Gardner, and the present writers.
1. The SIMCLIN is online and controls the diagnostic and/or remediation directly. "Computer Diagnosis"
2. The SIMCLIN serves as an aid to a clinician or researcher: "Computer Aided Diagnosis"
3. The SIMCASE is queried directly. Primarily used in research studies to reduce costs through batch processing.
4. The SIMCASE is presented indirectly by a human clinician or researcher. Used to test the human interface problem.
5. The clinical interaction involves the diagnosis and remediation of a real client. ("Computer Aided Diagnostic Counseling")
6. The clinical interaction involves a direct conversation with a client. This interaction is not appropriate to SIMCLIN usage.
FIGURE 30. RAW DATA AND MEANS FOR FOUR DIAGNOSES BY TWO SIMDOCS (ACADM AND CLINIC)

Our serious studies of simulation in reading have only recently begun. However, we may offer some preliminary results of our attempts to generate SIMCLIN's which attack problems in much the same manner as the subjects used in the pilot CLIPPIR observational study. As an example we have chosen the SIMCLIN's using the "GS" memory, created by one of our most senior clinicians in reading. "GS" was observational study subject OS1-3. After the procedure discussed previously was applied by "GS" to the creation of a clinical memory, our clinician assisted us in the creation of two strategies. Simclin GS-2 emphasized the early generation of hypotheses in diagnoses. SIMCLIN GS-3 emphasized the continuous generation of hypotheses, throughout the clinical interaction. Thus our clinician attempted to create (1) a strategy typical of his performance (like OS1-3) and a strategy very much more conservative (like subject OS1-3 in the CLIPPIR pilot observational study). Figure 31 shows the strategies comprising SIMCLINs GS-2 and GS-3.
\*\*ABSTRACT

**RTYPE, DSDR**

**CLINICAN, GS2**

**STRATEGY, GS2**

**SIMCLIN, GS2**

**DSL**

DIAGNOSE

THRESHOLD FOR DIAGNOSIS IS 0.60

SELECT CUES WITH METHOD 1

EVALUATE HYPOTHESES WITH METHOD 6

SELECT HYPOTHESES WITH METHOD 1

EVALUATE DIAGNOSIS WITH METHOD 4

THRESHOLD FOR TERMINATING DIAGNOSIS IS 0.45

INITIAL CONTACT: TYPE 2

SELECT DIAGNOSIS WITH METHOD 2

EVALUATE CUES WITH METHOD 3

INITIAL CONTACT

HDR GS H1

RETAIN 20 CUES

IO CUE SEARCH

RETAIN 8 HYPOTHESES

PDR SEARCH

HYPOTHESES EVALUATION

RETAIN 6 CUES

IO CUE SEARCH

DIAGNOSIS JUDGEMENT

RETAIN 4 HYPOTHESES

PDR SEARCH

HYPOTHESES EVALUATION

IO CUE SEARCH

DIAGNOSIS JUDGEMENT

RETAIN 3 HYPOTHESES

PDR SEARCH

HYPOTHESES EVALUATION

IO CUE SEARCH

DIAGNOSIS JUDGEMENT

IO CUE SEARCH

DIAGNOSIS JUDGEMENT

REFER

END

TREAT

TREATMENT EVALUATION

TREATMENT SELECTION

END
5.5.3.1 Method

The SIMCLINs GS-2 and GS-3 were run under BMIS using the SIMCASE Stephen in the Manually Based SIMCASE mode, which allows the intrusion of errors due to misinterpretation of SIMCLIN queries. (i.e., the 4th mode in Figure 29). Complete protocols were collected and the total interaction was divided into quarters using cues-collected as the basis, since real time is not relevant in a time-shared computer environment (later studies using dedicated mini-computers will eliminate this artifact so that real time may be used).

5.5.3.2 Results

The results of the GS studies are summarized below. Figure 32 shows the Diagnostic Commonality index of GS-2 and GS-3 in comparison with GS(OSL-3) and the other ssbjects of the CLIPTR pilot study. Figure 33 shows the proportion of hypotheses generated during successive quarters of the simulated diagnosis. As noted, GS-2, like the human clinician OSL-3, generates hypotheses early; while GS-2, like the human clinician OSL-8 generates hypotheses later in the interaction.

