PREMEDICAL EDUCATION AND PERFORMANCE ON MEDICAL TASKS: A COGNITIVE APPROACH

by

VIVIAN MEDLEY-MARK

A THESIS SUBMITTED TO THE FACULTY
OF GRADUATE STUDIES AND RESEARCH,
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF MASTER OF ARTS

DEPARTMENT OF EDUCATIONAL PSYCHOLOGY AND COUNSELLING
McGILL UNIVERSITY, MONTREAL

ISBN 0-315-38293-7

PREMEDICAL EDUCATION AND PERFORMANCE ON MEDICAL TASKS

ACKNOWLEDGEMENTS

The author would like to acknowledge a number of people who were influential in the completion of this thesis. To begin, I would like to thank the students who volunteered their valuable time in order to participate in this study. I would also like to acknowledge Dr. Yogesh Patel and Dr. Hugh Scott who constructed the clinical cases. I'm not sure the students who participated in this study appreciated their efforts, but I most certainly did. A special thanks, as well, to Dr. Mylene Trottier for her expertise in explaining the medical aspects of the clinical cases and in evaluating the accuracy of students' pathophysiology protocols.

I also wish to thank my advisor, Vilma Patel, for her assistance in the planning and organizing of this study and, additionally, for the opportunities she created for me during the past three years.

To my team of dedicated editors, Tony Mark, Lorry Coughlin, and Alenoush Saroyan, I thank you for your time, effort, and invaluable comments on all sections and versions of this thesis. As well, I would like to acknowledge Josie Gambino and Sophie Pelletier for transforming my abstract into a resume.

Finally, to all of my fellow 'academia nuts', too many to mention, thanks for the support and encouragement over the past year. I owe you one!

ABSTRACT

This study was motivated by a perceived deficit in two research areas within Medical Education. One line of research evaluates the academic performance of medical students from varying premedical backgrounds on department-generated and standardized examinations. A second approach examines expert-novice differences in comprehension and problem-solving through the use of complex tasks and detailed methods of analyses. To date, however, these tasks and analyses have not been used to study medical students of varying premedical backgrounds. This study was designed to address this issue.

First year medical students with three types of premedical backgrounds read two clinical texts describing a patient problem, recalled them in writing, explained the underlying pathophysiology, and provided diagnoses. Detailed analyses of subjects' protocols are presented. In general, group differences were found in case representation and interpretation, suggesting a need for the continuation of this line of research and directions for future research.

RÉSUMÉ

Cette recherche a été entreprise pour palier à un déficit perçu dans deux secteurs de recherche médicale. Un de ces secteurs se préoccupe des étudiants en médecine ayant reçu différents types de formation prémédicale. Il évalue le rendement académique de ces étudiants lors d'examens donnés par le département de même que lors d'examens généraux standardisés. Le deuxième secteur de recherche étudie les différences entre experts et novices dans le domaine de la compréhension et de la résolution de problèmes grâce à l'usage de tâches complexes et de méthodes d'analyse détaillées. Toutefois, jusqu'à présent, ces tâches et analyses n'ont pas été utilisées pour étudier les étudiants en médecine possédant diverses formations prémédicales. Cette recherche examine précisément cette dernière situation.

Des étudiants en première année de médecine, divisés selon trois types de formation prémédicale, lurent deux textes cliniques décrivant le problème d'un patient. Leur tâche consista a réécrire le problème de mémoire, à en expliquer la pathophysiologie sous-jacente et à fournir un diagnostique. Des analyses détaillées des protocoles des sujets sont présentées. En général, des différences entre les groupes ont été obtenues pour ce qui est du rappel et de l'interprétation du cas. Ces résultats suggèrent la nécessité de continuer ce type de recherche ainsi que des directions vers lesquelles ces recherches pourront être faites.

TABLE OF CONTENTS

LIST OF TABLES

TABLE PAGE

LIST OF FIGURES

FIGURE PAGE

CHAPTER I

REVIEW OF THE LITERATURE

Introduction

Much of the research in the professional domain of Medical Education has been concerned with performance; the performance of students prior to, during, and subsequent to medical school. These investigations have followed either one of two approaches. The more traditional approach is grounded in the theory and methods of evaluation. A more recent approach has its roots in cognitive science These two approaches have been carried out independently, but a close examination of the goals and methodologies of each suggests that each can benefit from the other, with the end result being an enhanced understanding of performance in the domain of medicine. The motivation for the study presented here was derived from this perception.

Generally, this study draws upon the unique insightfulness of evaluation researchers regarding the need to characterize students' premedical background when considering subsequent medical performance, and the cognitive scientists' awareness of the need to examine the cognitive processes underlying the performance of complex tasks in addition to examining the accuracy of that performance. The strengths of the two approaches are combined to address, in a preliminary fashion, two important issues. The first issue concerns the influence of type of premedical education on performance on medical tasks, while the second issue concerns the nature of the cognitive processes of novices in complex tasks. To address these issues the comprehension and problem solving skills of three groups of entering medical students, from varying premedical backgrounds, were assessed using methods derived from cognitive science.

The Evaluation Approach to Performance Assessment in Medicine

At one time, students enrolled in medical schools constituted a fairly homogeneous population. They had completed high school and a four-year undergraduate degree in which they had majored in one of the basic sciences. The students who have been accepted into medical school over the past few decades, however, and who continue to be admitted, have a wide and varied premedical academic history This change to a heterogeneous population of medical students has been the result of two developments within medical education One development has been a growing dissatisfaction with the notion that students who plan to make a career out of medicine should have a strong background in the sciences Some researchers continue to favor this type of undergraduate premedical background (e g , Vaisrub, 1978) because of its perceived relevance to a medical career while other researchers advocate the acceptance of students who are broadly educated in the social sciences or humanities (e.g., Pellegrino, 1980; Thomas, 1978). This issue continues to be debated but, in the interim, students with nonscience majors presently comprise a portion of the once homogeneous population of natural science graduates within medical school (e.g., Thomae-Forgues and Erdmann, 1980).

A second development which has served to modify the population of students in certain medical schools has been a change in the medical curriculum itself. These changes have included an acceleration of premedical programs, medical programs, and acceleration through the integration of these two programs. The specifics of these changes will be outlined and placed into an historical perspective in the next section.

Suffice it to mention at this point that these curricular changes permitted the acceptance of students directly from high school into medical school, thus resulting in the creation of an additional group of students with a unique type of premedical background.

The acceptance of such a wide variety of students into medical school has spurred curiousity and concern among researchers and educators alike regarding the capabilities of these new breeds of students, in comparison to traditional students. Both prior to and during medical school, students are exposed to a variety of episodic examination procedures designed to determine their qualification for a course grade, advancement to the next academic year, and ultimately, for graduation and an M.D. degree. Whether these examinations be faculty-generated or extramural, they share the goal of determining levels of competency or qualification at a precise point in the preparation for a medical career. The results of these examinations served as a starting point for researchers in their quest to evaluate the academic ability of medical students from nontraditional premedical backgrounds, by providing a means of comparison with students from traditional, science-oriented premedical backgrounds.

An Historical Overview of the Emergence of Accelerated Programs in Medicine

Traditionally, the first stage in the professional education of physicians has involved four years of medical school subsequent to the completion of a Baccalaureate degree. In 1970, however, the Carnegie Commission on Higher Education published a report advocating accelerated medical education (The Carnegie Report, 1970). This proposal was motivated by two objectives; first, to reduce the financial expenditures incurred by students, universities, and supportive agencies in the training of physicians, and second, to increase the existing pool of practicing physicians. The system which was favored by the Commission was a four-year undergraduate degree followed by a three-year medical school program with a concomitant revision of curriculum content (Daubney, Wagner, and Rogers, 1981). In its most basic form, however, accelerated medical programs were in existence approximately four decades earlier. During World War II, there was an immediate need for an increase in the number of physicians in the armed forces (Matlack, 1972). In response to this need, a number of medical schools in both Canada and the United States adopted what has since been referred to as the Compression approach (Swanson, 1972). This entailed the reduction of the number of calendar years required to complete an M.D. degree from four years to three years. This was accomplished by leaving the traditional medical curriculum intact but lengthening the academic year by reducing vacation periods. Thus, the content which was covered in medical schools remained unchanged and was simply compressed into three longer academic years. In addition, medical schools began admitting students into these accelerated programs every nine months, thus allowing two classes to graduate in one calender year. While this form of acceleration did indeed increase the number of graduating physicians, the design of the program placed a tremendous burden on the faculty and was discarded shortly after the war ended.

In 1971, in accordance with the desires of the Carnegie Commission, the implementation of a new medical program similar to that utilized during World War II was proposed (Blumberg, 1971). That is, the completion of the program would entail only 36 calendar months of study as opposed to the traditional 45 months. This would be accomplished chiefly by the reduction of vacation time between school years. Under this program, as with conventional medical programs, only one class would enter each calendar year. This is the primary distinction between the proposed accelerated medical program and that used previously. This difference eliminates much of the overload which was placed on the faculty in World War II. The primary advantage of this modified accelerated program, as espoused by Dr. Blumberg, is the students' younger age at graduation. This benefit would be manifested in three ways: earlier entrance into the workforce and thus an increase in mean man years of practice, fewer undergraduate medical expenses, and earlier access to a stable income. Further, it was predicted that the reduction in calender years would serve to increase physician supply within the decade without having to increase the actual number of entrants into medical school per year. With these advantages made explicit, in combination with a monetary incentive by the federal government to be provided to medical schools for each medical student who graduates in three years, many medical colleges began experimenting with three-year accelerated programs (e.g., Medical Education in the United States 1969-1970).

Approximately a decade previous to the Carnegie report and subsequent to its release as well, a number of medical schools experimented with a different form of acceleration. More specifically, while some schools began accelerating students' premedical education, others were integrating premedical and medical education. Both types of programs allowed an M.D. degree to be earned by students only six years after graduating from high school. Students following the traditional route in their attainment of an M.D. degree require eight years of study subsequent to high school; four years are engaged in an undergraduate premedical program followed by an additional four years in medical school. By allowing students to concurrently study courses in premedical and medical programs, there is a savings in the total number of calendar years it takes to become a physician. The accelerated programs which emerged have been variously referred to as premedical/medical programs, 6-year medical programs, combined medical programs, B.A./M.D. programs, and B.S./M.D. programs. Although they did not flourish in the seventies as did the accelerated 3-year programs, they also did not experience the same decline in popularity as did the 3-year programs (Daubney, et al., 1981). For example, at the height of the three-year program activity in 1973, 27% of the nation's medical schools had fully adopted or incorporated a three-year program into their curriculum. An editorial published six years later, however, reported the number of three-year program schools as constituting only 6% of the nation's medical schools (Beran, 1979). A factor contributing to this decline in popularity was the fact that the monetary incentive offered by the government to those schools which adopted three-year medical programs never materialized.

While many medical schools have adopted an integrated premedical-medical curriculum format, the specifics of the program adopted vary from school to school. For example, most of the medical schools which offer accelerated programs tend to accelerate students' premedical education and then deposit them in a conventional four-year medical program. This is true of the programs offered at Johns Hopkins University (Asper, 1964), Jefferson-Penn State University (Herbut, Sodeman, Conly, and Ascah, 1969), the Rensselaer Polytechnic Institute-Albany Medical College (Kanter, 1969), the Lehigh-Medical College of Pennsylvania (Pritchard, 1976), Boston University (Lanzoni and Kayne, 1976), and McGill University (Patel, Dauphinee, and Medley-Mark, 1984). The few exceptions to this procedure include the program at the University of Michigan (Campbell and DeMuth, 1976) and at the City College of New York (Gellhorne and Scheuer, 1978). These latter schools more literally combine premedical and medical programs by permitting students to simultaneously enroll in courses offered (indeed, required) by both programs.

One commonality between most programs, in addition to the fact that there is a reduction in calendar time between high school and the award of an M.D. degree, is that a student continues to fulfill the requirements of an undergraduate degree in addition to an M.D. degree. However, not all universities concur with the type of undergraduate degree students must successfully complete. Some schools have adopted the philosophy that medical students should be permitted the opportunity to develop their own personal philosophy through a liberal arts education (Pritchard, 1976). Consequently, these schools (e.g., Lehigh-Medical College of Pennsylvania, Boston University, Universty of Michigan and Johns Hopkins University) offer students a combined B.A./M.D. program. Other medical schools (e.g., City College of New York, the Rensselaer Polytechnic Institute-Albany Medical College and Jefferson-Penn State University) are more traditional in their belief that students should be thoroughly educated in the sciences and thus award students a B.Sc. degree.

Finally, while the overwhelming majority of accelerated programs reduce the premedical-medical years from eight to six, Jefferson-Penn State University and McGill University have further reduced this time lag to five years. These two programs, however, are very different. While Jefferson-Penn State University requires students to complete a B.Sc. degree, McGill University does not. The accelerated medical program offered by McGill University, referred to as the Medical-Prepatory Program (Med-P program), accepts students who hold only a College d'Enseignment General et d'Enseignement Professionnel (CEGEP) certificate. This is comparable to the first year of Junior College in the United States. The first year of the Med-P program involves some basic science instruction. The required courses taken during this year are identical to those offered to students enrolling in their first year of McGill's B.Sc. program. Following the successful completion of the Med-P program (i.e., maintaining a Grade Point Average of 3.5) these students proceed with the regular four-year medical program.

To conclude, this section has outlined two forms of curriculum change which have occurred in medical programs over the last few decades. These include an acceleration of medical programs primarily through the reduction of vacation time or through the integration of premedical and medical programs, and an acceleration of premedical programs. Both modifications reduce the traditional time period between graduation from high school and the completion of an M.D. degree. One of the primary consequences of these new medical programs has been a diversification of the premedical backgrounds of students in medical school.

The Evaluation of Accelerated Medical Programs

The most common means of evaluating the success of accelerated medical programs, whether they be the three-year medical program, combined premedical/medical progam, or accelerated premedical program, has been to compare the academic performance of these students with those enrolled in the same medical school but in the traditional program. More specifically, the criteria for comparison have included grade-point-averages (GPA), scores on standardized tests such as the Medical College Admissions Test (MCAT) and National Board of Medical Examiners (NBME) test, receipt of various awards and honors, and rate of attrition. As this list of criteria imply, the majority of studies have been concerned with the academic performance of students in accelerated programs during their undergraduate medical years; few studies have examined post-undergraduate performance (i.e., at the resident or intern level). Since the population of students participating in the present study includes students who are enrolled in a premedical/medical progam, the remaining discussion will be restricted to studies which have evaluated students in this type of accelerated program at the exclusion of three-year accelerated medical programs.

The few published studies which have examined students' performance on achievement tests (i.e., nonstandardized tests) during their undergraduate medical education, have not reported statistically significant differences between accelerated students and their nonaccelerated counterparts. For example, students enrolled in the Jefferson-Penn State five-year accelerated program obtained equivalent yearly and combined G.P.A.'s as those students enrolled at the same school but in the traditional medical program (Herbut et al., 1969). Another study (Lanzoni and Kayne, 1976) reported that the medical clerkship grades (i.e., the composite score of interns', residents', and physicians' subjective evaluation of students medical knowledge, motivation, and behavior with patients) of accelerated and traditional students were comparable.

Studies which have compared MCAT scores of students enrolled in accelerated and nonaccelerated programs have yielded somewhat mixed results. This test is comprised of many subtests, including Verbal Ability, Quantitative Ability, Science, and General Information. Students enrolled in traditional medical programs write this test in their final undergraduate year, before they enter medical school. The results of this test constitute one of the criteria used by the Medical School Admissions Office in deciding whether or not a candidate should be admitted into a program. Students enrolled in accelerated programs, however, usually write this test approximately one year (Herbut et al., 1969) or two years (Lanzoni and Kayne, 1976) after they have been accepted into this program and the results of this test are therefore not used as admission criteria.

Herbut et al. (1969) found through comparison that accelerated students scored significantly higher on the Quantitative Ability and Science subtests of the MCAT, but were comparable to traditional students on the Verbal and General Information subtests. Students enrolled in the six-year combined Liberal-Arts-Medical program at Boston University were initially evaluated three years after the implementation of the program. At that point in time, the accelerated students and their nonaccelerated counterparts were found to be comparable on all four subtests of the MCAT (Keefer, 1964). However, a study published approximately one decade later, comprised of the first three graduating classes of this B.A./M.D. program, revealed that these students were superior to traditional students on the General Information, Quantitative, and Science subtests of the MCAT (Lanzoni and Kayne, 1976). This finding has since been substantiated in the same medical school using an additional two groups of graduates (Blaustein and Kayne, 1980). Despite these mixed results, the findings of these studies suggest that students in accelerated programs are, minimally, as competent as students in traditional programs and, maximally, superior in some respects, as measured by the MCAT.

All students who are awarded an M.D. degree must first pass the National Board of Medical Examiners (NBME) test at a recognized level. This standardized examination is composed of three parts, each designed to assess students' knowledge in different domains and written at different points in one's medical education (Hubbard, 1978).

Part I of the NBME is a multiple choice test which is designed to measure a student's knowledge in the basic medical sciences including anatomy, biochemistry, microbiology, pathology, pharmacology and, more recently, the behavioral sciences. This part of the examination is typically written once the student has completed the second year of a medical program; that is, upon completion of basic science instruction.

Part II of the NBME is also in a multiple-choice format and is designed to evaluate students' knowledge in the clinical sciences, including internal medicine, obstetrics and gynecology, pediatrics, preventive medicine and public health, psychiatry, and surgery. This part of the exam is typically written when students have completed their clinical instruction.

Although both Part I and Part II of this exam are comprised of a number of component subjects, they are scored as multidisciplinary exams. Consequently, students receive a total score on these two parts of the NBME as opposed to a grade for each of the subject-matter areas.

Part III of the NBME evaluates students' clinical competence. The Patient Management Problem (PMP) approach which is used to accomplish this, is designed to simulate a realistic clinical situation in which a patient presents with a limited array of symptoms and some information regarding his previous medical history. Students must diagnose the patient via requests for laboratory studies and diagnostic procedures and then make some decisions regarding therapy and patient management. This part of the examination is typically taken once students have graduated from medical school.

There are a limited number of published studies which document the performance of accelerated students versus nonaccelerated students on all or portions of the NBME test. While most studies have reported that accelerated students receive higher total scores on Part I (e.g., Kanter, 1969) and Part II (e.g., Herbut et al., 1969) of this test than do regular students, only one school found these differences to reach statistical significance (Blaustein and Kayne, 1980; Lanzoni and Kayne, 1976). These latter studies additionally reported statistically significant differences in test scores for these two groups of students on Part III of the NBME; these differences were in favor of the accelerated students.

As a more indirect measure of scholastic ability, accelerated and nonaccelerated students have been compared on graduation honors, election to Alpha Omega Alpha, and achievement of specialty board certification. While this comparison has been utilized in only one study (Lanzoni and Kayne, 1976), thus limiting the generalizability of the results at present, nonsignificant results were found on all three measures.

Another factor which has been used in the comparison of B.S./M.D. students and regular students is attrition rate. While few studies have made this comparison directly (e.g., Blaustein and Kayne, 1980; Lanzoni and Kayne, 1976), the reported attrition rates for accelerated students appears to be much higher than for traditional students (Daubney et al., 1981). Using an operational definition of attrition to include both withdrawal from a program (for any reason) and failure to maintain the prescribed pace of study (i.e., having to repeat an academic year), attrition rates for B.S./M.D. students has been reported to be as high as 36% (Campbell and DeMuth, 1976). A variety of causes for student attrition have been explicated including a change of career plans, academic difficulties, personal problems, and a loss of enthusiasm for acceleration. With the exception of a few studies (e.g., Flair, 1969; Kanter, 1969), however, most have neglected to detail the percentage of these students who withdrew for each of these specific reasons. This is a potentially important issue because of the implications each factor carries with it. Students enrolled in accelerated programs must choose the career of physician by the time they complete high school. If a large percentage of students withdraw from accelerated programs as compared to regular programs, because they have made a change in their career goals, then perhaps such a major decision is better made once students have more thoroughly explored other possible career opportunities.

Another method of evaluating accelerated programs has been to survey the students themselves and inquire about their perceptions of the program. A study conducted at the Jefferson-Penn State Universities, for example, surveyed students at various points in the five-year B.S./M.D. program regarding their satisfaction with the program (Grossman, Conley, Menduke, and Graff, 1972). While all students felt the program should be continued, sixty-four percent felt that it was in need of some modification. When asked if they would have preferred a five-year program or a six-year program, slightly over one-half of the students stated that they would have preferred the six-year program. However, if given the choice to enroll in the present five-year program or attend four years of college, seventy-six percent stated that they would once again choose the five-year program. A similar line of research has been conducted at McGill University (Patel, Dauphinee, and Medley-Mark, 1984) involving a comparison of accelerated and nonaccelerated students. While approximately seventy-five percent of the nonaccelerated students viewed their premedical education as an advantage, only fourty-four percent of the accelerated students held this view. The disadvantages outlined by these students included the belief that they were immature during the clinical phase of their studies and, additionally, that they had received a narrow education. The advantages espoused by these students revolved around their younger age at graduation. This factor had advantages in terms of family planning and for post-medical studies. The study conducted by Grossman, et al. (1972), however, found that students in the accelerated program felt it had encouraged them to mature faster and that they had not missed much in terms of educational or life experiences. The differences found in these two studies may, however, simply be a result of time at which the questionnaire was administered; the study at McGill University was retrospective in nature (i.e., students were surveyed after they had graduated), while the study at Jefferson University was conducted while students were enrolled in the program.