5.5.4. Conclusions

While it is much too soon to draw conclusions on the simulation of clinical problem solving in reading we can offer some tentative observations.

First, it would appear that important aspects of clinical decision making in reading can be captured by computer systems.

Second, highly trained clinicians can translate their knowledge into SIMCLINs which in turn mirror important aspects of the human clinicians behavior.

Third, the simulation of reading clinicians offers a largely unexplored field of study which promises substantial rewards to those choosing this area for study.
TERMS OF THE DIAGNOSTIC COMMONALITY INDEX

FIGURE 12. RESULTS OF SIMULATIONS OF CLINICAL PROBLEM SOLVERS EXPRESSED IN

BMIS THEORY MANUAL

80
FIGURE 33. COMPUTER SIMULATION RESULTS: EARLY VERSUS LATE GENERATION OF HYPOTHESES

SUCCESSIVE QUARTERS OF THE OBSERVATIONAL SESSIONS
6.0 SUMMARY and DISCUSSION

The present document is designed as an archival source for the Inquiry Theory of Clinical Problem Solving — which has been the principal subject of research for one or another of the present writers for nearly a decade.

The first part of the document presents a distillation of research findings of many scholars on clinical problem solving in medicine and other fields. These findings have been summarized by means of three basic concepts which seem to describe a rudimentary theory of clinical problem solving performance and instruction. The first concept involves the principle of case simulation: that realistic valid clinical problem solving behavior can be induced by simulated cases presented to clinicians. The second concept concerns the principle of clinician simulation: that the problem solving behavior of clinicians is probabilistically dependent upon clinical memory, clinical strategies, and the case. The third concept involves the performance evaluation and instructional corollaries of the Inquiry Theory: that clinical problem solving performance can be measured and that such performance can be improved by providing (1) external aids to clinical memory or strategy, (2) training aids for clinical memory or strategy, and (3) combinations of external and training aids.

The second part of the document, presents an overview of the computer system (BMIS) which constitutes a formalization of the Inquiry Theory summarized in part one.

The third portion of the present document, discusses the history and current status of applications for the Inquiry theory and the BMIS programs. Currently two fields are considered: clinical problem solving in medicine and in reading.

In conclusion, a summary of very tentative suggestions for applications of this research may be noted for diagnosis. Basically, our research would suggest that inaccurate diagnoses in both medicine and reading involve the following:

* Failure to generate correct hypotheses in the initial encounter

* Failure to obtain a complete diagnosis including casual factors.

*Failure in Level of Dx, Dx at a level of abstraction; not sufficiently specific for any but a general treatment.

These three factors are not offered as principles supported by research, but merely as hunches suggested by our observations. We include them in print merely in the hope that they will spur other researchers to study more precisely the diagnostic remediation act.
7.0 BIBLIOGRAPHY


Alcorn, F. S., O'Donnell, E., and Ackerman, L. V.; The Protocol and Results of Training Nonradiologists to Scan Mammograms; Radiology, 1971, 99, 523-529.


Angoff, W. H.; Test Reliability and Effective Test Length; Psychometrika, 1953, 18, 1-4.


Barrows, H. S., and Bennett, K.; The Diagnostic (Problem Solving) Skill of the Neurologist: Experimental Studies and Their Implications for Neurological Training; Archives of Neurology, 1972, 26, 273-277.
Barrows, H. S., Feighter, J. W., Neufeld, V. R., and Norman, G. R.; Data Gathering, Problem Solving, and Management - Observation of the Clinical Encounter in Time Segments; McMaster University, Hamilton, Ontario, Canada, 1976


Beach, B. H.; Expert Judgment About Uncertainty: Bayesian Decision Making in Realistic Settings; Organizational Behavior and Human Performance, 1975, 14, 10-59


Bleich, H. L.; Computerized Clinical Diagnosis; Computer Conference, December 1974, 33(12), 2317-2319


Bordage, G., M.D.; Computers and Medical Diagnostic Problem Solving; Dept of Biometry, School of Medicine, Case Western Reserve University, June 1976.