The cumulative findings of the studies reported here suggest that students in accelerated medical programs are comparable, and in some respects superior to students in traditional medical programs as evidenced by direct indices of academic ability on both standardized tests (e.g., MCAT, NBME) and nonstandardized tests (e.g., medical clerkship grades, undergraduate GPA in medical school), as well as indirect indices of academic ability (e.g., rate of attrition, receipt of various awards and honors).

The Cognitive Approach to Performance Assessment in Medicine

The performance of individuals in the domain of medicine has also been explored using methods derived from cognitive psychology. Generally, this has involved the use of the contrastive method in which the performance of groups of individuals who are at either extreme of an expert-novice continuum are compared on a specific task. In the context of medicine, the novice population has typically consisted of medical students while the expert population has usually been comprised of post-graduates of medical school (e.g., interns, residents, practicing physicians). The tasks which have been employed vary, but typically revolve around a patient problem (e.g., in the form of a written clinical text or an interview with a simulated patient). The goal of this research is to characterize differences in behavior between novices and experts within a particular domain in the hopes of providing some insight into the underlying processes which must somehow be acquired and/or altered by the novice en route to developing expertise (Feltovich, 1983).

The contrastive approach has been used to study expertise in a wide variety of content domains. While it was originally used to study expertise in the game of chess (Chase and Simon, 1973; deGroot, 1965), it has subsequently been used to study expert-novice differences in other games including the board games of GO (e.g., Reitman, 1976) and GOMOKU (Eisenstadt and Kareev, 1975, 1977), and the card game of bridge (e.g., Charness, 1979). In addition to these nonverbal tasks, the contrastive method has been used to study expert-novice differences on tasks which require the input and output of complex verbal information. The content domains which have been investigated include physics (e.g., Chi, Feltovich, and Glaser, 1981; Larkin, McDermott, Simon, and Simon, 1980), the social sciences (Voss, Green, Post, and Penner, in press), baseball (e.g., Chiesi, Spilich, and Voss, 1979), geometry (Greeno, 1978), and medicine (e.g., Feltovich, 1981; Patel, Groen, and Frederiksen, 1986). These studies (of both verbal and nonverbal tasks) have been concerned with identifying differences between experts and novices in knowledge representation and/or problem solving ability. In terms of the present study, the expert-novice research which has been conducted in the domain of medicine is of primary relevance and will, consequently, serve as the focal point for the remainder of this section.

An Historical Overview of Cognitive Psychology in Medicine

In 1969, a research project was initiated at Michigan State University (Elstein, Shulman, and Sprafka, 1978). This 'Medical Inquiry Project', which spanned five years, represented the introduction of modern cognitive psychology into medical education and proved to be the seminal work on the cognitive processes involved in medical problem solving. The results of this study were subsequently replicated in the Clinical Methods Study conducted at McMaster University (Barrows, Neufeld, Feightner, and Norman, 1977). Additionally, this study extended the work of Elstein and colleagues (1977) by including populations of medical students, as well as internists. The significant finding from this study was the similarity in clinical reasoning noted in first year medical students through to practicing physicians. That is, the number of competing hypotheses or possible diagnoses considered for a clinical case and the specific point in time during the clinical interview at which these hypotheses were generated were found to be comparable across groups, irrespective of level of medical expertise. The one differentiating feature between these diverse groups was the actual content of the diagnostic hypotheses. These results were interpreted as analogous to some of those found in the study of expertise in the board game of chess (Chase and Simon, 1973; deGroot, 1965). That is, chess masters could not be differentiated from chess novices on the basis of the number of moves considered nor on the time at which they considered these moves. The single factor which did differentiate these groups was the specific nature of the moves under consideration.

The chess studies also examined differences between experts and novices in their internal representation of information, an aspect which was not included in the Clinical Methods Study. More specifically, chess masters were found to be superior to novices in their recall of typical mid-game boards (as measured by the recall of individual chess men) but, when asked to recall atypical boards (i.e., those boards on which the positioning of individual chess men were random, but within the legal constraints of the game), there was no difference in performance between experts and novices. Due to the previous similarity in results between the studies conducted in medicine and chess, a series of studies were conducted both at McMaster University (e.g., Muzzin, Norman, Jacoby, Feightner, Tugwell, and Guyatt, 1982; Norman, Jacoby, Feightner, and Campbell, 1979) and, subsequently, at the University of Limburg (e.g., Claessan and Boshuizen, 1985) in an attempt to more thoroughly generalize the findings of deGroot (1965) and Chase and Simon (1973) to the domain of medicine. These replications, however, did not meet with complete success. More specifically, the superior memory performance of experts over novices for typical stimuli was replicated in only one analogous study in medicine (Norman, et al., 1979) Also, this study replicated the chess finding that the superior performance of experts over novices disappeared when they were requested to recall atypical stimuli. In two other studies (Claessan and Boshuizen, 1985; Muzzin, et al., 1982), however, no significant differences were found between medical students and practicing physicians in terms of the number of items recalled from a 'typical' clinical case.

A number of reasons have been postulated for the occurrence of these mixed results, including a confusion on the part of researchers in Medical Education regarding their interpretation of Chase and Simon's definition of typicallity and atypicallity (Coughlin and Patel, 1985) as well as the methods of analyses which were employed (Groen and Patel, 1985). It is this latter issue which is of particular interest. Chess is essentially a nonverbal task and, consequently, the data analysis and the methods used to characterize the stimulus materials are quite simplistic in nature. Conversely, both the input and output for medical problem solving involves complex verbal information. Thus, a more sophisticated method of data analysis may be necessary. To evaluate this possibility, a reanalysis of some of the recall protocols collected by Muzzin et al. (1982) was undertaken at McGill University by Patel and Frederiksen (1984a). The method of data analysis which they used is referred to as propositional analysis; a technique developed to facilitate the the identification of the cognitive processes underlying text comprehension (Frederiksen, 1975; Kintsch, 1974). This reanalysis was successful in yielding expert-novice differences and the results more closely resembled those obtained in the board game of chess. More specifically, differences were found between experts (consulting internists) and novices (second-year medical students) in terms of their representation of typical and atypical clinical cases as measured by the percentage of text-based propositions which were inferred and recalled, respectively.

The methods of discourse analysis employed by Patel and Frederiksen (1984a) in their reanalysis of the Muzzin et al. (1982) data have been used in a number of subsequent studies in medicine and have become the primary means of assessing expert-novice differences in the comprehension of medical texts. The paradigm employed typically involves requesting subjects to read a clinical text (often within a set time limit) and then to recall that text. Indices of comprehension processes include the percent of propositions which are recalled and inferred as well as the relevancy of these propositions to the correct diagnosis. In addition, the assessment of differences in medical problem solving has been accomplished by requesting experts and novices to explain the underlying pathophysiology of a clinical text. This task enables one to study subjects' causal knowledge (Patil and Szolovits, 1981).

The comparative profile of novices emerging from studies which have used these paradigms includes the following general characteristics about their comprehension and problem solving abilities:

(1) novices and experts represent clinical cases differently; experts make more inferences than novices (indicating a higher level of abstraction) while novices recall more than experts (e.g., Patel, in press),

(2) novices are less adept than experts at discriminating relevant from nonrelevant information contained in a clinical text (e.g., Patel, Groen, and Frederiksen, 1986; Patel, in press),

(3) novices and experts have qualitatively different causal knowledge; the problem models constructed by novices are based on experiential or common sense knowledge while the problem models constructed by experts are based on domain-specific knowledge (e.g., Patel and Groen, 1986a).

(4) novices' knowledge tends to be fragmented whereas experts' knowledge is highly organized (e.g., Patel and Groen, 1986a).

(5) novices and experts represent problems differently; novices' problem representations are based on surface characteristics of the problem while the problem models constructed by experts are more parsimonious, concise, specific, and integrated (e.g., Patel and Groen, 1986b).

An additional modification of the original deGroot (1965) and Chase and Simon (1973) paradigm which has been implemented by researchers in cognitive psychology as applied to medical education has been the nature of the populations which have been compared. More specifically, in addition to examining groups of individuals who are at either extreme of the expert-novice continuum, researchers have compared groups of individuals at differing points along the continuum, and groups of individuals at the upper-end of the continuum. With reference to the former modification, the combinations of populations which have been studied and contrasted include:

(1) first- and second-year medical students and physicians (e.g., Patel, HoPingKong, and Mark, 1984),

(2) first-, second-, and fourth-year medical students and physicians (e.g., Patel and Medley-Mark, 1986),

(3) laypersons, second- and fourth-year medical students, final year clinical clerks, first-year residents, family physicians and practicing dermatologists (Norman, Muzzin, and Rosenthal, 1985),

(4) second-, fourth-, and final-year medical students (Claessen and Boshuizen, 1985).

(5) first- and third-year medical students, second-year residents, and experienced physicians (e.g., LeClere and Bordage, 1984).

These studies have provided additional insight into the development of expertise in the domain of medicine and have indicated a non-monotonicity in learning (e.g., Patel, Groen, and Scott, in press).

A number of studies have also compared groups of individuals who are at the upper end of the novice-expert continuum (e.g., Kassirer, Kuipers, and Gorry, 1982; Miller, 1975; Muzzin et al., 1983). These studies represent an attempt to characterize differential abilities among groups of individuals who are expert in different aspects of one particular domain. For example, Joseph and Patel (1986) and Patel, Arocha, and Groen (1986) examined the problem solving processes of endocrinologists and cardiologists solving a problem both within and outside their area of expertise. While both groups of subjects may be considered experts, in the sense that they hold M.D. degrees, the physicians with subspecialty degrees in endocrinology are experts in understanding and diagnosing endocrine disorders whereas the physicians with subspecialty degrees in cardiology are experts in understanding and diagnosisng cardiac disorders. As a consequence of this differential expertise, these physicians may solve the two types of problems differently. The interesting facet of these studies is that they represent an emerging awareness among cognitive researchers of the need to characterize the populations under study in a more exact manner. Such studies also suggest that there may be unique characteristics of these populations at the upper end of the novice-expert continuum which are important to identify. In a similar vein, there may be unique characteristics of populations at the lower end of the novice-expert continuum which are important to identify. In the domain of medicine, however, this possibility has not been empirically investigated.

Glaser (1985) has noted that the increased understanding of expertise has led to an increased curiosity of how it is acquired. At present, however, our knowledge about novices in the domain of medicine is limited. The reason for this is two-fold. First, the populations which have previously been designated as novices have ranged from first year medical students (e.g., Patel, HoPingKong and Mark, 1984) through to internists (e.g., Elstein, et al., 1977). The only consistency between studies has been that the individuals selected for the novice group are those with less medical training than the individuals selected for the expert group. While this is acceptable, it may not yield substantial insight into the cognitive behavior of novices. Further, if we are to understand the true nature of expertise within a given field of study, it seems reasonable, in fact necessary, to begin with a thorough examination of 'true' novices, individuals at the outset of developing expertise. A second reason our knowledge of novices in the domain of medicine is limited is that the profile which has been developed is based on comparisons with experts. As the beginning of this chapter outlined, students entering medical school comprise a rather heterogeneous group in terms of their premedical academic history; students either hold B.Sc. degrees, B.A. degrees, or enter medical school directly from high school. However, all previous expert-novice studies have failed to consider the premedical backgrounds of the populations under study. As a result of these two conditions, our knowledge of novices may be, at best, more incomplete than realized and, at worst, misleading.

Rationale

As the preceeding pages have outlined, two research approaches have been used to study the performance of students in medical school. The evaluation approach has been concerned with identifying academic differences between populations of medical students who are equivalent in academic year but who differ in terms of premedical background. These studies have been conducted primarily for the purpose of program evaluation. The results of these studies, based on very global, evaluative-type measures, suggest that differences in academic performance in these populations do exist.

Medical students, equivalent in academic year, have also been studied in comparison to more advanced medical students and post-graduates of medical school (i.e., interns, residents, practicing physicians). The research approach that has been used is cognitively oriented and has as its goal the identification of the underlying cognitive processes which must be acquired by the novice (i.e., medical student) developing expertise. Complex tasks and methods of analyses have been the primary means for accomplishing this goal. While differences in performance have been noted, these studies have neglected to differentiate between populations of medical students in terms of their premedical background. Evidence from the evaluation research literature and more recently, research on the role of prior knowledge (e.g., Glaser, 1985) strongly suggests that premedical background is an important factor influencing students performance in medicine.

The purpose of the present study was to bridge the gap between the evaluation and cognitive research conducted in the domain of medicine by exploring the clinical case comprehension and problem solving abilities of students from various premedical backgrounds upon entry into medical school. The tasks and associated methods of analyses were borrowed from the cognitive science literature and were used in conjunction with a grouping variable borrowed from the evaluation research. This combined approach is unique and has the ability to enhance two ongoing lines of research within the professional domain of medicine.

CHAPTER II

METHOD

Subjects

A total of thirty students entering medical school participated in this study.*[1] The first step taken in the recruitment of these students was the attainment of a list of all of the students who had been accepted into McGill University's four-year and five-year medical programs effective September 1985. This list, as well as a list of the students who had successfully completed the first year of the five-year Medical-Preparatory program was obtained from the Medical Faculty's Admissions office. The experimenter then randomly selected students from these lists and telephoned them to inquire if they would like to participate in the study. The telephone conversations included a general description of the purpose and nature of the study and an estimate of the time involved. Emphasis was placed on the fact that participation was strictly on a voluntary basis and that refusal or acceptance to participate would in no way influence their subsequent studies at McGill University. The first ten subjects from each of the three lists of students who accepted were included in the study.

Thus, three groups of students, characterized by three different premedical backgrounds, participated in the study. At the time of testing, none of these students had received any formal medical training. Group 1 consisted of students (n=10) who had just completed College d'Enseignement General et d'Enseignement Professionnel (CEGEP) in which they had majored in Health Sciences. These students, subsequently referred to as **CEGEP** students, had been accepted into McGill's five-year, Medical-Preparatory program. Group 2 consisted of students (n=9)*[2] who had successfully completed the first year of the five-year Medical-Preparatory program, and entered the regular four-year medical program in September 1985 (approximately three months after testing). The students in this group are subsequently referred to as **Med-P** students. Finally, Group 3 consisted of students (n=10) who had completed a Bachelor of Science undergraduate degree. Six of these students majored in physiology, two in biochemistry, and one in neurophysiology. These students, subsequently referred to as **Degree** students, had been accepted into McGill's four-year medical program and are currently in the same classes as those students from Group two (i.e., the Med-P students).

Materials

The stimulus material consisted of one 20-page, self-contained booklet per student. The first page of each booklet contained a cover sheet (Appendix A) which described both the contents of the booklet and the corresponding instructions. This cover sheet was followed by one of two clinical texts, the details of which are described subsequently. This clinical text was accompanied by three separate sheets of paper headed with the following requests (1) Summarize the case, (2) Explain the case in terms of the underlying pathophysiology, and (3) Provide a diagnosis. All materials relating to one clinical text, including the text itself, were presented together and were followed by the second text and its accompanying materials. The two texts were used in counterbalanced order in all groups and, accordingly, two different types of booklets were constructed.

The two clinical texts outlined different clinical problems. Text 1 (Appendix B) described a patient with an endocrinology disorder. Text 2 (Appendix C) described a patient with a cardiac disorder. Each text was developed by a physician whose area of medical expertise lay in the specific nature of the clinical problem (i.e., an endocrinologist and a cardiologist, respectively). Both texts were derived from actual patient cases and described the history, physical examination findings, and some laboratory test results of the patient.

The two texts differed with respect to their surface-level text characteristics. The cardiology text (Text 2) was longer than the endocrinology text (Text 1) but had a lower propositional density (i.e., average number of propositions per segment). Table 1 outlines the exact surface-level text characteristics of these two texts.

Text 1: Hashimoto's Hypothyroidism.

Text 1, the endocrinology text, outlined the case of an elderly woman with Hashimoto's Hypothyroidism which had progressed to a myxedema pre-coma state. Hypothyroidism is a condition of deficient thyroid hormone secretion (Harrison, 1980). A number of reasons have been postulated for the occurrence of this condition which, depending on the specific cause of the disorder, is manifested at different points in one's life. For example, hypothyroidism may be present at birth or in early infancy if the thyroid gland has failed to develop or if there are inherited defects of thyroid hormone biosynthesis. Alternatively, hypothyroidism may be manifested in middle-age which is typically the case if the disorder has an auto-immune origin (also referred to as Hashimoto's hypothyroidism). Finally, hypothyroidism may appear at any age if the thyroid gland has been surgically removed or has been destroyed by radio-iodine therapy. Since the patient in this case had an enlarged thyroid gland (in fact, it was enlarged to twice the normal size), with a firm and irregular consistency, the most plausible form of hypothyroidism is Hashimoto's hypothyroidism.

TABLE 1

Textual Aspects of Stimulus Materials

ASPECT	HASHIMOTO'S HYPOTHYROIDISM (Text 1)	PERICARDIAL EFFUSION (Text 2)
Number of Words	250	368
Number of Segments	29	58
Number of Propositions	84	129
Propositional Density	2.90	2.22
Number of Relevant Propositions	32	48
Number of Nonrelevant Propositions	52	81
Percent of Relevant Propositions	38.10	37.21
Percent of Nonrelevant Propositions	61.90	62.79

A number of clinical features typical of hypothyroidism were embedded in this case. These include general symptoms arising from the involvement of the muscles, nervous system, cardiovascular system, and gastrointestinal system. Such general symptoms included cold intolerance and weight gain. The signs indicative of these symptoms are pale, dry, and coarse skin and hair. Symptoms which stem from nervous system involvement typically include slowness of thought, as indicated by slowness of speech, neuromuscular condition (indicated by delayed relaxation tendon reflexes), and drowsiness. Cardiovascular symptoms include angina of effort which is indicated by such signs as bradycardia (i.e., slowness of the heart, as indicated by a low pulse rate), evidence of ischaemic heart disease and sometimes pericardial effusion. Further, the heart sounds of a hypothyroid patient tend to be muffled. Gastrointestinal symptoms usually include constipation and hyponatremia or reduced serum sodium (due to the patient's inability to excrete free water by the kidneys). In addition, a complication frequently associated with profound hypothyroidism was evidenced by the patient in this case. More specifically, if hypothyroidism is left untreated, it may lead to myxoedema coma which is characterized by hypothermia (reduced body temperature). In this case, the patient had recently been prescribed a potassium iodide mixture to alleviate what had been diagnosed as chronic laryngitis. What was not made explicit in the text, however, was the fact that the patient had been misdiagnosed. Many patients with hypothyroidism experience a change in their voice. That is, due to the low levels of thyroid hormone, their voice becomes low or husky. Because the patient had not been diagnosed as having hypothyroidism, this quality of voice was misinterpreted as a sign of laryngitis. The medication which the patient was subsequently prescribed further decreased the already low levels of circulating thyroid hormone and resulted in a coma-like state which was the reason for the patient's admittance to emergency.

Text 2: Pericardial Effusion.

Text 2, the cardiology case, outlined the case of an elderly man with pericardial effusion. This is a condition in which the sac that surrounds the heart (i.e., the pericardium) is filled with fluid (Harrison, 1980). This fluid build-up results in constriction of the heart and restriction of its pulsation. When there is an extreme amount of fluid in the pericardium, cardiac tamponade ensues. This results in a fall of cardiac output and systemic venous congestion.

A number of clinical features which are commonly associated with pericardial effusion were embedded in this case. Many of these findings are due to the inefficiency with which blood circulates both to and from the heart. In this patient, the capacity of both the right side and left side of the heart was decreased. The right side of the heart is responsible for the uptake of blood once it has circulated through the body. However, the large amounts of fluid in the pericardium restricted or blocked the flow of blood and fluids to the right side of the heart. Consequently, there were abnormally large quantities of fluid in the body (edema), including the abdominal cavity (ascites), legs, presacrum, and scrotum. As well, the jugular veins were distended because of the blockage of flow to the right side of the heart.

The pulmonary symptoms which this patient manifested indicated impairment of the left side of the heart. The left side of the heart is responsible for the uptake and circulation of blood to the body once it has been purified by the lungs. However, the flow of blood and fluids to the left side of the heart was also restricted by the fluid accumulation in the pericardium. Consequently, the patient had difficulty breathing (dyspnea) on mild exertion (because of the fluid in his lungs) and by the time he presented to the hospital, he could no longer sleep in a horizontal position. This latter condition (orthopnea) is a direct result of the fact that when the patient lies down, the fluid in his lungs distributes over their entirety and more severly affects the respiratory cycle. Some findings which were indicative of these pulmonary symptoms included right pleural effusion (fluid build-up in the right lung) and partial atalectasis in the right lower lobe (collapsed right lung).