Bouckarett, A.; Computer Diagnosis of Thyroid Tumors; Nuclear Medizin. 1969, 7(4), 319-330

Bourne, L. E., Jr.; Human Conceptual Behavior; Boston: Allyn and Bacon, 1966

Bourne, L. E., and Dominowski, R. L.; Thinking; Annual Review of Psychology, 1972, 23, 105-130

Bowman, E. H.; Consistency and Optimality in Managerial Decision Making; Management Science, 1963, 9, pp. 318-321

Brandt, E. N., Jr. (Director); Proceedings Conference on the Use of Computers in Medical Education; University of Oklahoma Medical Center, 1968
Brehmer, B.; Hypotheses About Relations Between Scaled Variables in the Learning of Probabilistic Inference Tasks; Organizational Behavior and Human Performance, 1974, 11, 1-27


Bruce, R., and Yarnell, S.; Journal of Chronic Disabilities; 1966, 19, pp. 473-484


Bruner, J. S.; Toward a Theory of Instruction; Cambridge, Mass.: Belknap Press, 1966

Bruner, J. S.; The Relevance of Education; New York: Norton, 1971


Chamberlin, T. C.; The Method of Multiple Working Hypotheses (1890); Reprinted in Science, 1965, 148, 754-759

Chan, S.; Annotated Bibliography on Computer-Assisted Diagnosis; Michigan State University. Mimeo, 1974

Chapman, L. J., and Chapman, J. P.; Genesis of Popular but Erroneous Psychodiagnostic Observations; Journal of Abnormal Psychology, 1967, 72,


Claparede, E.; La genese de l'hypotheses; Archives de Psychologie, 1934, 24, 1-154


Cook, M. G., and Dixon, M. P.; An analysis of the reliability of detection and diagnostic value of various pathological featurees in Crohn's disease and ulcerative colitis; Gut, 1973, 14, 255-262

Cope, O., and Zacharias, J.; Medical Education Reconsidered; Philadelphia: Lippincott, 1966

Cornfield, J.; Proceedings of 6th IBM Medical Symposium; 1964, pp. 567-589

Cornfield, J. and Tukey, J. W.; Average Values of Mean Squares in Factorials; Annals of Mathematical Statistics, 1956, 27, 907-949


Craig, R. M.; Recent Research on Discovery; Educational Leadership, 1969, 26, 501-508

Cronbach, L. J.; Coefficient, Alpha and the Internal Structure of Tests;
Psychometrika, 1951, 16, 297-334


Cuadra, C. A. (Editor); Annual Review of Information Science and Technology; Volume 6, 8. Encyclopedia Britannica, 1971, 1973


DeDombal, F. T.: Production of Artificial "Case Histories" by Using a Small Computer; British Medical Journal, 1971, 2, 578-581


Duncker, K.: On problem solving; Psychological Monographs, 1945, 58, No. 270


Edwards, W.; Behavioral Decision Theory; Annual Review of Psychology, 1961, 12, 473-498


Einhorn, H. J.; The Use of Nonlinear, Noncompensatory Models in Decision Making; Psychological Bulletin, 1970, 73, 221-230

Einhorn, H. J.; Expert Measurement and Mechanical Combination; Organizational Behavior and Human Performance; 1972, 7, 86-106


Elstein, A. S., Loupe, M. J., and Erdmann, J. B.; An Experimental Study of Medical Diagnostic Thinking; Journal of Structural Learning, 1971, 2, 45-53


Elstein, A. S., and Shulman, L. S.; Scoring and Analysis of Medical Inquiry
Protocols; Unpublished manuscript, Michigan State University, 1971.


Engel, G.L.; The Deficiency of the Case Presentation Method of Clinical Teaching; New England Journal of Medicine, 1971, 284, 2024

Erdmann, J.B.; An Appraisal of Three Scoring Procedures as Discriminations Between Good and Poor Problem Solvers; (Loyola Psychometric Laboratory Publications No. 40). Chicago: Loyola Psychometric Laboratory, 1964

Essex, B. J.; Approach to Rapid Problem Solving in Clinical Medicine; British Medical Journal 3, July 1975, 34-36


Finn, J.D.; Univariate And Multivariate Analysis of Variance and Covariance: A Fortran IV program (occasionnal paper No. 9); East Lansing, Mich.: Office of Research Consultation, College of Education, Michigan State University

Fisher, R.; Annual Eugenics; 1936, 7, 179-188

Fitzgerald, L., Overall, J., and Williams C.; American Journal of Roentgenology; 1966, 97, pp. 901-905

Flexner, A.; Medical Education in the United States; New York: Carnegie Foundation for the Advancement of Teaching, 1910

Flom, P.K.; Performance in the Medical Internship; Unpublished doctoral dissertation, University of California, Berkeley, 1970

Freud, S.; The Psychopathology of Everyday Life; In A.A. Brill (Ed. and trans.) The Basic Writings of Sigmund Freud. New York: Modern Library, 1938 (originally published in 1905

Friedman, R. B.; A Computer Program for Simulating the Patient-Physician Encounter; Journal of Medical Education, January 1973, 48, 92-97

Fries, J.F.; Experience Counting in Sequential Computer Diagnosis; Archives
of Internal Medicine; 1970, 126, pp. 647-651


Gagne, R.M.; Instruction Based on Research in Learning; Engineering Education, 1971, 61, 519-523


Getzels, J.W.; Creative Thinking, Problem Solving and Instruction; In E.R. Hilgard (Ed.), Theories of Learning and Instruction 63rd Yearbook of the National Society. Chicago: University of Chicago Press, 1964


Glendening, L.; POSTHOC: A Fortran IV Program for Generating Confidence Intervals Using Either Tukey or Scheffe Multiple Comparison procedures (occasional paper No. 20); East Lansing, Mich.: Office of Research Consultation, College of Education, Michigan State University, 1973

Goffman, E.; The Presentation of Self in Everyday Life; Garden City, New York: Doubleday, 1959

Goldberg, L.R.; Simple Models or Simple Processes? Some Research on Clinical Judgments; American Psychologist, 1968, 23, 483-496

Goldberg, L.R.; Man Versus Model of Man: A Rationale, Plus Some Evidence, for a Method of Improving on Clinical Inferences; Psychological Bulletin, 1970, 73, 422-432
Goldberg, L. R.; Five Models of Clinical Judgment: An Empirical Comparison Between Linear Representations of the Human Inference Process; Organizational Behavior and Human Performance, 1971, 6, 458-479

Goldberg, L. R.; Student Personality Characteristics and Optimal College Learning Conditions: An extensive search for trait by treatment interaction effects; Instructional Science, 1972, 1, 153-210


Gorry, G., and Barnett, G.; Computers and Biomedical Research; 1968, 1, pp. 490-507


Gough, H. G.; California Psychological Inventory; Palo Alto, California: Consulting Psychologists Press, 1957


Gough, H. G., Hall, W. B., and Harris, R. E.; Admissions procedures as forecasters of performance in medical training; Journal of Medical Education, 1963, 38, 983-998


Hammond, K. R.; Computer Graphics as an Aid to Learning; Science, 1971, 172,
Hammond, K.R.; Inductive Knowing In J. Royce, and W. Rozeboom (Eds.), The Psychology of Knowing, New York: Gordon and Breach, 1972


Harvey, A.M., and Bordley, J.; Differential Diagnosis (2nd ed.) Philadelphia: Saunders, 1970

Harvey, A.M., Johns, R.J., Owens, A.H., and Ross, R.S. (Eds.), The Principles and Practice of Medicine (18th ed.); New York: Appleton-Century-Crofts, 1972


Hebb, D.O.; What Psychology is All About; American Psychologist; 1974, 29, 2, 71-79

Helfer, R.E., and Slater, C.H.; Measuring the Process of Solving Clinical
Diagnostic Problems; British Journal of Medical Education, 1971, 5, 48-52


Hoffman, P.J.; The Paramorphic Representation of Clinical Judgment; Psychological Bulletin, 1960, 57, 116-131

Hoffman, P.J.; "Toward Quasi Automation of the Decisions of a Medical Specialists Examination Board in Evaluating Clinical Competence; Oregon Research Institute, 1974


Hoffman, P.J., Slovic P., and Rorer, L.G.; An Analysis of Variance Model for the Assessment of Configural Cue Utilization in Clinical Judgment; Psychological Bulletin, 1968, 69, 338-349