The fluid build-up in the pericardium was confirmed by the enlarged cardiac silhouette, the low voltage QR's and voltage fluctuation noted on the ECG readings (due to the electrical interference caused by the fluid), the faint heart sounds, and the absence of the apex beat. Finally, the clinical finding of pulsus paradoxicus was indicative of cardiac tamponade.

The patient additionally demonstrated signs of hepatic (liver) congestion and intestinal dysfunction (protein-losing gastroentropathy) as evidenced by a number of laboratory test results, including the levels of albumin, bilirubin, and urobilinogin. These readings suggested that the levels of blood proteins in this patient were low.

Procedure

Subjects were tested individually in one of McGill University's Centre for Medical Education offices. The instructions for completing the booklet (outlined on the coversheet) were first presented verbally by the experimenter to ensure that the correct procedure was fully understood and would be followed. Any uncertainty expressed by the students concerning the instructions was immediately clarified by the experimenter. Once the instructions had been discussed, students were left on their own to (1) read the initial clinical case, (2) summarize it in writing, (3) provide a documentation of the underlying pathophysiology of the case (that is, to explain the occurrence of the patient's symptoms in basic science terms), (4) provide a diagnosis, and (5) repeat these steps for the second clinical case. The order of text presentation was counterbalanced across students within each group, with the exception of Group 2 (the

Med-P group) in which there were only nine subjects. In this group, only four students received one order while the remaining five students received the alternate order.

Once students had completed the required task, the experimenter thoroughly debriefed them on the intended purpose of the study and answered any questions students had pertaining to the study.

CHAPTER III

METHOD OF ANALYSIS

Comprehension of Clinical Cases

Students' comprehension of the two clinical texts was assessed using methods of discourse analysis; a series of analyses which were developed to allow investigators to draw inferences about the cognitive processes that underlie the comprehension of verbal text. A number of systems currently exist but the two most comprehensive methods of discourse analysis, namely propositional analysis, are those developed by Kintsch (1974) and Frederiksen (1975, 1985). The basic tenet of these models is that a sentence is comprised of one or more sets of elements, referred to as propositions, which represent the semantic content of that sentence. Kintsch's (1974) model, while having the advantage of being relatively simple in nature is lacking in its degree of specificity and precision (Tierney and Mosenthal, 1982). The system developed by Frederiksen (1975) is more comprehensive in nature and includes both a semantic network and a logical system which relates sets of propositions. It is precisely this aspect of Frederiksen's method, in combination with the fact that it has been successfully adapted for the domain of medicine (e.g., Patel and Frederiksen, 1984a), which provides the justification for its use in the present study. The following is a description of how this analysis is conducted, with reference to the present study.

The first step in the analysis involves segmenting the stimulus materials (i.e., the two clinical texts) and subjects' summary protocols of these materials. This is accomplished by employing Winograd's (1972, 1983) system of clausal analysis; a form of analysis which has its roots in Halliday's (1967) systemic grammar. It is essentially a syntactic analysis and, as such, is concerned with the division of text into relevant syntactic units. Due to the wide range of complexity which can characterize what is typically identified as a sentence, these syntactic units are deemed a more manageable unit for the purpose of further analyses.

Within this system, a segment is defined as a clause or syntactic unit, typically containing a finite verb (either tensed or conjugated). This includes all major clauses (e.g., declarative, imperative, and interrogative) with their associated minor clauses. Additionally, one type of secondary clause is considered to be a separate segment, specifically, the bound adjunct. A bound adjunct is a clause which modifies another clause and is linked to that clause by means of an explicit binder. Clauses with nonfinite verbs (e.g., infinite, participle), including bound adjuncts, are not considered to be separate segments, but rather a part of a segment. For example, the sentence:

He then noted he was winded
after walking about 40 feet.

contains a secondary clause, a bound adjunct, but because the verb is infinite (e.g., walking), it is not considered a separate segment. Conversely, the sentence:

A month later she had been diagnosed
as having chronic laryngitis and was
prescribed a potassium iodide mixture
as an expectorant.

is considered to be two segments because the secondary clause or the bound adjunct contains a finite verb (e.g., prescribed).

In summary, segmentation is the preliminary phase in analyzing a piece of text. It is a means of dividing a text into convenient, well-defined, and theoretically motivated units (Dillinger, 1984) which serve as the input to the remaining two forms of analyses.

The second step in this analysis involves the semantic representation of each segment contained in the stimulus materials. This is accomplished using Frederiksen's (1975, 1985) method of propositional analysis. This representation allows one to see, at a precise level, how a subject's protocol is related to the presented text.

Propositions are numbered within segments and each proposition consists of a predicate, an argument, and a labelled relation linking the two. This is commonly referred to as a 'triple' and is defined as the smallest, identifiable semantic unit. According to Frederiksen's model, a predicate may be an action (e.g., diagnose), an object (e.g., mixture), or a relation which connects propositions (e.g., a conditional, temporal, or causal relation). Arguments may be case relations, such as patient (the patient of a processive action), agent (the agent of a resultive action), theme (the theme of a cognitive process) or result (the result of an action). Additionally, arguments may be identifying relations such as locative (the location of an action or object), tense (past, present, future), aspect (e.g., continuous), or modality (e.g., truth value). An example of this procedure, using segments number five and six from the Hashimoto's Hypothyroidism case (Text 1), is shown below.

5. A month later she had been diagnosed as having chronic laryngitis

PREDICATE	ARGUMENT
5.1 diagnose	REC:she, THM:(as)5.2, =TNS:past, ASPCT:comp;
5.2 have	ACT:laryngitis, ASPCT:cont;
5.3 laryngitis	ATT:chronic;
5.4 TEM:DIFF (1 month)	[5.1], [later]

6. and was prescribed a potassium iodide mixture as an expectorant.

PREDICATE	ARGUMENT
6.1 prescribe	REC:(and), OBJ:mixture, =TNS:past;
6.2 mixture	=ATT:iodide, ATT:potassium;
6.3 IDENT(as)	[6.2], [expectorant];

It should be noted that this system of propositional analysis allows for the embedding of propositions (e.g., proposition 5.2 is embedded within proposition 5.1), and the possibility of empty slots (e.g., the recipient in proposition 6.1). For a more complete and comprehensive description of the system, the reader is referred to Frederiksen (1985).

The third and final aspect of this analysis is referred to as recall analysis. This involves matching the segments from subjects' written protocols against the propositional content of the original stimulus text. Those segments from subjects' protocols which correspond exactly to the message base as defined in the original text are coded as recalls. Transformations made by subjects on any message base are coded as inferences. These transformations are specified by a set of rules developed by Frederiksen (1979). For example, if a segment in a subject's protocol read:

Blood pressure (B.P.) was 160/95.

and the corresponding segment in the original text was represented as:

PREDICATE	ARGUMENT
B.P.	=DEG:160/95, TNS:past;

then that segment from the subject's protocol would be scored as a recall because it contains every proposition that the text segment does. However, if the segment in a subject's protocol read:

Blood pressure (B.P.) was high.

then that segment would be scored as an inference because the subject has altered the degree value of the text proposition. These two response categories, recall and inference, are not distinct and a segment from a subject's protocol may be coded as both a recall and an inference. Additionally, a segment from a subject's protocol may be scored as a part-recall. This is typically the case when only the arguments of a text proposition are contained within a segment of a subject's protocol (i.e., the predicate or head element of a proposition is not included as part of the subject's text). For example, a segment in the stimulus text may contain the propositions:

PREDICATE	ARGUMENT
palpate	OBJ:beat, =TNS:past, NEG (not), MOD:QUAL (could);
beat	CAT:apex

while the corresponding segment in the subject's protocol reads:

The apex could not be palpated.

While the argument of the second text proposition has been retained in the subject's protocol, the head element (beat) has been omitted. The absence of this head element necessarily excludes the possibility of scoring the subject's protocol for the presence of a triple; a necessary condition for a recall or inference to be coded. Thus, this portion of the segment is scored as a part-recall.

In summary, subjects' comprehension of the clinical texts was analysed through the use of techniques of discourse analysis. These range from surface level clausal analysis, through propositional analysis and, finally, an analysis of the recalled, partially recalled and, inferred propositions which the subject produces.

In terms of collecting data for the comprehension analysis, the number of propositions which formed the basis for recalls, part-recalls, and inferences within each subject's summary protocol, for each of the two texts was calculated. These figures were subsequently transformed into mean percentages for the purpose of standardizing the results across the two texts. Each of these three responses were subsequently subdivided according to their relevance to the diagnosis.

Measures of Relevancy

All propositions in each of the two texts were categorized as either relevant or nonrelevant to the correct diagnosis. This was accomplished by requesting two experts to identify the relevant aspects of the text that would lead to an accurate diagnosis. This resulted in a comparable proportion of both relevant and nonrelevant propositions across texts. More specifically, there were a total of 32 (38.1% of the total number of propositions) relevant propositions and 52 (61.9%) nonrelevant propositions for the Hashimoto's Hypothyroidism text and 48 relevant (37.2%) and 81 (62.7%) nonrelevant propositions for the Pericardial Effusion text.

Pathophysiology Explanations of Clinical Cases

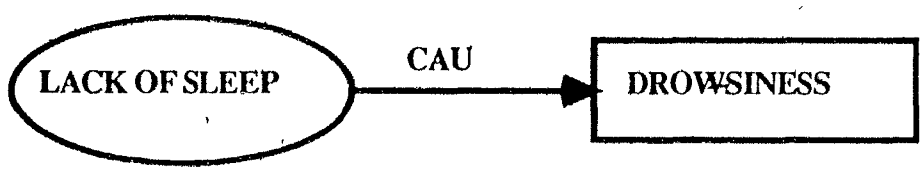
It has been proposed (Feltovich and Barrows, 1984; Patil and Szolovits, 1981) that subjects' causal knowledge can be assessed by requesting them to explain the underlying pathophysiology of a clinical case. In more basic terms, this requires subjects to explain the occurence of an unspecified number of symptoms which are explicitly stated within the text by using their basic science knowledge. At a more theoretical level, this requires subjects to build a problem model or frame (Minsky,

1975) which underlies the clinical information provided in the text (Patel and Groen, 1986a). Analyzing these pathophysiological protocols involves, first, representing them in terms of a causal network. A causal network is a relational structure with labelled links and nodes. Although these protocols are not actually propositionally represented, propositional analysis does form the basis for determining these relational structures. That is, a working knowledge of propositional analysis is sufficient to enable one to draw out the causal and conditional links within the pathophysiological protocols, and it is not necessary to represent the semantics of these protocols in their entirety. This representation allows one to see how a problem model is built either using the subjects' world knowledge or specific knowledge of the case.

As an example of this procedure, consider the following statement contained in a subject's pathophysiology protocol:

Lack of sleep causes drowsiness.

In constructing a causal network of this subject's pathophysiology explanation, this would be represented as:



The square box indicates information contained in the original text and the circle indicates information provided by the subject (i.e., non-text based information). The information contained in these two nodes typically corresponds to a proposition. The labelled arrow indicates a causal link between these nodes. The representation of a subject's entire pathophysiology protocol contains a series of these link-node structures. It should be noted, however, that the order of these structures reflects the order in which they appear in the subject's protocol. That is, the order of the structure simply mirrors the order which the subject has imposed and does not imply any hierarchical organization.

The second step in analyzing these protocols involves examining the accuracy of their content. This is accomplished by examining the representation of each protocol in conjunction with a physician. Both the accuracy of the information contained in the nodes and the links which connect these nodes are assessed. These two analyses provide qualitative information regarding subjects' protocols.

Finally, the number of text-based symptoms contained in each subject's protocol, for each of the two clinical texts is calculated. This allows for a more quantitative discussion of these protocols.

Diagnoses

Students' diagnoses were initially categorized as either accurate or inaccurate. Due to the diverse nature of the diagnoses which fell into this latter category, it was subsequently subdivided into four subcategories, including no diagnosis, treatment-related diagnosis, pathophysiology-related diagnosis, and incorrect diagnosis. A treatment-related diagnosis typically involved a discussion of the medical procedures considered to be necessary to cure the patient's illness, at the expense of labelling the disease itself. A pathophysiology-related diagnosis involved a description of the cause of some of the patient's symptoms without reference to the patient's specific clinical disorder. An inaccurate diagnosis, while having the essential property of a diagnosis, was incorrect for the particular text for which it was provided. For each of the two clinical texts, then, the number of diagnoses which fell into each of the five categories was simply tabulated. In addition, the accurate and inaccurate diagnosis categories were collapsed in an attempt to estimate each group's ability to identify the major malfunctioning organ for the two patients.

Statistical Analysis

A repeated measures multivariate analysis of variance was conducted on the comprehension and relevancy data. The dependent measures were the percentage of relevant and nonrelevant propositions which were recalled, inferred, and part-recalled. There was one within subjects factor of text and this factor had two levels. Additionally, there were two between subjects factors. These included group, with three levels, and text order, with two levels. A summary of the variables included in the analysis is presented in Table 2.

Analyses were conducted using the statistical package SPSSX, version 2.1 on an AMDAHL 470 V7 computer.

TABLE 2

Summary of the Variables Included in the Multivariate Statistical Analysis

INDEPENDENT VARIABLES:

Group

GROUP 1: CEGEP Students
GROUP 2: Med-P Students
GROUP 3: Degree Students

Text

TEXT 1: Hashimoto's Hypothyroidism
TEXT 2: Pericardial Effusion

Text Order

ORDER 1: Text 1 followed by Text 2
ORDER 2: Text 2 followed by Text 1

DEPENDENT VARIABLES

Response Type

RECALL
INFERENCE
PART-RECALL

Relevance

RELEVANT to the diagnosis
NONRELEVANT to the diagnosis

CHAPTER IV

COMPREHENSION OF CLINICAL CASES

Results and Discussion

The summary results of the repeated measures multivariate analysis of variance are presented in Table 3. The means and standard deviations are presented in Table 4. Four significant main effects emerged which were all involved in several higher-order interactions. These main effects included group ($F(2,23)=3.815$, $p=.037$), response type ($F(2,22)=72.458$, $p<.0005$), text ($F(1,23)=23.810$, $p<.0005$), and relevancy ($F(1,23)=142.980$, $p<.0005$). Significant two-way interactions included group by response type ($F(4,44)=2.741$, $p=.040$), text by response type ($F(2,22)=17.053$, $p<.0005$), text order by response type ($F(2,22)=3.897$, $p=.036$), and text by relevancy ($F(2,22)=16.173$, $p=.001$). Finally, a significant three-way interaction was found for text by response type by relevancy ($F(2,22)=4.552$, $p=.022$).

The group by response type interaction is depicted in Figure 1. All groups of students consistently recalled more propositions than they inferred, and their level of inferencing was higher than their level of part-recalls. There were differences between groups, however, in the mean percent of propositions which formed the basis for recalls, inferences, and part-recalls. More specifically, after examining Figure 1, it is immediately apparent that the interaction was due to the performance of the Med-P students. The CEGEP students and the Degree students recalled and inferred comparable percentages of text propositions. The Med-P students, however, recalled

TABLE 3

Results of the Repeated Measures Multivariate Analysis of Variance

Source	F	df	Significance
Constant	32579.02	1, 23	.000
Between Subjects			
Group (G)	3.82	2, 23	.037 *
Text Order (TO)	< 1	1, 23	.557
G X TO	2.72	2, 23	.087
Within Subjects			
Text (T)	23.81	1, 23	.000 *
Relevance (R)	142.98	1, 23	.000 *
Response Type (RS)	72.46	2, 22	.000 *
T X R	16.17	2, 22	.001 *
T X RS	17.05	2, 22	.000 *
R X RS	3.21	2, 22	.060
T X R X RS	4.55	2, 22	.022 *
Interactions of Between and Within Factors			
G X T	< 1	2, 23	.699
G X R	3.01	2, 23	069
G X RS	2.74	4, 44	.040 *
TO X T	< 1	1, 23	.476
TO X R	1.18	1, 23	.289
TO X RS	3.90	2, 22	.036 *
G X TO X T	< 1	2, 23	.602
G X TO X R	1.10	2, 23	.349
G X TO X RS	2.27	4, 44	.076
G X T X R	< 1	2, 23	415
G X T X RS	< 1	4, 44	.441
G X RS X R	< 1	4, 44	.807
TO X T X R	2.13	1, 23	.158
TO X T X RS	2.13	2, 22	.143
TO X R X RS	< 1	2, 22	.966
G X TO X T X R	< 1	2, 23	.526
G X TO X T X RS	< 1	4, 44	.850
G X TO X R X RS	< 1	4, 44	.492
G X T X R X RS	< 1	4, 44	.910
TO X T X R XRS	< 1	2, 22	.587
G X TO X T X R X RS	< 1	4, 44	.607

TABLE 4

Means and Standard Deviations for Recall, Inference, and Part-Recall of the Three Groups by Clinical Case, Relevance of Information, and Text Order.

		Hashimoto's Hypothyroidism				Pericardial Effusion			
		Order 1		Order 2		Order 1		Order 2	
		Mean	S.D.	Mean	S D	Mean	S D	Mean	S.D.
Group 1 (CEGEP)									
High Relevance	Recall	20 000	11 180	16 875	4 739	17 083	6 972	7 083	3 486
	Inference	15 000	8 089	8 125	7 194	13 333	4 796	7 917	1 743
	Part-Recall	6 875	3 423	8 750	5 135	417	932	2 500	2 716
Low Relevance	Recall	18 077	11.586	7 308	2 507	9 877	5 856	8 148	4 150
	Inference	9 615	5.267	5.385	4 169	6 667	3 962	5 926	3 074
	Part-Recall	2.308	1 609	3 077	2 917	.494	1.104	494	676
Group 2 (Med-P)									
High Relevance	Recall	21.094	16.413	35 625	5 229	13 021	5 201	19 167	11 730
	Inference	17.969	11 250	18 750	3 827	16 146	7 488	18 750	5.312
	Part-Recall	10.156	2 992	11 797	5 229	1 042	2 083	1 250	1 141
Low Relevance	Recall	11.539	9.931	22.308	19 115	9.568	7.897	13 580	8.189
	Inference	12 019	3.284	8 077	5.160	11 420	6.325	10 617	4 750
	Part-Recall	2.404	962	3 462	2 507	.309	617	494	.676
Group 3 (Degree)									
High Relevance	Recall	15.625	7 967	23 750	4 739	10 000	4.517	12.917	7 424
	Inference	10.000	4 075	12.500	5 846	12 500	3 898	14 167	2.716
	Part-Recall	9 375	7.329	12.500	11.049	1 667	932	833	1 141
Low Relevance	Recall	5.385	6 853	12.692	3 493	6.420	5.617	8 642	6.648
	Inference	8 077	3.699	6.923	4.427	7.901	5.060	7 407	3 904
	Part-Recall	3 077	2.193	4.615	3.988	.247	.552	.494	1 104

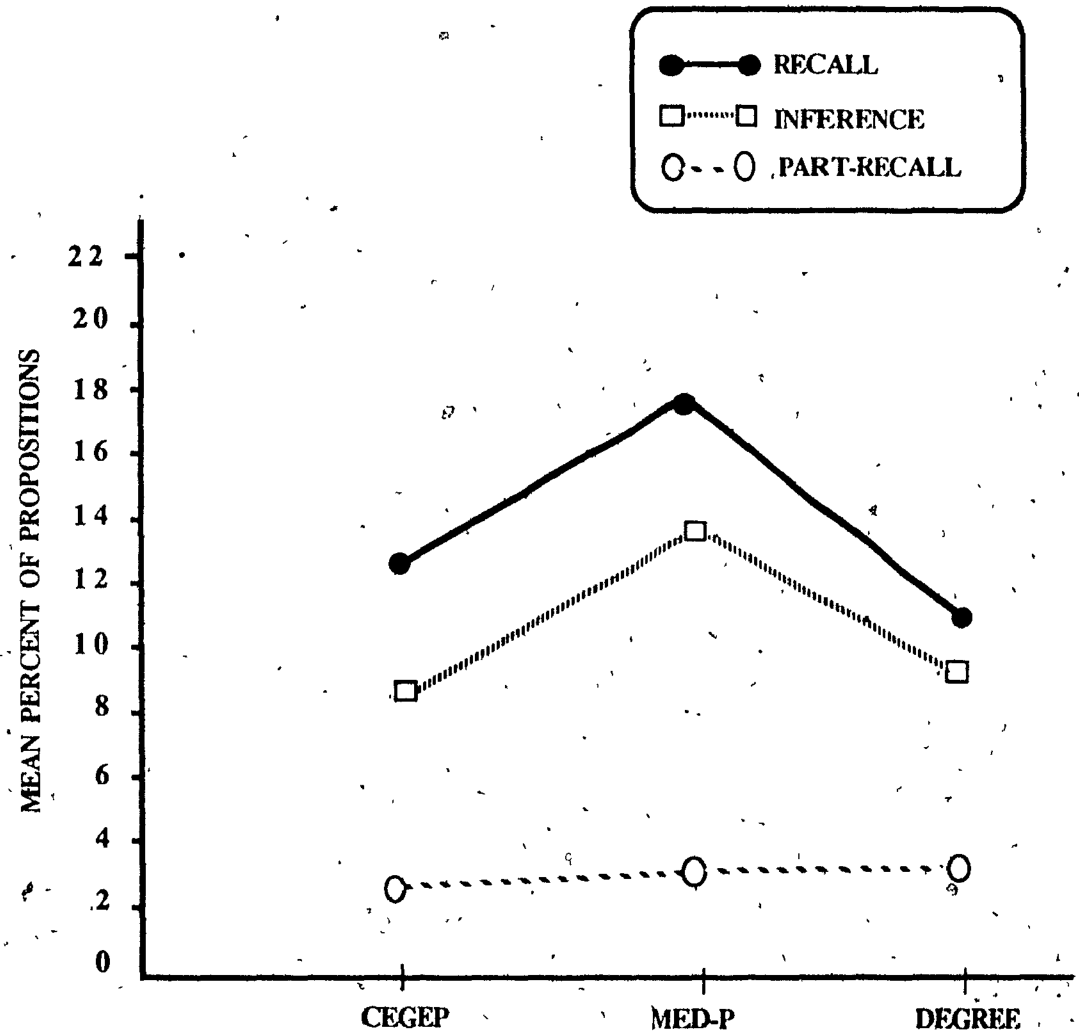


Figure 1. Mean Percent of Propositions by Group and Response Type
F(4,44)=2.74, p=.040

and inferred many more propositions than either the CEGEP students or the Degree students. In fact, the Med-P students inferred more propositions than either of the other two groups recalled. It should be noted, however, that the degree of inferencing for all groups is relatively low. The level of part-recalls for all groups was very low. There was a progressive increase in the number of propositions which formed the basis for this response type from the CEGEP group through to the Degree group, but this increment was minimal.