Hogarth, R.M.; Process Tracing in Clinical Judgment; Behavioral Science, 1974, 19, 298-313

Holsti, O.R.; Content Analysis for the Social Sciences and Humanities; Reading, Mass.: Addison-Wesley, 1969


Hoffman, P. J.; Physicians Appraise Other Physicians: Improving The Decisions of a Medical Specialty Board; ORI Research Bulletin, 1974a, 14(4)

Hoffman, P. J.; Automation of the Decisions of a Medical Specialist Examination Board in Evaluating Clinical Competence; Oregon Research Institute, 1974b

Honey, D. N., (Ed.); The Merck Manual of Diagnosis and Therapy; Merck and Company, Inc., 1972

Horrocks, J.C., McCann, A.P., Staniland, J.R., Leaper, D.J., and DeDombal, F.T.; Computer-Aided Diagnosis: Description of an Adaptable System, and
Operational Experience with 2,034 Cases; British Medical Journal, April 1972, 2, pp. 5-9

Hoyt, C.J.; Test Reliability Estimated by Analysis of Variance; Psychometrika, 1941, 6, 153-160

IFIPS; Proceedings for Medinfo 74; The International Federation of Information Processing Societies World Conference on Computers in Medicine; North Holland Publishing Company, 1974

John, E.R.; Contributions to the Study of the Problem-Solving Process; Psychological Monographs, 1957, 71, No. 447


Kagan, N.; Influencing Human Interaction—Eleven Years With IPR; Canadian Counsellor, 1975, 9, 74-97


Kegel-Flom, P.; Predicting Supervisor, Peer, and Self Ratings of Intern Performance; Journal of Medical Education, 1975, 50, 812-815


Kendell, R.E.; The Role of Diagnosis in Psychiatry; Oxford: Blackwell Scientific Publications, 1972

Kessel, F.S.; The Philosophy of Science as Proclaimed and Science as Practiced: "Identity" or "Dualism"?; American Psychologist, 1969, 24, 999-1005
King, L.S.; What is a Diagnosis?; Journal of the American Medical Association. 1967, 202, pp. 714-717


Knowles, J.H. (Ed.); Views of Medical Education and Medical Care; Cambridge, Mass.: Harvard University Press, 1968

Kohout, L.; Algebraic Models in Computer-Aided Medical Diagnosis; Medinfo 74, North Holland Publishing Company, 1974, pp. 575-579

Kuhn, T.S.; The Structure of Scientific Revolutions (2nd ed.); Chicago: University of Chicago Press, 1970

Kumar, V.K.; The Structure of Human Memory and Some Educational Implications; Review of Educational Research, 1971, 41, 379-418

Laboratory of Computer Science; Medical Education Programs: User's Manual; Boston: Massachusetts General Hospital, June, 1976

Larson, C.M.; The Heuristic Standpoint in the Teaching of Elementary Calculus; Ann Arbor, Mich.: University Microfilms, 1960


Lindberg, D.A.B.; The Computer and Medical Care; Charles C. Thomas, 1968

Lindsay, P.H., and Normal, D.A.; Human Information Processing; New York: Academic Press, 1972


Lodwick, G.; Computer Analysis of Tumor Roentgenograms; Proceedings of 5th IBM Medical Symposium. 1963, pp. 213-224

Lodwick, G., Haun, C., Smith, W. et al.; Radiology; 1963, 80, pp. 273-275


Lusted, L.B.; Introduction to Medical Decision Making; Springfield, Ill.: Thomas, 1968


Maier, N.R.F.; Reasoning in Humans I.; One Direction; Journal of Comparative Psychology, 1938, 10, 115-143

Maier, N.R.F.; Mechanization in Problem Solving: The Effects of Einstellung; Psychological Monographs; 1942, 54, 1-15


McGuire, C.H.; Personal communication; February 25, 1972


McNeil, B.J., Keeler, E., and Adelstein, S.J.; Primer on Certain Elements of Medical Decision Making; New England Journal of Medicine, 1975, 293, 211-215


Mead, G.H.; Mind, Self and Society; Chicago: University of Chicago Press, 1934

Medawar, P.B.; Induction and Intuition in Scientific Thought; London: Methuen, 1969

Medawar, P.B.; The Art of the Soluble; London: Methuen, 1967

Meehl, P.E.; Clinical Versus Statistical Prediction; Minneapolis: University of Minnesota Press, 1954

Meehl, P.E.; Psychodiagnosis: Selected Papers; Minneapolis: University of Minnesota Press, 1973

Miller, G.A.; The Magical Number Seven, Plus or Minus Two: Some Limits on our Capacity for Processing Information; Psychological Review, 1956, 63, 81-97


Miller, G.E. (Ed.) Teaching and Learning in Medical Schools; Cambridge, Mass.: Harvard University Press, 1962

Molidor, J.B. and Wagner, C.; DOCS Users Manual; OMERAD, Michigan State University, 1974
Molidor, J. B. and Wagner, C. DOCS Data Processing Manual; OMERAD, Michigan State University, 1974


Morgan, W. L., and Engel, G. L.; The Clinical Approach to the Patient; Philadelphia: Saunders, 1969

Nash, F. A.; Differential Diagnosis: An Apparatus to Assist the Logical Facilities; The Lancet, April 24, 1954

Neisser, U.; Cognitive Psychology; New York: Appleton-Century-Crofts, 1967


Neufeld, V. R., and Barrows, H. S.; The McMaster Philosophy: An Approach to Medical Education; Journal of Medical Education, 1974, 49, 1040-1050

Neufeld, V. R., Norman, G. R., Feightner, J. N., and Barrows, H. S.; Clinical Methods of Medical Students II a Cross-Sectional Analysis; McMaster University, Hamilton, Ontario, Canada, 1976

Neutra, R. R.; Lecture notes on Bayes' theorem with multiple tests and multiple diagnoses; Unpublished Manuscript, 1976. (Available from Department of Preventive and Social Medicine, Harvard Medical School 641 Huntington Ave., Boston, Mass. 02115)


Neyman, J.; Public Health Reports; 1947, 62, pp. 1449-1456

Norman, D. A.; Memory and Attention: An Introduction to Human Information Processing; New York: Wiley, 1969


Oskamp, S.; Overconfidence in Case Study Judgments; Journal of Consulting Psychology; 1965, 29, pp. 261-265


Overall, J., and Williams, C.; JAMA; 1963, 183, pp. 307-313

Overall, J., and Williams, C.; Medizinische Dokumentation; 1961a, 5, pp. 51-56

Overall, J., and Williams, C.; Medizinische Dokumentation; 1961b, 5, pp. 78-80


Price, R.B., and Vlahcevic, Z.K.; Logical Principles in Differential Diagnosis; Annals of Internal Medicine, 1971, 75, 89-95

Radford, J.; Reflections on Introspection; American Psychologist, 1974, 29, 245-250

Raiffa, H.; Decision Analysis: Introductory Lectures on Choices Under Uncertainty; Reading-Mass.: Addison-Wesley, 1968


Rimoldi, H.J.A.; A Technique for the Study of Problem Solving; Educational
and Psychological Measurement, 1955, 15, 450-461

Rimoldi, H.J.A.; The Test of Diagnostic Skills; Journal of Medical Education, 1961, 36, 73-79

Rimoldi, H.J.A.; Evaluation and Training of Clinical Diagnostic Skills; Chicago: Psychometric Laboratory, Loyola University, 1963, (Rep. No. 41)


Robinson, S.; Clinical Inference in Medical Education: A Look at Teaching and Evaluation Problem Solving Skills; Office of Medical Education, University of Arizona College of Medicine, Office Paper No. 9, 1975

Rokeach, M.; The Open and Closed Mind; New York: Basic Books, 1960


Rubel, R.A.; Decision Analysis and Medical Diagnosis and Treatment; Ann Arbor, Mich.: University Microfilms, 1970


Scheffe, H.; The Analysis of Variance; New York: Wiley, 1959


Schwartz, S.H., and Simon, R.I.; Studying Information Processing and Decision
Making in Medical Diagnosis; In H. Wesley (Ed.), Health Research and the Systems Approach; Detroit: Wayne State University Press, 1972