This pattern of results suggests that not all groups of students interpreted the texts in the same way. Since it has been proposed (Collins, Brown, and Larkin, 1980) that inferences involve higher level cognitive process than recalls, the Med-P students may have processed the clinical texts at a more abstract level than the other two groups of students. It is interesting that the Degree students, with three or four years of University science education performed at a comparable level as the CEGEP students who have no science education at the University level.

The text by response type interaction is illustrated in Figure 2, where it is apparent that the interaction is due to the more verbatim response types of recall and part-recall as opposed to the more abstract, inferential response. As a group, students recalled a much higher percentage of propositions from the Hashimoto's Hypothyroidism text than from the Pericardial Effusion text. They also partially recalled more propositions from Text 1 than from Text 2. These differing response patterns across the two texts ~~suggest that~~ the students processed them differently. Interestingly, the mean percent of propositions which formed the basis for inferences was virtually identical across texts. Overall, then, there were more propositions included in subjects' summaries for Text 1 than for Text 2, but the consistency in level of inferences across texts indicates that both texts appear to be processed at a comparable level of abstraction.

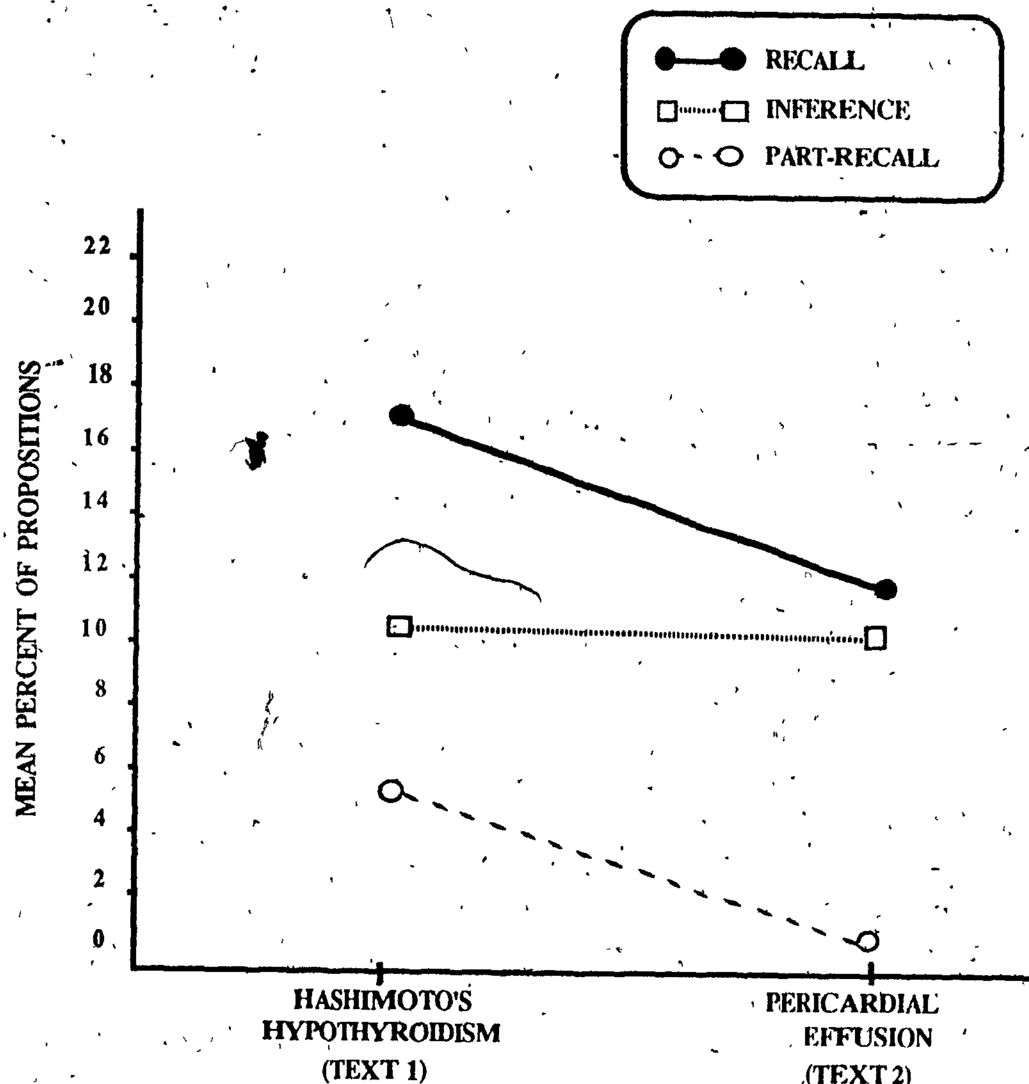


Figure 2. Mean Percent of Propositions by Text and Response Type
$F(2,22)=17.05$, $p<.0005$

Figure 3 depicts the text order by response type interaction. The consistency in level of inferencing across texts noted just previously is no longer entirely true. That is, depending on the text order students receive, there is a slight variation in the mean percent of propositions which form the basis for inferences. More specifically, students given the Hashimoto's Hypothyroidism text (Text 1) first tended to infer more overall than if they were given this text second. Conversely, students given the Pericardial Effusion text (Text 2) first, recalled more overall than students given this text second. This is true as well for part-recalls.

Potential factors contributing to this observed behavior are the levels of difficulty and comprehensibility of the clinical texts used. Given the level of the students being examined and the nature of the task, it seems reasonable to assume that if one text is easier and /or more comprehensible than the other, then this will be reflected in their performance. On this premise, an expert physician was requested to compare the two texts in terms of level of difficulty and comprehensibility. The results of this comparison indicated that the Pericardial Effusion text was both more difficult and less comprehensible than the Hashimoto's Hypothyroidism text. This decision was based on the quantity and type of laboratory tests and test results, the length of the text, and the complexity of the disease process itself. Thus, it would appear that exposing students, initially, to a difficult and somewhat incomprehensible text, detrimentally affects their level of processing of a subsequent text. In comparison, exposing students to an easier, more comprehensible text facilitates or enhances their level of processing of a subsequent text.

The text by relevancy interaction is shown in Figure 4. Students, as a whole, discriminated between relevant and nonrelevant text propositions as indicated by the differential percentage of these propositions included in their summaries. Students tended to operate more on the relevant as opposed to the nonrelevant information and,

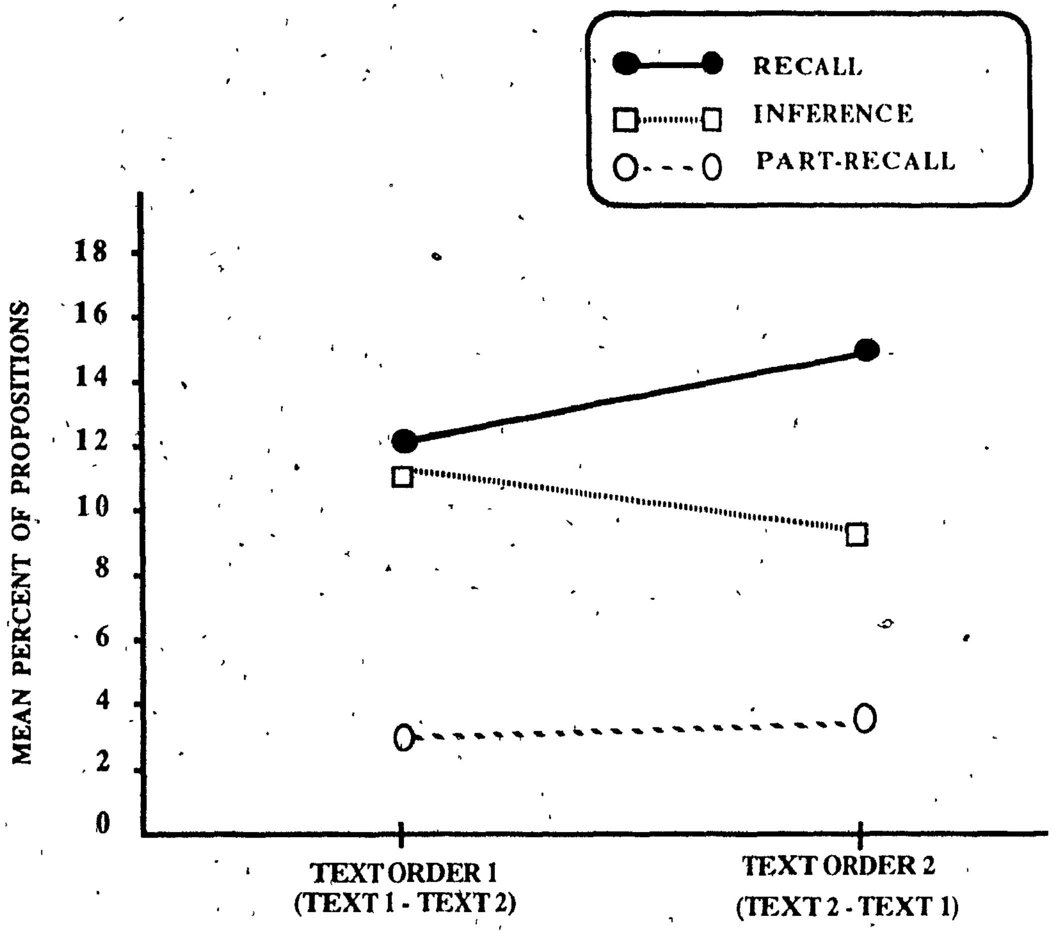
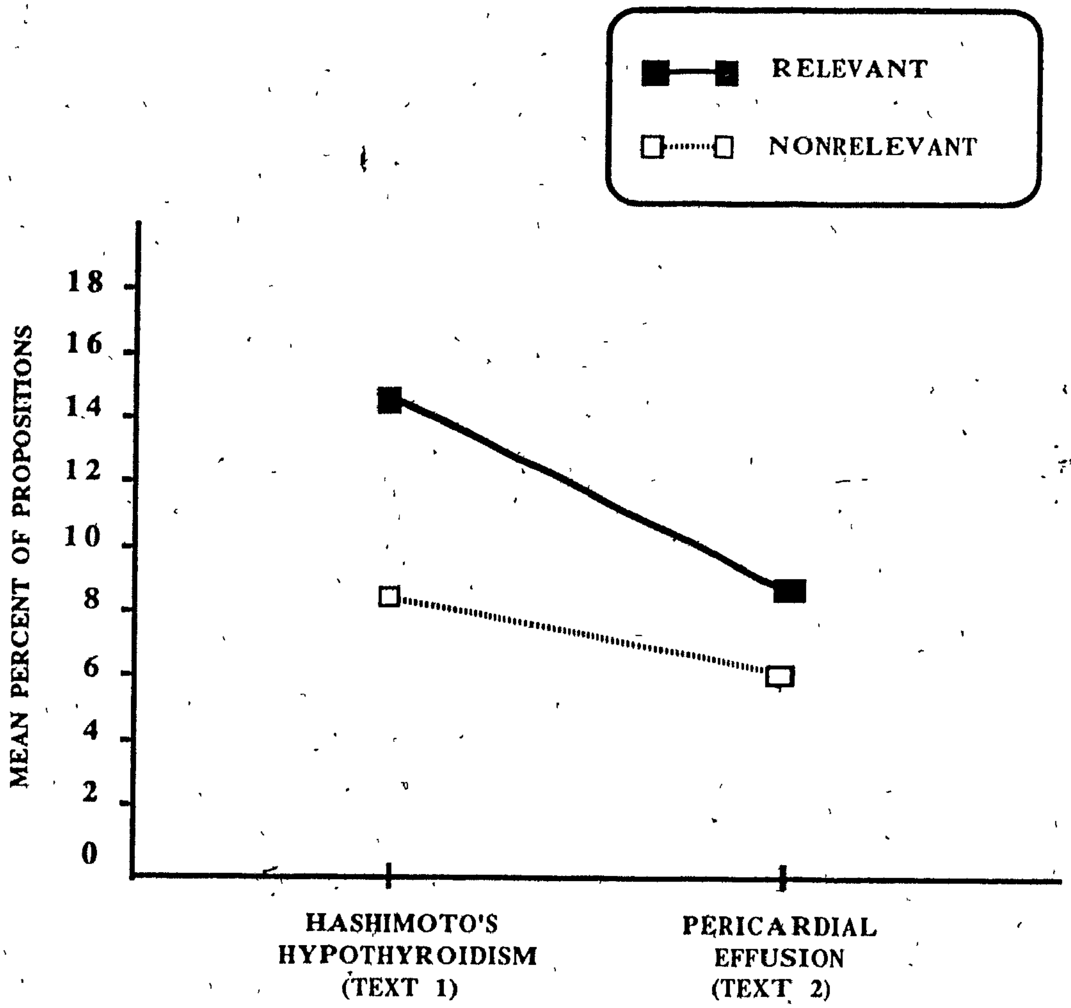


Figure 3. Mean Percent of Propositions by Text Order and Response Type $F(2,22)=3.90$, $p=.036$



<u>Figure 4.</u> Mean Percent of Propositions by Text and Relevance of Information $F(2,22)=16.17$, $p=.001$

while this is true for both texts, this distinction is more prominent for the Hashimoto's Hypothyroidism text than the Pericardial Effusion text.

As with the text order by response type interaction discussed previously, the text by relevancy interaction may be due to the differential characteristics of the two texts. That is, students were better able to discern the relevant from the nonrelevant information in the Hashimoto's Hypothyroidism text than in the Pericardial Effusion text presumably because this former text was less difficult and more comprehensible than the latter text.

The significant three-way interaction of text by response type by relevancy is illustrated in Figure 5. The consistency in the mean percent of propositions which form the basis for inferences across texts and the discrimination between relevant and nonrelevant information is once again present. Overall, relevant text propositions were inferred upon more than nonrelevant propositions, but across texts, the degree of inferencing on relevant propositions was virtually identical, as was the degree of inferencing on nonrelevant propositions.

General Discussion

The effect of group was included in only one of the five significant interactions found in the present study. Based on this finding, two conclusions are drawn. The first conclusion is based on the interaction which includes the factor of group. This interaction implies, as noted previously, that the three groups of students did not represent the clinical cases equivalently. The CEGEP students and the Degree students were more similiar in the way they represented the information than either was to the Med-P group. This pattern of results has been found in another study reported by Patel, HoPingKong, and Mark (1984) which investigated the comprehension processes of subjects with varying levels of medical expertise, and used the same method of

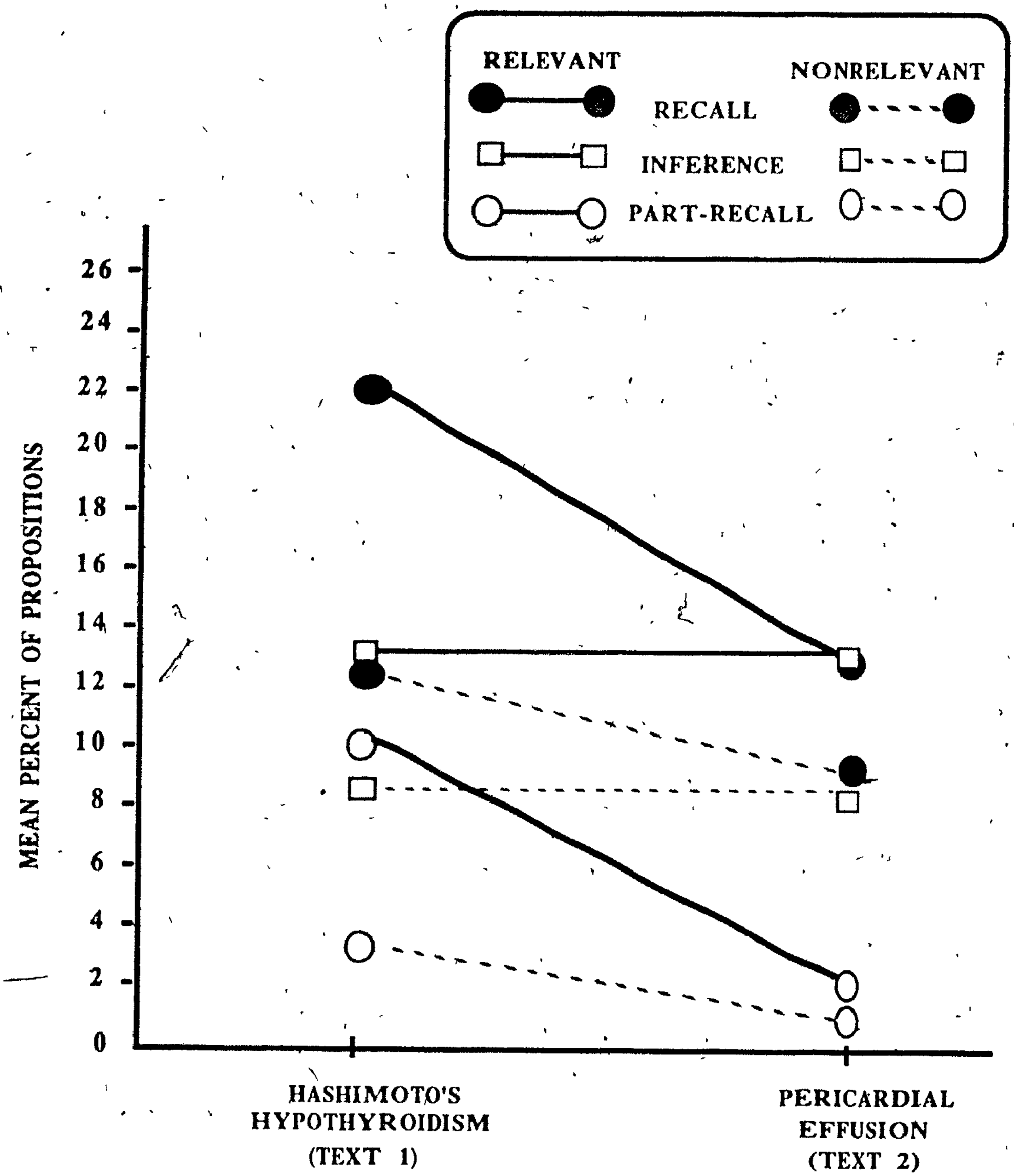


Figure 5. Mean Percent of Propositions by Text, Response Type, and Relevance of Information $F(2,22)=4.55$, $p=.022$

analysis as that used in the present study. The subjects in the Patel, HoPingKong, and Mark (1984) study included first year medical students (novices), second year medical students (intermediates), and physicians (experts). When indices of the comprehension processes (e.g., recall, inference) of these three groups were compared, it was found that the novices and experts resembled each other more so than either resembled the intermediates and, further, that the intermediates tended to operate on the greatest number of aspects of a medical text. When this is depicted graphically, it resembles Figure 1 of the present study, and the result is a peak formation, with the intermediate students being located at the tip of the peak and the novice and expert subjects both at comparable positions at the base of the peak.

The propositions which formed the basis for these global reponse types (e.g., recall, inference) were subsequently categorized as either relevant or nonrelevant to the diagnosis. This manipulation served to clarify the seemingly curious similarity between the high and low level groups, as well as the superior performance of the intermediates over the experts. More specifically, two significant interactions emerged which involved the factor of relevancy. The first interaction was a two-way interaction involving experience level and relevance in which experts were found to operate the most on high relevance propositions and the least on low relevance propositions. The intermediates operated slightly less than the experts on high relevance propositions and the most on low relevance propositions. Finally, novices operated the least on high relevance propositions and less than the intermediates but more than the experts on low relevance propositions. It was concluded that the three groups of subjects demonstrated differential abilities in selecting relevant from nonrelevant information.

The second interaction noted in the Patel, HoPingKong, and Mark (1984) study involved experience level, relevancy, and response type. This finding indicated that, in addition to possessing differential abilities in selecting relevant from nonrelevant information, subjects processed these two types of information differently.

While in the present study, there was no significant group by relevancy interaction, or group by relevancy by response type interaction, the former interaction did approach significance ($p=.069$). The Med-P group (which could be classified as an intermediate group if level is determined by the number of years of premedical science education) operated the most on both relevant ($\overline{X}=15.4$) and nonrelevant ($\overline{X}=8.8$) propostions (Fig. 6). The novice or CEGEP group operated the least on relevant propositions ($\overline{X}=10.3$) and less than the intermediates but more than the experts on the nonrelevant propositions ($\overline{X}=6.5$). Finally, the Degree, or expert subjects, operated the least on the nonrelevant propositions ($\overline{X}=6.0$) and while not the most on relevant propositions, at least more than the novices ($\overline{X}=11.3$). Thus, the processes of selectivity attributed to the expert subjects in the Patel, HoPingKong and Mark (1984) study seem to be operative, at least at a minimal level, in the Degree group. These students, however, do not yet process this information in a differential manner, as indicated by the nonsignificant triple interaction of group, relevancy, and response type.

The second conclusion to be drawn from the comprehension data is based on the remaining four significant interactions which did not include the group factor. The absence of this factor in these interactions indicates that any conclusions based on these findings may extend to all three groups of students. All four of these interactions involve, in one form or another, the factor of text and, therefore, each interaction may be related to the differential level of difficulty and comprehensibility of these texts. More specifically, it was found that students' pattern of response (both global levels of recalls, inferences, and part-recalls, and response type on relevant and nonrelevant propositions) was different for each of the two texts and was, additionally, affected by the order in which the texts were presented. The interpretations of these results are

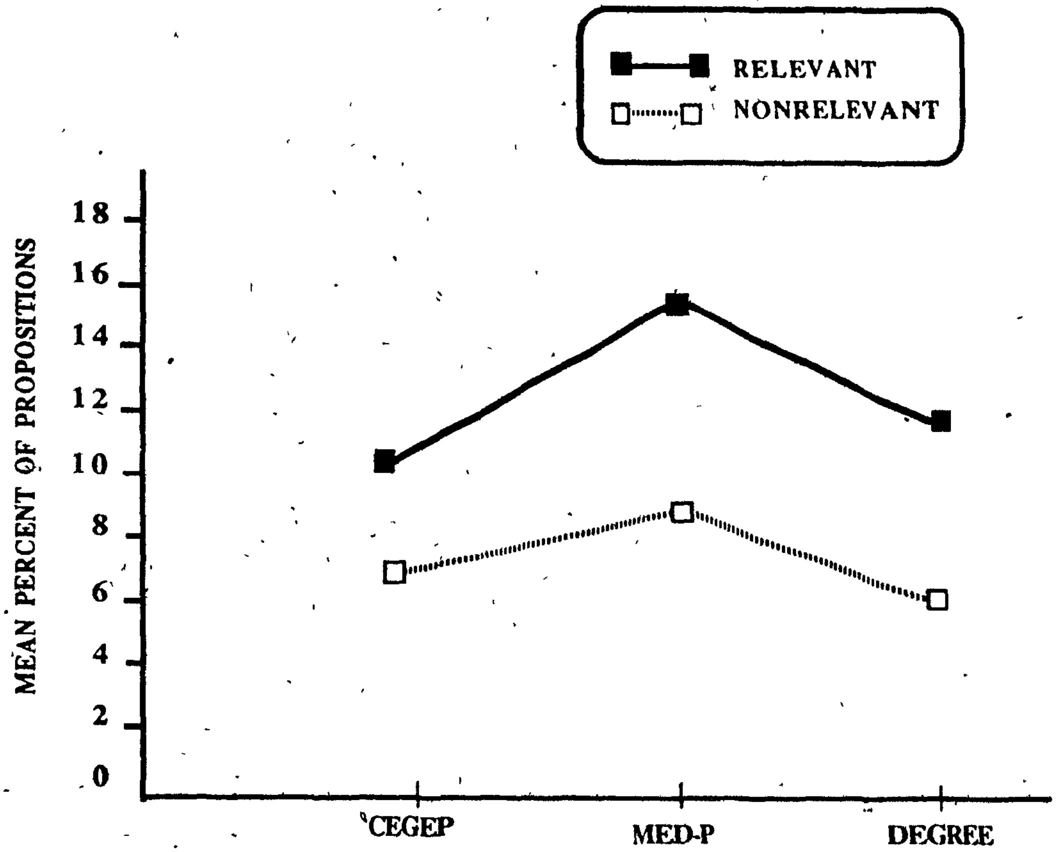


Figure 6: Mean Percent of Propositions by Group and Relevance of Information $F(2, 23)=3.01$, $p=.069$

presently ambiguous. On one hand, the strong text effects could be due to the characteristics of the sample. If so, then the findings indicate that not only do the differing knowledge structures which the three groups of students bring to the task affect their interpretation of the texts (as evidenced by the main effect of group and the group by response type interaction) but, the texts themselves exert an influence on the manner in which the students interpret them. In other words, the inexperience of the students at this level may leave them vulnerable to the quality of the materials they are exposed to. This, in turn, may have implications for instruction including, for example, the ordering of materials presentation. The response type by text order interaction found in the present study suggests that if the goal of a task is to have students abstract information (from two texts of differing levels of comprehensibility and difficulty) at a high level, as opposed to recalling the information verbatim, then students should initially be presented with the easier, more comprehensible text.

On the other hand, the text effects could be due to the characteristics of the texts themselves. That is, the quality of the texts may affect students' comprehension regardless of premedical background. The untangling of these possibilities requires further research, employing comparable populations and additional text materials.

CHAPTER VI

PATHOPHYSIOLOGY EXPLANATIONS OF CLINICAL CASES

Results and Discussion

The following is a description of the findings from the analysis of subjects' pathophysiology protocols. It was assumed that the vast majority of students' explanations would be inaccurate and consultation with a physician verified this assumption. As a consequence, the following discussion will emphasize the differential nature of students' explanations, at the expense of the accuracy of these explanations.

The protocols of each of the three groups of students will be discussed individually with cross comparisons where appropriate. Since the quality of students' explanations were similar across texts, the results are not segregated by text.

Group 1: CEGEP Students

Six out of a total of twenty protocols were not included in this analysis. Two protocols were excluded because these students did not attempt to provide pathophysiology explanation. An additional four protocols were excluded because the students simply regurgitated some of the patient's symptoms in a list format (i.e., they did not attempt to explain the cause of these symptoms). Thus, these protocols could not yield any insight into the type of prior knowledge these students had nor the manner in which they used this knowledge. These four protocols were provided by two students: one for each of the two cases.

Three findings were readily discernable from the data. First, students at this level had an obvious lack of basic science knowledge. Their explanations tended to be based on common knowledge; a finding which was somewhat expected based on theories concerned with the development of preinstructional science misconceptions. For example, Osborne (1984) has proposed that through interacting with the world, individuals develop 'mini-theories' which help to predict, explain, and describe events. He distinguishes between three types or clusters of mini-theories: 'gut dynamics', 'lay dynamics', and 'physicists dynamics'. It is the middle category of mini-theories which appears to have been operative within this population of CEGEP students. Lay dynamics are defined as consisting of the language and ideas that people develop indirectly, that is, through their exposure to other people, the media, and books. Further, these 'scientific' ideas may consist of facts, fantasies, and beliefs. There are numerous examples of such ideas contained in the pathophysiological explanations provided by the CEGEP students, particularly for the Hashimoto's Hypothyroidism case (Text 1). For example, the students who attempted to account for the patient's constipation noted that it was caused by a lack of fibre. Additionally, students who attempted to account for the patient's rough, scaly skin noted it to be either an allergic reaction to the medication the patient had received for her laryngitis or to a lack of moisture. As a final example, students attributed the patient's high blood pressure to either her obese condition or to a high level of sodium (despite the fact that the patient's sodium level was actually lower than normal; a fact which was included in the text but in numeric form). The presence of common knowledge explanations in first year medical students has previously been documented (e.g., Patel, HoPingKong, and Mark, 1984) but their premedical backgrounds were unspecified.

The second finding which emerged from these protocols was an inability on the part of the students to discern any relationship between the symptoms presented in any one case. CEGEP students typically explained each symptom independently, without reference to the other symptoms or to any one underlying problem. This is exemplified in Figures 7 and 8 where two CEGEP students' pathophysiology explanations are represented in the form of causal networks. In Figure 7, the student (CG #3) explains only three of the patient's symptoms from the endocrinology text: drowsiness, constipation, and skin condition. There is no overlap between these symptoms in terms of their cause; each symptom is explained without relating it to another symptom or to a common cause. The same is true for the pathophysiology protocol (Figure 8) of another student (CG #6) for the cardiology text. In this example, more text-based symptoms are explained but the explanations remain independent of each other. These representations suggest that students view each symptom as an entity all of its own, rather than as part of a network of symptoms or as part of a system.

The third finding derived from the analysis of the CEGEP students' pathophysiology protocols was the little emphasis which was placed on the patient's major problem. Only three of the seven students in this group who offered pathophysiological explanations for the Hashimoto's Hypothyroidism case referred to the patient's thyroid problem. All three of these explanations revolved around the iodine level in the patient. Specifically, these students inaccurately stated that the iodine level caused the thyroid gland to enlarge. In actuality, however, the thyroid gland was enlarged prior to the patient's ingestion of potassium iodide. This medication increased the level of iodide in the patient's body and served to further impair the already malfunctioning thyroid gland. Thus, it would appear that these students have some knowledge of the relationship between iodide and the thyroid gland, but they have difficulties with the direction of the involvement. Additionally, only two students

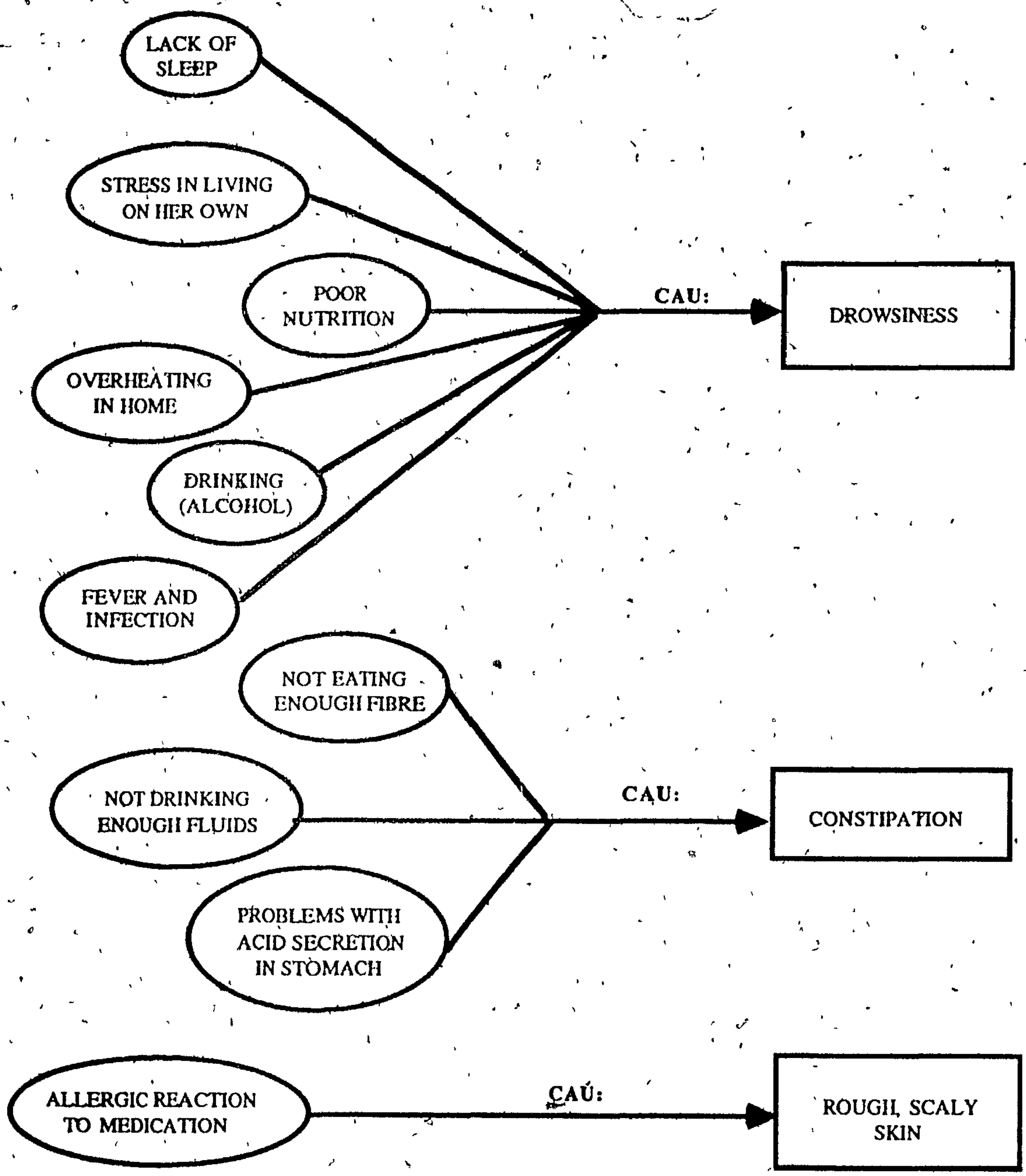


Figure 7. Structural Representation of a Pathophysiological Protocol of a CEGEP Student (CG #3) for the Hashimoto's Hypothyroidism Text

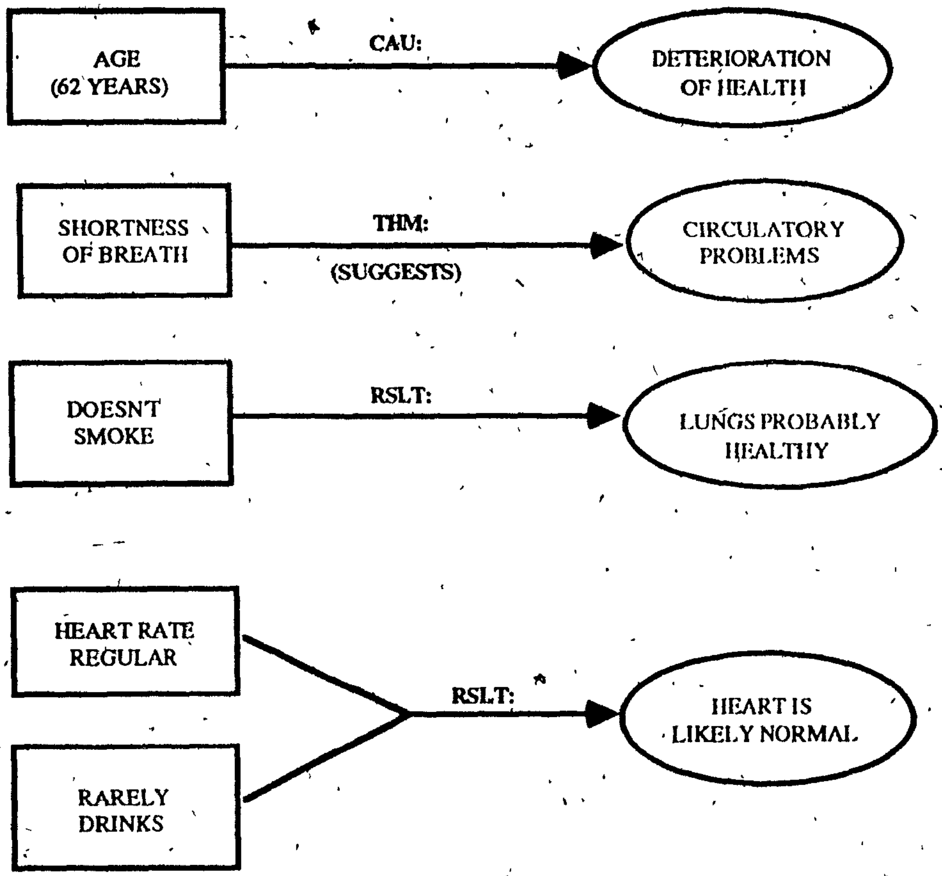


Figure 8. Structural Representation of a Pathophysiological Protocol of a CEGEP Student (CG #6) for the Pericardial Effusion Text

ascribed the enlarged thyroid as a cause of other symptoms which were manifested by this patient, specifically the difficulty the patient had with speaking and the increased blood pressure.

There was a similar lack of emphasis on the patient's major problem in the Pericardial Effusion text. Again, only three of the seven CEGEP students who provided a pathophysiology explanation for Text 2 referred to the patient's heart condition. Unlike the explanations provided for Text 1, however, none of these students delineated the cause of the patient's major malfunctioning organ. Instead, they either attributed it (i.e., the heart) as the cause of another symptom which the patient manifested (e.g., the patient's shortness of breath) or concluded that the heart was functioning normally. This latter conclusion was deduced, in part, on the basis of a select few other findings mentioned in the text (e.g., the fact that the patient rarely drank and that the heart rate was regular).

In summary, then, the pathophysiological protocols provided by the CEGEP students in Group 1 demonstrated (1) a lack of basic science knowledge, (2) a reliance on common knowledge as the basis for their explanations, (3) an inability to recognize any relationship between symptoms, and (4) a lack of emphasis on the patients' major malfunctioning organs and the resultant symptoms arising from these disturbances. Interestingly, the majority of CEGEP students did attempt to explain the pathophysiology of the two cases; only two students did not.

Group 2: Med-P Students

All nine students from this group provided a pathophysiological explanation of both Text 1 and Text 2. In comparison to the protocols obtained from the CEGEP students in Group 1, two striking features emerge. These include the large quantity of information contained within the protocols and the more highly integrated manner in which this information is presented. At a quantitative level, the students in Group 2 attempted to explain an average of 5.4 symptoms from Text 1 and 6.9 symptoms from Text 2. This is in contrast to the students in Group 1 who explained an average of 3.4 symptoms from Text 1 and 2.3 symptoms from Text 2. In terms of integrating this information, the Med-P students, like the CEGEP students, tended to interpret and explain signs and symptoms in a fragmentary manner. This is illustrated, for example, in Figure 9. This Med-P student's (MP#4) explanation of the Hashimoto's Hypothyroidism case revolves around four distinct text-based symptoms: edema, drowsiness, the enlarged thyroid, and the biochemical composition of the blood. The major distinction between this protocol and a protocol generated by a CEGEP student (e.g., see Fig. 8), however, is the larger number of node-link structures associated with each symptom Thus, the Med-P students explained individual signs and symptoms in a more integrated fashion then the CEGEP students but, their protocols continued to lack an overall structure.