Sedelow, S. Y., and Sedelow, W. A. Jr.; Language Research and the Computer; University of Kansas, 1972

Senior, J.R.; The Development and Validation of a Computer-Based System for Testing and Teaching Clinical Competence; Philadelphia: National Board of Medical Examiners and American Board of Internal Medicine, 1974


Sherman, H., and Komaroff, A.L.; Ambulatory Care Project; (progress Rep. 11A). Boston: Beth Israel Hospital, 1974

Shulman, L.S.; Seeking Styles and Individual Differences in Patterns of Inquiry; School Review, 1965, 73, 258-266


Simon, H.A.; How Big is a Chunk?; Science, 1974, 183, 482-488


Slovic, P.: Personal communication; October 17, 1974

Slovic, P., and Lichtenstein, S.: Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment; Organizational Behavior and Human Performance, 1971, 6, 649-744


Sutcliffe, J.P.: On the Role of "Instructions to the Subject" in
Psychological Experiments; American Psychologist, 1972, 27, 755-758


Taylor, T.; J. Roy Coll. Phycns.;

Takahashi, K.; Logic of Diagnosis and Its Processing by Computer With Respect to Congenital Heart Disease and Brain Tumors; Proceedings Automated Data Processing in Hospitals, Elsinore, Denmark, 1966, pp. 477-498

Templeton, A., Lehr, J., and Simmons, C.; Radiology; 1966, 87, pp. 658-670

Templeton, A., Simmon, C., and Lehr, J.; American Journal of Roentgenology; 1968, 102, pp. 865-874


Thomas, C. L., (Ed.); Tabors Cyclopedia Medical Dictionary; F. A. Davis Company, 1973

Thomas, J. C.; An Analysis of Behavior in the Hobbits-Orcs Problem; Cognitive Psychology, 6, 257-269, 1974

Thorndike, E.L.; Fundamental Theorems in Judging Men; Journal of Applied Psychology, 1918, 2, 67-76

Toronto, A., Veasy, L., and Warner, H.; Progress in Cardiovascular Discovery; 1968, 102, pp. 797-803


Tuddenham, W.J.; The Use of Logical Flow Charts as an Aid in Teaching Roentgen Diagnosis; American Journal of Roentgenology, Radium Therapy, and Nuclear Medicine, 1968, 102, 797-803


Vinsonhaler, J. F., Wagner, C. C., et al.; Technical Manuals for the Basic Information Management System (BMIS); The Institute for Research on Teaching, Michigan State University, 1977


Wallace, H.A.; What is in the Corn Judge's Mind?; Journal of the American Society of Agronomy, 1923, 15, 300-304


Warr, P.; Thought and Personality; Baltimore: Penguin, 1974


Ways, P.O., Baker, T., Finkelstein, P., Fiel, N.J., and Jones, J.W.; The Problem-Oriented Record: A Self-Instructional Unit; East Lansing: College of Human Medicine, Michigan State University, 1972

Ways, P.O., Loftus, G., and Jones, J.M.; Focal Problem Teaching in Medical Education; Journal of Medical Education, 1973, 48, 565-571


Wiggins, N., and Wiggins, J.S.; A Typological Analysis of Male Preferences for Female Body Types; Multivariate Behavioral Research, 1969, 4, 1, 89-102

Wilson, J.W.; Generality of Heuristics as an Instructional Variable; Unpublished doctoral dissertation, Stanford University, 1967


Wortman, P.M.; Cognitive Utilization of Probabilistic Cues; Behavioral Science, 1970, 15, 329-336

Wortman, P.M.; Medical Diagnosis: An Information-Processing Approach; Computers and Biomedical Research, 1972, 5, 315-328


Youssef, Z. I.; The Effects of Cascaded Inference on the Subjective Value of Information; Organizational Behavior and Human Performance; 1973, 10, pp. 359-363

Zemper, E.D.; CAI in Medical Education: Perspectives and Potentials at
Michigan State University; OMERAD Office Paper No. 1. Michigan State University, 1973

8.0 GLOSSARY OF TERMS

This portion of the BMIS Technical Manual is not presently available.
9.0 INDEX

This section of the BMIS Technical Manual is not presently available.