Additionally, the explanations offered by the Med-P students tended to be more accurate and sophisticated than those of the CEGEP students. This can be attributed, in part, to the Med-P students' greater reliance on basic science knowledge as opposed to common knowledge for the basis of their explanations. However, it should be noted that many of the explanations offered by the Med-P students were only accurate at a very general level. For example, one Med-P student (MP#4) noted that the patient's drowsy state indicated a lack of oxygen to the body's cells (Fig. 9). Further, this lack of oxygen was interpreted as an indicator of either an abnormally functioning pulmonary system or circulatory system. While both of these interpretations are accurate, this explanation obviously lacks the specifics of these abnormally functioning systems. These explanations are, however, more sophisticated than those offered by

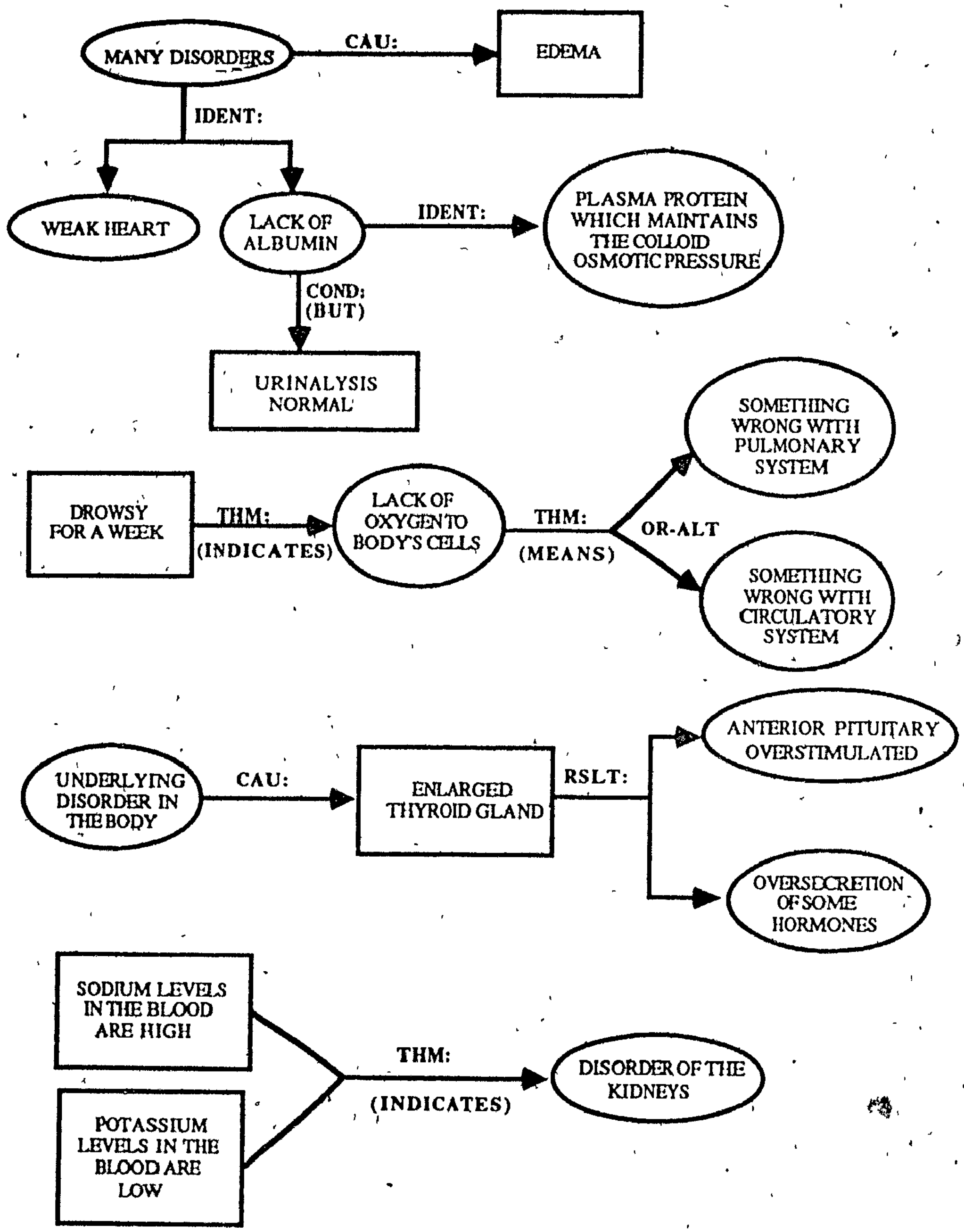


Figure 9. Structural Representation of a Pathophysiological Protocol of a Med-P Student (MP #4) for the Hashimoto's Hypothyroidism Text

the CEGEP students (e.g., see Fig. 7, CG#3). Figure 9 presents another example of the accurate but general quality of the explanations offered by the Med-P students. This student (MP#4) accurately interpreted the laboratory readings of the levels of sodium and potassium (i.e., the sodium level was high and the potassium level was low), and demonstrated some understanding of the relationship between these elements and the functioning of the kidneys. However, it is a very general explanation in the clinical context of this particular case.

Both the CEGEP students and the Med-P students included descriptions in addition to explanations in their pathophysiology protocols. However, these descriptions were of a very different nature. The descriptions found in the CEGEP protocols tended to include definitions of concepts. For example, students noted that because the patients in Text 1 and Text 2 experienced swelling (edema), there was water retention. The causal relationship established by these students, however, merely consists of a definition of edema.

Descriptions contained in the Med-P protocols typically revolved around normal physiological functioning. For example, one student (MP#4) noted that albumin is a plasma protein which maintains the colloid osmotic pressure (Fig. 9). While this is accurate, it does not aid in the generation of an accurate diagnosis. This student further noted that edema could develop if the albumin level is low. However, since the urinalysis was found to be normal, it was deduced that the albumin level must also be normal and, therefore, could not be responsible for the edema. Thus, a number of unnecessary intermediary links were used to explain and then rule out the cause of a symptom.

Figure 10 also illustrates the presence of descriptions in the pathophysiological explanations offered by the Med-P students. In this example, the student (MP#6) describes the normal cycle of bilirubin, including how it is metabolized and excreted.

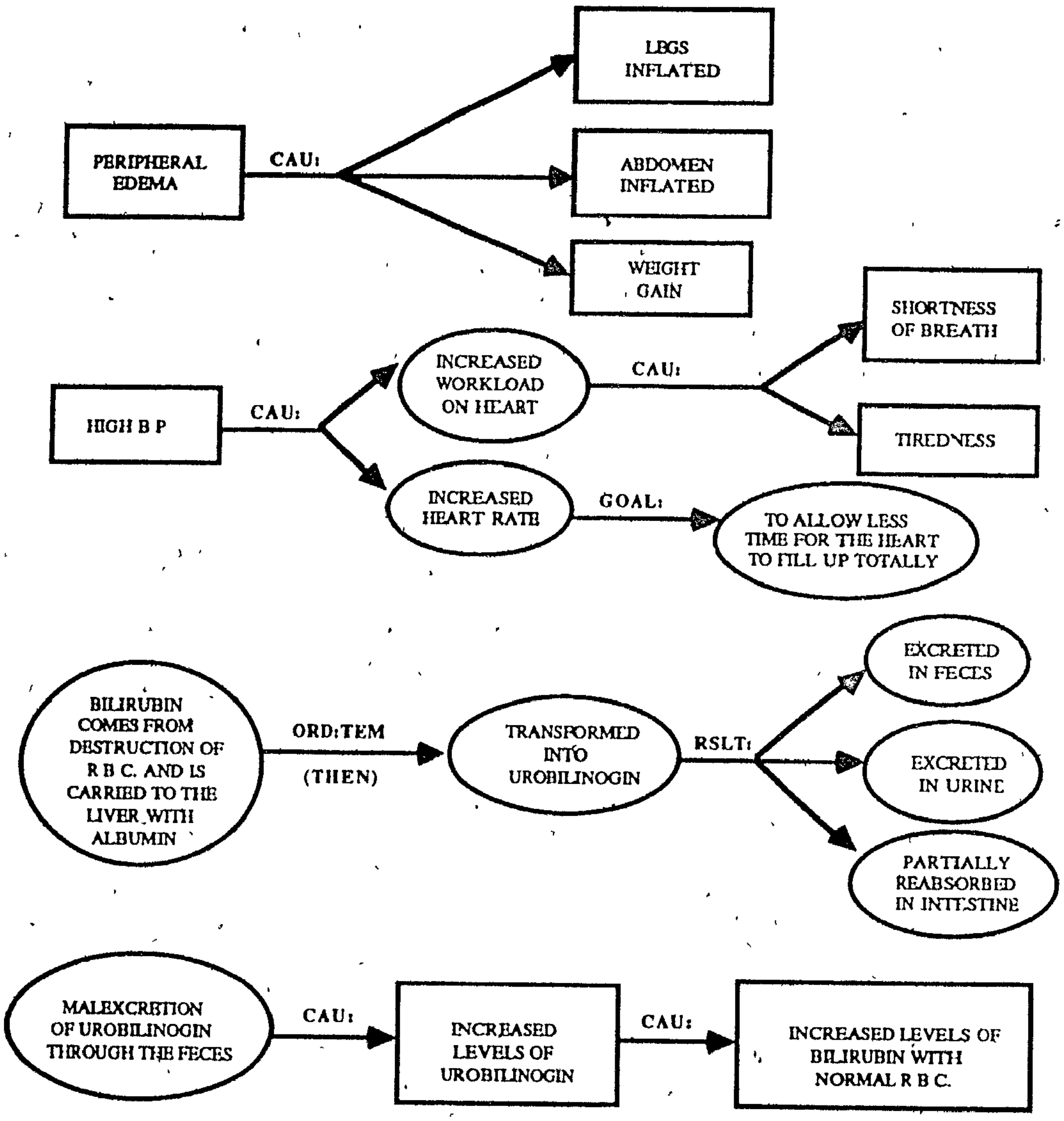


Figure 10. Structural Representation of a Pathophysiological Protocol of a Med-P Student (MP #6) for the Pericardial Effusion Text

While providing an interesting physiological explanation of this phenomenon, it is peripheral to the task at hand.

Differences between the CEGEP students and the Med-P students, in addition to the qualitative nature of their explanations mentioned previously, lie in the symptoms on which they focussed their explanations. The CEGEP students tended to focus on signs and/or symptoms which most people have experienced or for which one has some general knowledge (e.g., blood pressure, drowsiness, constipation, weight gain, edema, shortness of breath). The Med-P students focussed on many of the same signs and symptoms which the CEGEP students did but, additionally, they attempted to interpret many findings which were more technical in nature (e.g., levels of bilirubin, urobilinogen, albumin, arterial blood gases, and hemoglobin, and ECG readings). These latter interpretations may indicate a more detailed understanding of the physiological functioning of the human body.

Another feature which was characteristic of the Med-P explanations was the little emphasis placed on the evaluation of alternative causal explanations of a symptom. This is illustrated, for example, in Figure 9. This student (MP#4) noted that the patient's drowsiness indicated a lack of oxygen which in turn suggested either a malfunctioning pulmonary system or circulatory system. Neither of these alternatives, however, were further evaluated and consequently, neither was ruled out. Similar examples are contained in two other pathohysiology protocols for the Hashimoto's Hypothyroidism text and in three pathophysiology protocols for the Pericardial Effusion text.

In terms of the patient's major malfunctioning organ, over one-half (n=5) of the students from the Med-P group made reference to the thyroid problem of Text 1. While four of these students attempted to delineate the cause of the thyroid enlargement, only one student was accurate in stating that it was due to some auto-immune malfunction.

This student, however, suggested an alternative cause, specifically a low level of iodine in the patient's blood. This is, in fact, the inverse of the actual situation, and resembles the explanations offered by the CEGEP students. The explanations offered by the other three Med-P students were quite varied in nature (e.g., the enlarged thyroid was due to partially occluded airways, overstimulation of the anterior pituitary and oversecretion of various hormones) and all were inaccurate. Three of these students additionally ascribed the enlarged thyroid gland as the source of some of the patient's other problems. An even larger percentage of students from this group made reference to the patient's major problem in their explanation of the second text (56% and 66% for Text 1 and Text 2, respectively). Only two of the six students who made reference to the patient's heart problem attempted to delineate its cause. The majority of students (n=5) ascribed the malfunctioning heart as the source of other symptoms which the patient had manifested.

In summary, the pathophysiology protocols written by the Med-P students as compared to the CEGEP students (1) contained more explanations of text-based symptoms, (2) were more highly integrated (at least within explanations of individual signs and/or symptoms), (3) included functional descriptions as opposed to definitions, and (4) had more emphasis on the patients' major malfunctioning organs and their resultant symptoms. A feature which was unique to the Med-P protocols was the presence of alternative causal explanations which were not evaluated.

Group 3: Degree Students

A total of six protocols were excluded from this analysis; two protocols were excluded because these students did not attempt this section and four protocols were excluded because these students only recalled some of the symptoms from the text in a list format. These six protocols were provided by three students and were equally divided between the two texts.

The protocols generated by the Degree students were, in one respect, similar to those generated by the CEGEP students and, in another respect, similar to those generated by the Med-P students. More specifically, the Degree students and the CEGEP students were similar in terms of the small quantity of symptoms which were accounted for in the pathophysiology explanations of the two clinical cases. The Degree students attempted to explain an average of 3.0 symptoms from Text 1 (as compared to 3.4 for the CEGEP students) and 3.6 symptoms from Text 2 (as compared to 2.3 for the CEGEP students). The qualitative nature of the Degree students' pathophysiology explanations, however, more closely resembled those of the Med-P students than the CEGEP students, although they were slightly more sophisticated in the degree of structuredness which they contained. The majority of the protocols generated by the Degree students contained one or two top-level 'themes' from which all further symptoms or explanations were derived. This is illustrated, for example, in Figure 11. This student's (DG#8) pathophysiology explanation of the Hashimoto's Hypothyroidism text revolves around the patient's consumption of potassium iodide. The remainder of the protocol is linked either directly or through intermediary steps to this state. Figure 12 similarily demonstrates this structured form of reasoning of a Degree student (DG#4) for the Pericardial Effusion text. This student emphasizes the patient's problem with fluid accumulation, specifying both the cause of this state and the resultant respiratory symptoms arising from this state. These two protocols are very different in nature than those generated by the CEGEP students, where symptoms were explained without reference to other symptoms or one underlying problem. Although the protocols generated by the Med-P students were more sophisticated in this respect, they were typically limited to the explanation of a symptom with reference to

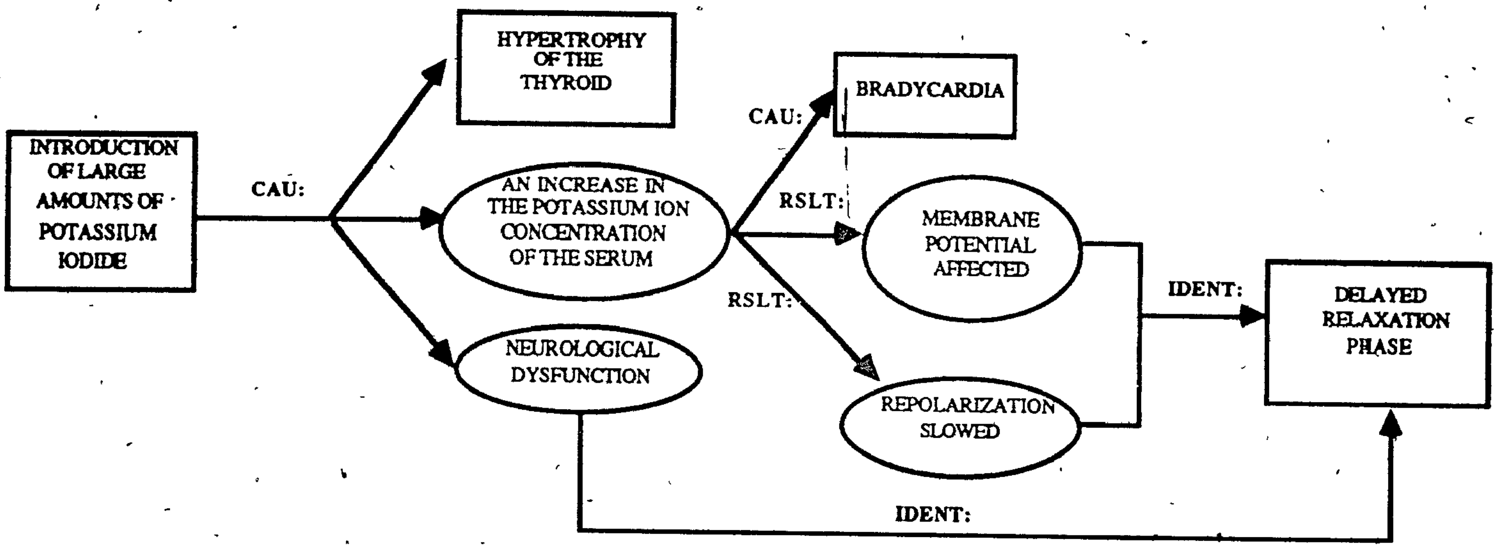


Figure 11. Structural Representation of a Pathophysiological Protocol of a Degree Student (DG #8) for the Hashimoto's Hypothyroidism Text

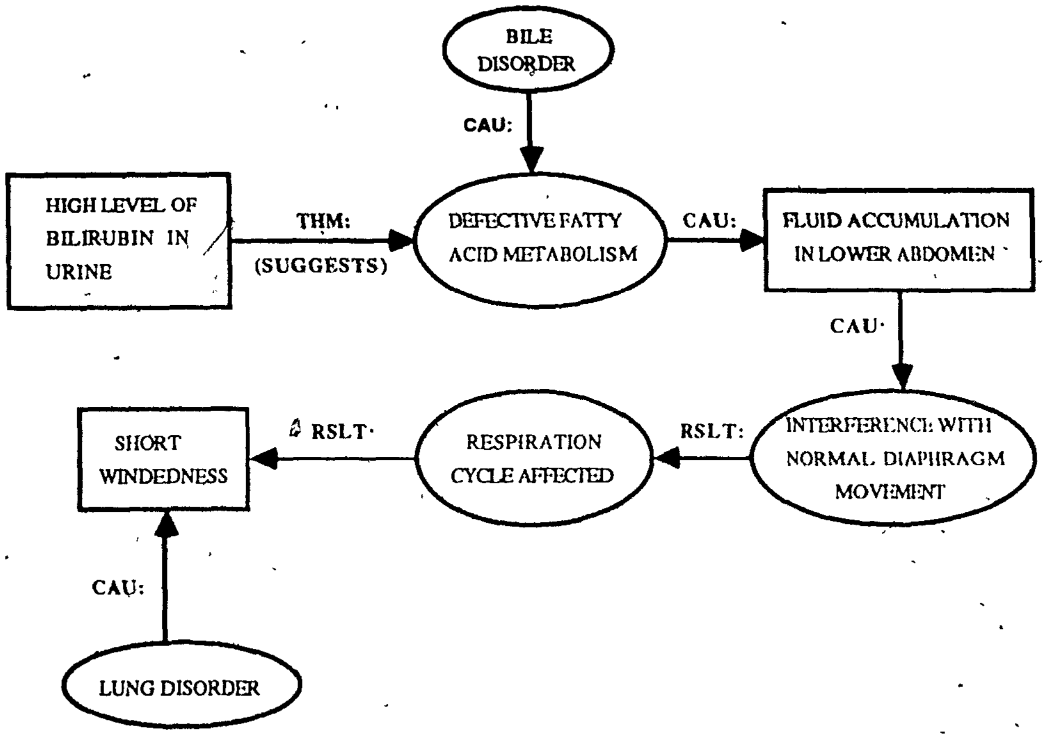


Figure 12. Structural Representation of a Pathophysiological Protocol of a Degree Student (DG #4) for the Pericardial Effusion Text

other symptoms, at the expense of one overall, general problem. The Degree students, in comparison, were more inclined to identify one major problem or state from which all other symptoms and explanations were linked.

Unlike both the CEGEP students and the Med-P students, the Degree students did not include descriptions (functional, definitional, or otherwise) in their pathophysiology explanations of the two clinical cases. In this sense, the Degree students may be more target oriented, refraining from detailing information they may know but that is considered irrelevant to the task at hand (a characteristic of the Med-P students) and avoiding the mistake of using a definition of a concept as a causal explanation of that concept (a characteristic of the CEGEP students).

The absence of descriptions in the pathophysiology explanations offered by the Degree students provides another point of distinction between the three groups of students. As mentioned previously, the explanations offered by the CEGEP students were typically based on common knowledge. Many of the explanations offered by the Med-P students were based on very detailed basic science concepts of normal biochemical functioning. The explanations offered by the Degree students, however, were more global or general in nature. For example, one Degree student (DG#4) explained the patient's shortness of breath (from Text 2) in terms of an interference in the movement of the diaphragm from an accumulation of fluid in the abdomen (Fig. 12). This same symptom was explained by a Med-P student (MP#5) in terms of the patient's low hemoglobin; an insufficient supply of red blood cells transporting oxygen to the body's muscles and tissues (particularly during exercise) resulted in the patient's shortness of breath. A typical CEGEP interpretation of this symptom was that it indicated some form of circulatory problem (e.g., CG#6).

In terms of detecting or emphasizing the patient's major malfunctioning organ, the Degree students were similar to the Med-P students. Slightly more than fifty percent of the Degree students who provided a pathophysiology explanation for the Hashimoto's Hypothyroidism case (Text 1) made reference to the patient's thyroid problem. Three of these four students attempted to delineate the cause of the thyroid enlargement. Additionally, three students attributed the enlarged thyroid as the source of other symptoms which the patient manifested. Four Degree students also emphasized the major problem of the patient with Pericardial Effusion, described in Text 2. All four of these students attributed the heart as the source of some of the patient's symptoms at the expense of stipulating why the heart was malfunctioning.

To recapitulate, the general findings obtained from this analysis of the Degree group's pathophysiology explanations included: (1) a similarity to the CEGEP group in terms of the small quantity of text-based symptoms contained in these protocols, (2) a similarity to the Med-P group in terms of the high degree of structuredness with which the underlying disease process was outlined, (3) the inclusion of one or two top level-states from which all remaining symptoms and explanations were linked, (4) the absence of any form of descriptions, (5) the presence of global or general explanations, and (6) a similarity to the Med-P group in terms of the degree of emphasis placed on the patients' major malfunctioning organs and their resultant symptoms.

General Discussion

The analysis of students' pathophysiological explanations of the two clinical cases revealed some qualitative information about the different types(s) of knowledge each group possesses and the manner in which they use this knowledge. The CEGEP students, for example, had not yet been exposed to any formal training in the basic sciences. Not surprisingly, then, there was an obvious lack of basic science information in their explanations. These students have, however, experienced illness both directly (through their own personal experience) and indirectly (through family members, friends, and the media). This experiential knowledge provided the basis for their explanations and resulted in a limited and superficial understanding of the disease processes implicitly contained in the clinical texts. Their primary means of reasoning through a clinical case was through a process of association; CEGEP students associated the signs and symptoms of the patients described in the clinical texts to those which they have experienced. This was evidenced in all facets of their explanations: the specific symptoms on which they focussed their explanations, the lack of cohesiveness between possible causes of different symptoms, and the explanations themselves.

The Med-P students in this study were, only one year previous, CEGEP students. At the time of testing, the Med-P students had had one year of basic science instruction at the university level. While this can not be considered extensive, it did increase the quantity of basic science explanations contained in their protocols as compared to those contained in the CEGEP students' protocols.

The basic science courses which the Med-P students had successfully completed, however, dealt with the **normal** physiological mechanisms of human functioning rather than the pathophysiological mechanisms. This was evidenced by the presence of what was previously referred to as 'functional descriptions' in their protocols. In addition, many of the signs and symptoms which the Med-P students could not explain in terms of basic science concepts were explained in such a way that they made sense at an intuitive level. For example, one student (MP#4) noted that because the patient described in Text 1 had been drowsy for a week, there was an insufficient supply of oxygen to the body's cells. This, in turn, was interpreted as being indicative of either a pulmonary system problem or a circulatory system problem. While both of these systems are related to the delivery of oxygen, this fact was probably derived from common knowledge, as opposed to any real understanding of the disease process itself.

Thus, the explanations provided by the Med-P students reflected a familiarity with some normal physiological mechanisms (due to the one-year of basic science instruction which had received) and an unfamiliarity with pathophysiological mechanisms (due to the absence of this focus in their basic science courses). As a consequence of this latter condition, many of the pathophysiological explanations offered by these students were based on common knowledge. Further, when alternative explanations of this nature were provided for one symptom, students were unsuccessful in evaluating their relative plausibility because they lacked the necessary pathophysiological knowledge.

The Degree students had completed either a three-year or four-year undergraduate degree in the basic sciences and presumably have a fairly extensive knowledge base concerning the mechanisms of normal physiological functioning. They appeared to have relied on this knowledge as opposed to common knowledge as the basis for their explanations. Whether this common knowledge has actually dissipated and been replaced with more scientific knowledge or exists but is known to be inaccurate, can not be determined. The important point to note, however, is that the Degree students do not use this knowledge as the basis for their explanations. When they encounter a symptom which can not be explained in terms of their basic science knowledge, they did not provide an explanation based on common knowledge but, instead, excluded this symptom from their protocol altogether. This strategy seems plausible considering the small quantity of text-based symptoms which were included in their protocols but which were explained in basic science terms.

In comparison to the other two groups of students, the Degree students appeared to have a more thorough understanding of the interconnectedness between different body systems; their protocols typically revolved around one or two top level 'themes', suggesting an attempt to view the symptoms, initially presented in a discrete manner, as a network of symptoms. However, the small quantitiy of text-based symptoms which were included in their protocols and the erroneous nature of many of their explanations demonstrates their lack of pathophysiological and clinical training.

CHAPTER VI

DIAGNOSES

Results and Discussion

The results presented in this section will emphasize both the types of diagnoses provided by the three groups of students combined, as well as the content of the diagnoses provided by each of the three groups individually. The results will refer to Text 1 followed by Text 2.

Text 1: Hashimoto's Hypothyroidism

The overwhelming majority (90%) of diagnoses for this text were inaccurate. It seems that many students did not even understand the concept of a diagnosis. This is evident from Table 5 where a number of diagnoses can be more accurately characterized as treatment-management plans (n=4) or extensions of pathophysiology explanations (n=6). These two categories account for slightly more than one-third of the diagnoses provided for this text. More specifically, 30% of the CEGEP students, 33% of the Med-P students, and 40% of the Degree students failed to provide what could accurately be termed a diagnosis.

As an example of diagnoses which are actually treatment-management plans, consider the following 'diagnoses' offered by a CEGEP student and a Degree student, respectively:

TABLE 5

Diagnoses Provided by CEGEP, Med-P, and Degree Students for Text 1: Hashimoto's Hypothyroidism

	GROUP		
	CEGEP STUDENTS (n=10)	MED-P STUDENTS (n=9)	DEGREE STUDENTS (n=10)
DIAGNOSES			
No Diagnosis	2	1	2
Treatment Related	1	0	3
Pathophysiology Related	2	3	1
Incorrect	3	4	4
Correct	2	1	0

CEGEP Student

From personal experience, I know that eating too many or only starches and not drinking enough fluids causes drowsiness, fatigue, constipation, and obesity. The only thing I can suggest to the woman is for her to drink fluids, eat more vegetables and less starch, and to get enough sleep. [CEGEP #3]

Degree Student

Iodine may be administered to correct the thyroid gland enlargement. The expectorant appears to be ineffective (there still was a 30 lb. weight gain). Some lung treatment is necessary. [Degree #4]

As an example of diagnoses which are actually extensions of pathophysiology explanations, one such 'diagnosis' from each group will be presented.

CEGEP Student

The woman's problems seem to be due mostly to a difficulty in bringing enough oxygen into her blood, and possibly mucus formation in the lungs. I don't think her heart is the cause of her problems, although it could be somewhat affected. As for the constipation, it doesn't appear to be a related problem. [CEGEP #4]

Med-P Student

The elderly lady suffered from an inadequate supply of oxygen to her organs and muscles, and an inadequate system of elimination, which both led to a loss in energy and thus difficulty in moving and sleeping. The former symptom stems from edema in her lungs and tissues which can be altered by treating with expectorant. There was also some weakness due to the abnormally sized thyroid gland, and inefficient efferent nervous system which prevents the muscles from relaxing normally. The roughness of her skin may also be due to some nervous disorder (possibly shingles). [Med-P #3]

Degree Student

Liver malfunction causes lowered protein production resulting in edema. This causes difficulty in breathing (which yields drowsiness) and the edema causes the hoarse voice (pressure on the larynx). The swelling could also have caused this patient's lack of appetite and perhaps her constipation. The iodide (in the expectorant) may also contribute to the constipation. The iodide (in the expectorant) likely caused the hypertrophy of the thyroid. This exacerbated the hoarseness of voice. [Degree #2]

Additionally, it should be noted that while all three groups of students 'mistakenly' provided pathophysiology-like diagnoses, this was more characteristic of the CEGEP and Med-P students than of the Degree students. These students were more inclined to provide treatment-like diagnoses; an error which none of the Med-P students committed.

A total of eleven students provided an inaccurate diagnosis for the Hashimoto's Hypothyroidism text. Eight of these, however, were at least related to some form of thyroid problem. This includes all three inaccurate diagnoses provided by the CEGEP students, two of the four inaccurate diagnoses provided by the Med-P students, and three of the four inaccurate diagnoses provided by the Degree students. Converting these figures into percentages based on students within a group who, first, provided what could accurately be termed a diagnosis and, second, emphasized the major malfunctioning organ, it was found that the CEGEP students more frequently (100%) emphasized the thyroid problem of Text 1, followed by the Degree students (75%) and, lastly, the Med-P students (60%).

In reference to the CEGEP group, one student diagnosed the patient as having either a thyroid tumor or goitre. Goitre is simply a more sophisticated term used to refer to the enlargement of the thyroid gland (i.e., it is not the cause of the abnormally sized gland). While a tumor in the thyroid gland could cause goitre, this is not the cause in the present situation. Both of these diagnoses were also provided by students in the other two groups. Specifically, a diagnosis of goitre was provided by two Degree students and a diagnosis of a tumor in the thyroid was provided by one Med-P student.

A second CEGEP student provided two diagnoses; hyperthyroidism, and some form of heart problem. The former diagnosis reflects the student's concern over the physical size of the thyroid gland as opposed to its function. That is, the prefix 'hyper' is typically used to refer to something in excess and this student appears to have diagnosed the patient's problem based on the physical enlargement of the thyroid gland. A patient with hyperthyroidism, however, is actually suffering from an excess amount of thyroid hormones and the symptoms which are manifested are in the opposite direction to those manifested in a patient with hypothyroidism. This incorrect diagnosis was also provided by one Degree student.

The third CEGEP student who provided an inaccurate diagnosis had diagnosed the patient as having a hormonal disturbance. Although this is true, it was considered too general to be accurate. This was also a diagnosis provided by one Med-P student.

The remaining three diagnoses which were incorrect were all related to some form of heart problem. Specifically, one Degree student and one Med-P student diagnosed the patient as having congestive heart failure, while one additional Med-P student provided a diagnosis of arteriosclerosis (a disorder characterized by hardening of the arteries).

Only three students in total provided a diagnosis which could be considered accurate. This included two CEGEP students and one Med-P student. The diagnoses provided by the CEGEP students included a malfunctioning of the thyroid gland, and some form of thyroid condition, both of which are very general diagnoses. The Med-P student diagnosed the patient as having hypothyroidism. Although there was no reference to the patient's pre-coma state, and thus the diagnosis was not complete, it was the most accurate diagnosis provided.

Text 2: Pericardial Effusion

The general findings noted for students' diagnoses for the Hashimoto's Hypothyroidism text extend to the Pericardial Effusion text as well. This includes the large percentage of erroneous diagnoses and a misconception of what a diagnosis actually is.

All of the diagnoses which were provided by students for this text were inaccurate (Table 6). Compared to the Hashimoto's Hypothyroidism text, fewer students (particularly from the CEGEP group) even attempted to provide a diagnosis. This is revealed by comparing both the combined total for the 'incorrect' and 'correct' diagnoses for Text 1 and Text 2 (14 and 10, respectively) and the total for the 'no diagnosis' category (5 and 9, respectively).

Students continued to misinterpret the task of providing a diagnosis, as evidenced by the number of diagnoses which are more accurately characterized as treatment-management plans or extensions of pathophysiology explanations. All of the treatment-related diagnoses provided for the Pericardial Effusion text were provided by the same CEGEP and Degree students who had provided treatment-related diagnoses for the Hashimoto's Hypothyroidism text. The CEGEP students who provided pathophysiology-related diagnoses for the Pericardial Effusion text included the same two students who had provided this type of diagnosis for Text 1 and one student who had, interestingly, provided an accurate diagnosis for Text 1. For the Med-P group, only two students provided a pathophysiology-related diagnosis for the Pericardial Effusion text as compared to the three Med-P students who did this for Text 1. This additional student neglected to provide a diagnosis for Text 2. Finally, the one Degree student who had provided a pathophysiology-related diagnosis for Text 2, had also provided one for Text 1. This consistency across subjects suggests a profound misconception of what a diagnosis entails.

Four of the ten diagnoses which were inaccurate for the Pericardial Effusion text were at least related to some form of heart problem. This excluded the one inaccurate diagnosis provided by a CEGEP student but included two of the six inaccurate diagnoses provided by the Med-P students (congestive heart failure, mild cardiac infarction) and two of the three diagnoses provided by the Degree students

TABLE 6

Diagnoses Provided by CEGEP, Med-P, and Degree Students for Text 2: Pericardial Effusion

	GROUP		
	CEGEP STUDENTS (n=10)	MED-P STUDENTS (n=9)	DEGREE STUDENTS (n=10)
DIAGNOSES			
No Diagnosis	5	1	3
Treatment Related	1	0	3
Pathophysiology Related	3	2	1
Incorrect	1	6	3
Correct	0	0	0

(arteriosclerosis, circulatory problem). In terms of percentages, 0% of the CEGEP students, 35% of the Med-P students and 66% of the Degree students provided what could accurately be termed a 'diagnosis' referred to the patient's heart problem.

Additional inaccurate diagnoses included a kidney disorder (one Med-P and one Degree student), a blood disorder (two Med-P students), and a problem with the oxygen transport system (one CEGEP student).

General Discussion

The analysis of students' diagnoses indicated that many students can indeed generate a diagnosis but that this ability is quite limited both within and across groups of students. Collapsing across texts and accuracy of diagnoses (i.e., the accurate and inaccurate diagnosis categories), the percentage of students in the CEGEP group and the Degree group who provided diagnoses was comparable (30% and 35%, respectively). The students in the Med-P group demonstrated a superior ability at generating diagnoses, but this ability continued to be limited, encompassing only 56% of the students. These figures, however, may underestimate students' ability because the percentages were calculated on the basis of the performance of all students, including those who neglected to provide a diagnosis. Since information on students' diagnositic ability is not directly available under this latter condition (e.g., students may be able to generate a diagnosis but simply could not generate a plausible one for the present case(s) and thus did not attempt a diagnosis), the percentage of students who provided diagnoses was recalculated, excluding those students who neglected to provide a diagnosis. Based on these calculations, the percentage of CEGEP, Degree, and Med-P students who provided a diagnosis increased to approximately 41%, 46%, and 69%, respectively.

Overall, there was a higher percentage of students who provided a diagnosis for the Hashimoto's Hypothyroidism text (48%) than there was for the Pericardial Effusion text (38%). This difference, however, can be largely accounted for by the higher percentage of students who neglected to provide a diagnosis for Text 2. This suggests that the ability of students to generate a diagnosis may indeed be underestimated if calculated on the performance of every student, including those who do not provide a diagnosis.

Perhaps the more interesting finding was that a large percentage of students both within and across groups did **not** have an accurate conception of what a diagnosis entails. This was evidenced by the number of students who consistently attempted to prescribe some form of treatment-management plan for the patient or, alternatively, who attempted to explain some of the patient's symptoms as opposed to identifying the patient problem as an example of a particular disease. Since this misconception was evidenced by students in each of the three groups, it is concluded that this is a general characteristic of the novices in this sample. Previous studies (e.g., Claessen and Boshuizen, 1985; Norman et al., 1985) which have examined the diagnostic ability of medical students have been surprisingly vague in their descriptions, restricting their results to the percentage of students who generated accurate and inaccurate diagnoses; no further information is provided concerning the nature of the inaccurate diagnoses. While the findings of the present study may be unique to the level of students studied, thus dismissing this apparent oversight of previous researchers, it remains an issue to be further examined. Until then, the findings of the present study can be viewed as contributory to the existing literature in that they provide some insight into the types of errors first year medical students commit when they attempt to diagnose a clinical case.

Although the identification of the major malfunctioning organ was not an explicit requirement of the students in this study, estimates of their ability to do so were derived by examining the content of their diagnoses, irrespective of accuracy. Collapsing across the two texts and accuracy of diagnosis, 58% of the students were able to identify the major organ which was the source of the patient's problem The thyroid gland problem, which the first text revolved around, was more frequently identified by all groups of students than was the heart problem of the second text. This finding may be due to the differential saliency of information contained in the two texts regarding the major malfunctioning organ. Text 1 included the clinical finding that the thyroid gland was enlarged to twice the normal size and felt firm and irregular. The only comparable clinical feature of Text 2 was the enlarged cardiac silhouette. Previous studies (e.g., Patel and Frederiksen, 1984b) have noted that 'novices' typically focus on dramatic aspects of a clinical case. At an intuitive level, the explicit text-based information referring to the thyroid problem of Text 1 appears more salient than the heart problem of Text 2. The superior ability of students to identify the major malfunctioning organ of Text 1 over Text 2 concurs with this intuitive notion of saliency.

This differential ability to discern the major malfunctioning organ was also noted within groups, as well as collapsing across groups. The CEGEP students were superior at identifying the thyroid problem of Text 1 but, the Degree students were superior at identifying the heart problem of Text 2.

A total of only three accurate diagnoses were provided, and all three were provided for the Hashimoto's Hypothyroidism text. While it may seem surprising that two of these diagnoses were provided by CEGEP students, it should be noted that their diagnoses involved the identification of the major malfunctioning organ (e.g., some form of thyroid condition) as oppossed to the assignment of a clinical name to the patient's problem.

Both at an intuitive level, and subsequent to examining the pathophysiology protocols of all students in this study, it is not surprising that the overwhelming majority of the diagnoses were inaccurate. Generating an accurate diagnosis for a clinical case involves, minimally, discriminating relevant from nonrelevant case information, synthesizing this information into a coherent problem model and, finally, differentiating between various diseases. The text by relevance interaction noted in the comprehension analysis revealed that these students, as a whole, could discriminate relevant from nonrelevant information to some extent. While this discrimination was more evident in the Hashimoto's Hypothyroidism text than in the Pericardial Effusion text, the total percentage of propositions which formed the basis for both recalls and inferences was relatively low for all groups. The group by relevance interaction approached significance and indicated that the Degree students were slightly more adept than the other students at discriminating the relevant from the nonrelevant information.

The pathophysiology protocols generated by these students, however, suggested that they did not fully comprehend the significance of this relevant information. Their problem models, while varying from group to group, contained explanations which were limited in focus and were either inaccurate or, at best, accurate in a very general sense. Presumably, this latter condition arises because the students who have some basic science knowledge have not yet learned to apply it in a clinical context.

Finally, the ability to differentiate between various diseases involves both an awareness of the clinical names used to refer to them and knowledge of the signs and symptoms which characterize these diseases. Students at this level have not yet been exposed to the plethora of clinical diseases which exist. The possibility that any of these students have heard of Hashimoto's Hypothyroidism or Pericardial Effusion, let alone their associated complications (e.g., myxodema coma, cardiac tamponade) is minimal. The possibility that these students have some knowledge about the signs and symptoms typically associated with these diseases is even less probable. Thus, it should not be surprising that only one student was able to accurately diagnose one of the clinical texts.

CHAPTER VII

"The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly."

(Ausubel, 1968, p.vi)

CONCLUSIONS AND IMPLICATIONS

In summary, this study was an exploratory attempt at evaluating the effect of type of premedical education on clinical case comprehension and problem solving in a population of students upon entry into medical school. Previous studies which have been conducted using comparable populations of students have been concerned with evaluating the effectiveness of accelerated versus nonaccelerated medical programs as measured by student performance on department-generated and/or standardized examinations. Other studies have been concerned with the underlying cognitive processes involved in performing complex tasks (such as those employed in the present study) and have compared medical students early in their studies (i.e., novices) with more advanced students or post-graduates of medical school (i.e., experts). These latter studies, however, have neglected to consider the premedical background of the students comprising the designated novice (and expert!) population. The present study, then, is unique in that it draws from and supplements two areas of research in medical education. The results suggest that there are **both** quantitative and qualitative differences and similarities between accelerated and nonaccelerated medical students upon entry into medical school; a finding which was previously obscured because researchers either grouped students according to their academic year, irrespective of type of premedical education, or used materials which were designed to identify only quantitative differences between these groups of students.

The analysis of students' summary protocols in the present study, for example, indicated that the three groups of students had different representations of the clinical texts. More specifically, the Med-P students recalled and inferred more propositions from the two texts combined than either the CEGEP students or the Degree students. While it can be suggested that the Med-P students represented the texts at a more abstract level than the other two groups of students, there was also a trend for the Degree students to be the most selective in terms of which propositions (i.e., relevant or nonrelevant) formed the basis for recall and/or inference. Determining which group is superior or more advanced in their representation of clinical cases would be premature and perhaps constitutes an issue of minor importance at this time. The crux of these results is that type of premedical education does influence students' representation of clinical texts. A more immediate issue for cognitive scientists, then, lies in determining the extent to which previous expert-novice studies have erroneously characterized the cognitive behavior of 'novices'.

The analysis of students' pathophysiology explanations of the two clinical texts also yielded group differences. The marked improvement from the CEGEP students to the Med-P students suggests that the one-year medical preparatory program has a substantial and positive impact. For example, much of the 'common knowledge' which was characteristic of the CEGEP students' explanations appears to have dissipated and been replaced with basic science knowledge. While much of this knowledge is not yet accurate in the specific context of the clinical case, some of the concepts have been mastered at a general level. Additionally, students at the Med-P level appeared more aware of the interconnectedness between different body systems and functions as compared to the CEGEP students as evidenced by the former's more highly structured and integrated explanations. This awareness, however, was even more acute in the group of Degree students.

Finally, this study provided some insight into the diagnostic ability of students upon entering medical school. While it is readily apparent that such a task is beyond the scope of students' ability at this level (as evidenced by the extremely low level of accuracy), requesting them to do so yielded some interesting information. Specifically, approximatley one-half of the students within each of the three groups did attempt to provide a diagnosis for the two clinical cases and, despite their inaccuracy, a number of these diagnoses indicated that the students could, at least, identify the major malfunctioning organ. Since this is a skill which some of these students already possess (at least at a rudimentary level), perhaps this would be a good building block from which to teach more advanced diagnostic skills. However, before such an approach is implemented, the generalizability of this skill should first be examined through the use of additional clinical problems describing patients with disorders involving other major organs. Further, the possibility that this skill is influenced by the manner in which the case information is presented (i.e., the saliency of the cues) suggests that this is a factor to be manipulated and examined in future studies.

To conclude, the overall results of the present study warrant a continuation of this line of research. It has been shown that differences in premedical education are reflected in 'novices" representation, explanation, and diagnosis of clinical texts. A question which remains unanswered is "To what extent do the initial differences between groups with varying premedical backgrounds persist?" The answer to this question requires some developmental research involving either a longitudinal study in which these same students are tested (in a similar manner as the present study) at successive points in their medical studies, or a cross-sectional study in which groups of students with premedical backgrounds comparable to those students used in the present study, but who differ in terms of academic year, are tested and compared. A comparison of the results obtained in the present study with those obtained through a developmental study should determine if these group differences persist and, if so, perhaps the duration of these differences (i.e., do these groups of students remain heterogeneous or do they become more homogeneous in nature?). The possibility also exists that such a comparison will reveal qualitative differences of a nature other than those noted in the present study.

In addition to conducting a developmental research study, a useful modification of the research design used in the present study would be the inclusion of an additional group of students, specifically, students with an undergraduate degree in the social sciences or humanities. The comparative performance of these students would be of interest for two reasons. First, students with this type of premedical background are becoming more prevalent in medical schools around the country (e.g., Thomae-Forgues and Erdmann, 1980) and, while researchers have compared these students to traditional students (i.e., medical students with a B.Sc. degree), these studies have, again, restricted their comparisons to academic performance on department-generated and/or extramural examinations (e.g., Dickman, Sarnacki, Schimpfhauser, and Katz, 1980). Further, simply dichotomizing students based on their premedical undergraduate major (e.g., science or nonscience) may not be an adequate means of assessing the effect of type of premedical eduaction on students' subsequent performance in medical school. The fact that students do not major in one of the natural sciences at university does not exclude the possibility that they have taken some science courses. Therefore, consideration should also be given to the percentage of a student's coursework which was devoted to science courses. This awareness has recently emerged in the evaluation literature in studies concerned with the impact of type of

Baccalaureate preparation on performance in medical school (e.g., Canaday and Lancaster, 1985; Zeleznik, Hojat, and Veloski, 1983) but has yet to be considered as a possible influential factor in the cognitive performance of 'novices'.

A second reason for the inclusion of medical students with B.A. degrees in future studies is to equate groups of students with different premedical backgrounds in terms of age. Medical students who have first completed some form of undergraduate degree will be more similar in age and experience than students accepted into medical school directly from High School. Thus, maturational differences can be eliminated as a factor influencing the potential differential performance of B.A. and B.Sc. students. Further, since students enrolled in accelerated medical programs have perceived themselves, in retrospect, as lacking in maturity, particularly during the clinical phase of their studies (e.g., Patel, Dauphinee, Medley-Mark, 1984), this factor may be of particular importance in a developmental study on the effects of type of premedical education on subsequent performance in medical school.

Finally, it is suggested that future studies include additional clinical texts both from the same specialty areas as those used in the present study (i.e., cardiology and endocrinology) as well as from areas outside of these specialties. This manipulation would serve to determine the generalizability of the findings from the present study and, when used in combination with the previous suggestions, should provide a more accurate profile of the cognitive behavior of 'novices' in the domain of medicine and generate practical suggestions for their instruction.

References

Asper, S.P. (1964). The training of the physician. New England Journal of Medicine, 271, 447-451.

Ausubel, D.F. (1968). Educational psychology: A cognitive view. NewYork: Holt, Rinehart, and Winston.

Barrows, H.S., Neufeld, V.R., Feightner, J.W., & Norman, G.R. (1977). An analysis of the clinical methods of medical students and physicians. Final Report, Ontario: Ontario Ministry of Health.

Beran, R.L. (1979). The rise and fall of three-year medical school programs. Journal of Medical Education, 54, 248-249.

Blaustein, E.H., & Kayne, H.L. (1980). Boston University and accelerated medical education: The first five cohorts. Journal of Medical Education, 55, 202-204.

Blumberg, M.S. (1971). Accelerated program of medical education. Journal of Medical Education, 46, 643-651.

Campbell, C., & DeMuth, G.R. (1976). The University of Michigan integrated premedical-medical program. Journal of Medical Education, 51, 290-295.

Canaday, S.D., & Lancaster, C.J. (1985). Impact of undergraduate courses on medical students' performance in basic sciences. Journal of Medical Education, 60, 757-763.

Carnegie Commission on Higher Education (1970). Higher education and the nation's health. New York: McGraw-Hill.

Charness, N. (1979). Components of skill in bridge. Canadian Journal of Psychology, 1, 33, 1-16.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. Cognitive Psychology, 4, 55-81.

Chi, M.T.H., Feltovich, P.J., & Glaser, R (1981). Categorization and representation of physics problems by experts and novices. Cognitive Science, 5, 121-152.

Chiesi, H.L., Spilich, G.J., & Voss, J.F. (1979). Acquisition of domain related information in relation to high and low domain knowledge. Journal of Verbal Learning and Verbal Behavior, 18, 257-274.

Claessen, H.F.A., & Boshuizen, H.P.A. (1985). Recall of medical information by students and doctors. Medical Education, 19, 61-67.

Collins, J., Brown, J.S., & Larkin, K.M. (1980). Inferences in text understanding. In Spiro, Bruce, and Brewer (Eds.), Theoretical issues in reading comprehension. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Coughlin, L.D., & Patel, V.L. (1985). Adapting a paradigm from cognitive science to medical education: Problems and possible solutions. Proceedings of the Twenty-fourth Annual Conference on Research in Medical Education, (pp. 97-102).

Daubney, J.H., Wagner, E.E., & Rogers, W.A. (1981). Six-year B.S./M.D. programs: A literature review. Journal of Medical Education, 56, 497-503.

deGroot, A.D. (1965). Thought and choice in chess. Mouton Publishers, The Hague.

Dickman, R.L., Sarnacki, R.E., Schimpfhauser, F.T., & Katz, L.A. (1980). Medical students from natural science and nonscience undergraduate backgrounds. Journal of the American Medical Association, 243, 2506-2509.

Dillinger, M. (1984). Segmentation and clause analysis. Discourse processing laboratory, Report No. 3. Montreal, Quebec: McGill University, Department of Educational Psychology.

Eisenstadt, M., & Kareev, Y. (1975). Aspects of human problem solving: The use of internal representations. In Norman and Rumelhart (Eds.), Explorations in cognition. San Francisco: Freeman Press.

Eisenstadt, M., & Kareeve, Y. (1977). Perception in game playing. In P.N. Johnson-Laird & P.C. Wason (Eds.), Thinking. Cambridge: Cambridge University Press.

Elstein, A.S., Shulman, L.S., & Sprafka, S.A. (1978). Medical problem solving: An analysis of clinical reasoning. Cambridge, Massachusetts: Harvard University Press.

Feltovich, P.J. (1981). Knowledge based components of expertise in medical diagnosis. (Contract No. N00014-79-C-0215). Pittsburgh: University of Pittsburgh, LRDC.

Feltovich, P.J. (1983). Expertise: Reorganizing and refining knowledge use. Professions Education Research Notes, 4, 5.

Feltovich, P.J., & Barrows, H.S. (1984). Issues of generability in medical problem solving. In H.G. Schmidt & M.L. deVolder (Eds.) Tutorials in Problem Based Learning. Van Gorcum, Maastricht.

Flair, M.D. (1969). Honors program in medical education at Northwestern University. Journal of Medical Education, 44, 1127-1131.

Frederiksen, C.H. (1975). Representing logical and semantic structure of knowledge acquired from discourse. Cognitive Psychology, 7, 371-458.

Frederiksen, C.H. (1979). Discourse comprehension and early reading. In L. Resnik & P. Weaver (Eds.), Theory and Practice of Early Reading. Hillsdale, N.J.: Lawrence Erlbaum and Associates.

Frederiksen, C.H. (1985). Cognitive models and discourse analysis. In C.R. Cooper & S. Greenbaum (Eds.), Written Communique Annual, Vol. 1: Linguistic approaches to the study of written discourse. Beverley Hills, CA: Sage.

Gellhorn, A. & Scheuer, R. (1978). The experiment in medical education at the City College of New York. Journal of Medical Education, 53, 574-582.

Glaser, R. (1985). Thoughts on expertise, (Tech. Rep. No. 8). Pittsburgh, PA.: University of Pittsburgh, LRDC.

Greeno, J.G. (1978). A study of problem solving. In R. Glaser (Ed.), Advances in Instructional Psychology (Vol. 1) Hilldale, New Jersey: Lawrence Erlbaum Associates.

Groen, G.J., & Patel, V.L. (1985). Medical problem-solving: Some questionable assumptions. Medical Education, 19, 95-100.

Grossman, W.K., Conly, S.S., Jr., Menduke, H., & Graff, H. (1972). Medical students view their accelerated program. Journal of Medical Education, 47, 287-288.

Halliday, M. (1967) Notes on transitivity and theme in english. Part 2, Journal of Linguistics, 3, 177-274.

Harrison, T.R. (1980). Principles of internal medicine , 3rd Edition. N.Y.: McGraw-Hill Book Company.

Herbut, P.A., Sodeman, W.A., Conly, S.S., Jr., & Ascah, R.G. (1969). The Jefferson-Penn State accelerated medical student program. Journal of Medical Education, 44, 1132-1138.

Hubbard, J.P. (1978). Measuring medical education: The tests and the experience of the National Board of Medical Examiners., 2nd Edition. Philadephia.: Lea and Febiger.

Joseph, G.-M., & Patel, V.L. (1986). Specificity of expertise in medical reasoning. Proceedings of the Annual Conference of the Cognitive Science Society, (pp. 331-343), N.J.: Lawrence Erlbaum Associates.

Kanter, G.S. (1969). The Rensselaer Polytechnic Institute-Albany medical college six-year biomedical program. Journal of Medical Education, 44, 1139-1143.

Kassirer, J.P., Kuipers, B.J., & Gorry, G.A. (1982). Toward a theory of clinical expertise. The American Journal of Medicine, 73, 251-259.

Keefer, S.C. (1964). The training of the physician: Experiment with the medical-school curriculum at Boston University. New England Journal of Medicine, 271, 401-403.

Kintsch, W. (1974). The representation of meaning in memory. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Lanzoni, V., & Kayne, H.L., (1976). A report on graduates of the BostonUniversity six-year combined liberal-arts-medical program. Journal of Medical Education, 51, 283-289.

Larkin, J.H., McDermott, J., Simon, D.P., & Simon, H.A. (1980). Models of competence in solving physics problems. Cognitive Science, 4, 317-345.

LeClere, H., & Bordage, G. (1984). Text comprehension among medical students and experienced physicians: A preliminary study. Proceedings of the Twenty-third Annual Conference on Research in Medical Education, (pp. 133-138). Washington, D.C.: American Association of Medical Colleges.

Matlack, D.R. (1972). Changes and trends in medical education. Journal of Medical Education, 47, 612-619.

Medical education in the United States 1969-1970. Journal of American Medical Association, 214, 1483-1581.

Miller, P.B. (1975). Strategy choice in medical diagnosis. Report No. TR-153, Massachussets: Massachussets Institute of Technology.

Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (Ed.), The Psychology of Computer Vision; New York: McGraw Hill.

Muzzin, L.J., Norman, G.L., Jacoby, L.L., Feightner, J.W., Tugwell, P., & Guyatt, G.H. (1982). Manifestations of expertise in recall of clinical protocols. Proceedings of the Twenty-first Annual Conference on Research in Medical Education, (pp. 163-168).

Norman, G.R., Jacoby, L.L., Feightner, J.W., & Campbell, E.J.M. (1979). Clinical experience and the structure of memory. Proceedings of the eighteenth Annual Conference on Research in Medical Education, Washington, D.C., (pp. 214-218).

Norman, G.R., Muzzin, L.J., & Rosenthal, D. (1985). Expert-novice differences in perception and categorization in dermatology. Paper presented at the American Educational Research Association, Chicago.

Osborne, R. (1984). Children's dynamics. The Physics Teacher, 22, 504-508.

Patel V.P. (in Press). The role of prior knowledge in expert/novice differences in comprehension of medical texts. Memory and Language.

Patel, V.P., Arocha, J.F., & Groen, G.J. (1986). Strategy selection and degree of expertise in medical reasoning. Proceedings of the Eighth Annual Conference of the Cognitive Science Society.

Patel, V.L., Dauphinee, W.D., & Medley-Mark, V. (1984). Premedical education: A predictor of medical graduates future? (Report No. 4). Montreal, Quebec: McGill University, Center for Medical Education.

Patel, V.L., & Frederiksen, C.H. (1984 a). Cognitive processes in comprehension and knowledge acquisition by medical students and physicians. In H.G. Schmidt & M.L. deVolder (Eds.), Tutorials in Problem-Based Learning, Van Gorcum, Maastricht.

Patel, V.L., & Frederiksen, C.H. (1984 b). Role of prior knowledge in expert-novice differences in comprehension of medical texts. (Report No. 5). Montreal, Quebec: McGill University, Center for Medical Education.

Patel, V.L., & Groen, G.J. (1986a). Knowledge-based solution strategies in medical reasoning. Cognitive Science, 10, 91-115.

Patel, V.L., & Groen, G.J. (1986b). The role of mental models in causal explanations by medical students. To appear in Cognition and Instruction.

Patel, V.L., Groen, G.J., & Frederiksen, C.H. (1986). Differences between students and physicians in memory for clinical cases. Medical Education, 20, 3-9.

Patel, V.L., Groen, G.J., & Scott, H.M. (in Press). Biomedical knowledge and clinical reasoning: Role of situation models. In D. Evans & V.L. Patel (Eds.), Cognitive Sciences in Medicine. Josiah Macy Jr. Foundation: N.Y.: Academic Press.

Patel, V.L., HoPingKong, & Mark, V. (1984). Role of prior knowledge in comprehension of medical information by medical students and physicians. Proceedings of Research in Medical Education, Association of American Medical Colleges, Chicago, (pp. 127-132).

Patel, V.L., & Medley-Mark, V. (1986). Relationship between representation of textual information an underlying problem representation in medicine. (Report # CME86-CS1). Montreal, Quesbec: McGill University, Centre for Medical Education.

Patil, R.S., & Szolovits, P. (1981). Causal understanding of patient illness in medical diagnosis. Proceedings of the Seventh IJCAT, Vancouver, B.C., Canada.

Pellegrino, E.D. (1980). Pruning an old root: Premedical science and premedical school. Journal of American Medical Association, 243, 2518-2519.

Pritchard, H.N. (1976). The Lehigh-Medical college of Pennsylvania six-year B.A.-M.D. program. Journal of Medical Education, 51, 296-298.

Reitman, J.S. (1976). Skilled perception in GO. Cognitive Psychology, 8, 336-356.

Swanson, A.G. (1972). The three-year medical school curriculum. Journal of Medical Education, 47, 67.

Thomas, L. (1978). Notes of a biology watcher: How to fix the premedical curriculum. New England Journal of Medicine, 298, 1180-1181.

Thomae-Forgues, M., & Erdmann, J.B. (1980). Datagram. Journal of Medical Education, 55, 971-972.

Tierney, R.J., & Mosenthal, J. (1982). Discourse comprehension and production: Analyzing text structure and cohesion. In J. Langer, & M. Smith-Burke (Eds.) Reader meets author / bridging the gap: A psycholinguistic and sociolinguistic perspective, Newark, D.E.:International Reading Association.

Vaisrub, S. (1978). Medical education: Preload and afterload. Journal of American Medical Association, 240, 1755.

Voss, J.F., Greene, T.R., Post, T.A. & Penner, B.C. (1986). Problem-solving skill in the social sciences. In G. Bower (Ed.), The psychology of learning and motivation (Vol. 17) New York: Academic Press, (in press).

Winograd, T. (1972). Understanding natural language. Edinburgh: Edinburgh University Press.

Winograd, T. (1983). Language as a cognitive process: Volume 1: Syntax. Reading, Mass: Addison-Wesley Publishing Co. Inc.

Zeleznik, C., Hojat, M., & Veloski, J. (1983). Baccalaureate preparation for medical school: Does type of degree make a difference? Journal of Medical Education, 58, 26-33.

Footnotes

1 Concern is frequently expressed when a study is conducted with a relatively 'small' sample size. Since this is an aspect of the present study, an attempt is made to clarify why this was done and why such a tactic does not necessarily detract from or limit the results.

As outlined in the body of this thesis, the method of analysis used to assess subjects' comprehension of the clinical texts is quite detailed. Such an analysis is very demanding and time consuming. In lieu of this, a larger sample size is not considered to be feasible.

In terms of statistical analyses, sample size is directly related to the power or sensitivity of an experiment. While it is true that the larger the sample size, the greater the power or sensitivity of the experiment in detecting treatment differences in the population, it should be noted that if differences are detected using a smaller sample size (i.e., using a less sensitive and less powerful measure of treatment effects), they are probably that much more real and significant.

Taken together, then, the use of a relatively small sample size in the present study would appear justified.

2 The Med-P group was initially comprised of ten students, as are the other two groups. However, the tenth Med-P subject in this study was excluded because s/he was unable to complete the booklet within one sitting (due to a conflict in schedules). The stopping point of this student was after s/he had read the second clinical text and had summarized it. Thus, to resume testing would have involved having this subject either re-read this second text or simply carry on from where s/he had left off. Since

both situations could have possible confounding effects, this subject was excluded from the study. In addition, due to the time constraints necessarily imposed on this endeavour, a replacement subject was not possible.

APPENDIX A

Centre for Medical Education
McIntyre Medical Sciences Building
Room 529
McGill University

There are two clinical cases, each provided with:

(1) the history of a patient,
(2) the physical examination findings of that patient, and
(3) the laboratory findings of that patient.

1. **Read** the first case. When you have finished, turn the page over and do not refer to it again at any point during the study.

2. **Summarize** the case.

3. **Explain** the case in terms of the underlying pathophysiology.

4. Provide a **diagnosis.**

5. **Repeat** (1) through (4) above for the second clinical case.

Please write your answers in **longhand**. Do not use short note form.

Thank you for volunteering your time to help us with our study.

Birthdate:______________________________

If you hold an undergraduate degree, please specify type (e.g., B.Sc.) and your major.

____

APPENDIX B

CASE 1

A 63 year-old woman with a one-week history of increasing drowsiness and shortness of breath was brought to the emergency room by her daughter. The patient had not been well for over a year. She complained of feeling tired all the time, had a loss of appetite, a 30 lb. weight gain and constipation. A month later she had been diagnosed as having 'chronic laryngitis' and was prescribed a potassium iodide mixture as an expectorant.

Physical examination revealed a pale, drowsy, obese lady with marked periorbital edema. She had difficulty speaking, and when she did speak her voice was noted to be slow and hoarse. There were patches of vitiligo over both her legs. Her skin felt rough and scaly. Her body temperature was 36 deg. C. Pulse was 60/minute and regular. B.P. was 160/95. Examination of her neck revealed no jugular venous distention. The thyroid gland was enlarged to approximately twice the normal size. It felt firm and irregular. There was grade 1 galactorrhea. The apex beat could not be palpated. Chest examination showed decreased movements bilaterally and dullness to percussion. There was no splenomegaly. Neurological testing revealed symmetrical and normal tendon reflexes but, with a delayed relaxation phase. Urinalysis was normal. Chest X-ray showed large pleural effusions bilaterally. ECG revealed sinus bradycardia, low voltage complexes and non-specific T-wave flattening. Routine biochemistry (SMA=16) showed Na=125, K=3.8, BUN=8mg/100ml. Arterial blood gases PO_2=50 mm Hg, PCO_2=60 mm Hg. The patient was admitted to the intensive care unit for further management.

APPENDIX C

CASE 2

This 62 year-old retired Air Force mechanic was apparently well until about 5 months before presenting to the hospital. He then noted he was 'winded' after walking about 40 feet. He was increasingly breathless lying down, tried using 4 pillows to sleep and most recently is sleeping sitting up. He has occassionally awoken extremely short of breath. He has a mild non-productive cough and agrees that his voice is a little hoarse. During this time his legs have been swelling. His appetite has decreased yet his abdomen has increased and he has gained weight. He says "no food tastes good" and he has constant mild nausea but has not vomited. He has had no chest or abdominal pain. He does not smoke, drinks alcohol socially but less lately. His only admission to the hospital was for a heart attack 12 years ago. He recovered completely and was walking 6 miles a day a year ago. He is taking no medication.

On examination: H.R. 80/min. and regular. B.P. 120/98 mm Hg. Pulsus paradoxicus 12 mm Hg. No cyanosis. Pronounced peripheral edema of legs and presacrum. Some edema over abdominal wall and scrotum. Abdomen was large with shifting dullness and a fluid wave was demonstrated. Liver edge was smooth, 3 cm. below the right costal margin. Spleen was not palpated. No masses. Jugular veins distended to the angle of the jaw at 45 deg., apex not palpable, heart sounds faint, no S3, no S4, no murmurs. Some dullness to percussion at right lung base. Breath sounds diminished at both lung bases with decreased chest expansion. Fine end inspiratory crepitutions noted. Remainder examination was normal.

Hb=13.5 gm%, WBC=5,500 with a normal differential. Prothrombin time 12.5 (control 11.8), P.T.T. 34 (control 34), T4=7.5 (normal 4.5 - 10.5). Urinalysis was normal except urobilinogen 4.0 (normal 0.1 - 1.0). SMA 16 normal except: Albumin 3.5 (N=3.7 - 4.9), total bilirubin 1.7 (N=0.2 - 1.0), alkaline phosphotase 169 (N=30 -

105). Chest X-ray: "Enlarged cardiac silhouette, no evidence of pulmonary edema, right pleural effusion, partial atalectasis in right lower lobe". ECG: remote inferior myocardial infarction. Diffuse ST sagging with T-wave inversion. Generally low voltage QR's with voltage fluctuation.

This patient has been referred from an outlying hospital for definitive management.