

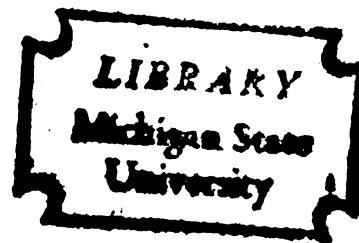
SEX AS A DETERMINANT OF RESPONSES TO
PATIENT MANAGEMENT PROBLEMS BY PHYSICIANS
AND MEDICAL STUDENTS

Dissertation for the Degree of Ph. D.

MICHIGAN STATE UNIVERSITY

DIANNE SEAVER SINGLETON

1975



This is to certify that the
thesis entitled
SEX AS A DETERMINANT OF RESPONSES TO PATIENT
MANAGEMENT PROBLEMS BY PHYSICIANS AND
MEDICAL STUDENTS

presented by

Dianne Seaver Singleton

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Counseling, Personnel
Services, and Educational Psychology

Norman L. Stewart

Major professor

Date 5/12/75

~~22276~~

~~0-010~~

~~0-010-213~~

~~0-010~~

N-035

eff

res

ger

up

res

use

nan

for

aff

the

the

the

A re

gest

give

ABSTRACT

SEX AS A DETERMINANT OF RESPONSES TO PATIENT MANAGEMENT PROBLEMS BY PHYSICIANS AND MEDICAL STUDENTS

By

Dianne Seaver Singleton

The purpose of the research was to assess the effect of gender upon management of medical patients by medical students and physicians. The effect of the gender of both the patients and the medical personnel upon responses to written simulations of patient management problems was examined. The instrument which was used included three types of patient management problems, namely suicide, asthma, and multiple sclerosis. Therefore, the type of problem was also examined to see if it affected the treatment of the patients.

The issue of major concern in this study was whether or not women and men patients received inappropriately different medical treatment as a function of their gender or the gender of the medical personnel. A review of health care literature provided some suggestion that gender might affect the quality of care given patients. The anecdotal and demographic

literature suggested that women as patients and women as health care workers are treated less well than men. Experimental evidence of differential treatment for men and women was not available in the literature.

The existence of this presumed inferior treatment needed to be tested. One means of testing this effect of sex upon responses to patient problems was through the use of a tight research design in which sex was the only variable. In order to do this, the three patient management problems were divided into two parallel forms. On Form A the three patients were women and on Form B the three patients were men. Form A and Form B were randomly assigned to the subjects. The problems approximated actual clinical reality as closely as possible. They required the respondent to make judgments and gather data as if he or she were in a clinical setting. The respondents were asked to assume that they were the physician described at the beginning of each problem. After reading a brief description of the presenting problem, the physician or medical student was asked to make a decision as to what he or she would like to do next. Once the physician selected one of the possible options, he or she was directed to another section of the problem or provided with additional data. This information, either directions or data, was made visible

by the respondent rubbing with a special chemical pen the area in the answer booklet which corresponded to the number of the selection made.

The effect of the gender of the medical personnel, male or female, was assessed along with the effect of the gender of the patients treated, males or females, across the three different types of patient problems. The independent variables in the study were the sex of the medical personnel, the sex of the patients, and the type of patient problem. The six dependent variables measured overall performance and skills in diagnosis, management, information gathering, arranging consultations, and managing the patient in correct sequence.

The study was conducted twice, once with 40 men and women physicians and once with 40 men and women medical students. The studies were analyzed separately, though the design and measures were the same for both studies. A random sample of 120 women and 120 men physicians living in a midwestern state were asked to contribute approximately one hour of their time to respond to Form A or B of the problems. Twenty men and twenty women voluntarily completed the problems. The medical student sample was obtained by posting notices requesting response to three patient management problems on the medical student bulletin boards at a midwestern university. Forty second-year medical students, 16 women

and 24 men, responded to the request and were paid \$5.00 each for their participation.

Observation of the data from a speculative rather than statistically rigorous vantage point revealed some interesting patterns. Based upon observations of mean scores of the groups, it was noted that the effect of the gender of patients and medical personnel was different for the physician sample and the medical student sample. Male physicians did not seem particularly influenced by the gender of the patients, whereas female physicians treated male patients more favorably than female patients. Within the medical student sample, however, both men and women students gave preferential treatment to patients of their same gender.

Despite the apparent effect gender had upon treatment of the patients, no statistically significant differences were found on any of the variables tested. The statistical analysis used was a multivariate approach to repeated measures. In other words, the differences in medical management which seemed to be linked to the gender of the patient and medical personnel were not sufficiently great to be statistically significant. More powerful research, probably with larger samples, is needed before conclusive statements can be made regarding the effect of gender upon medical management.

SEX AS A DETERMINANT OF RESPONSES TO PATIENT
MANAGEMENT PROBLEMS BY PHYSICIANS AND
MEDICAL STUDENTS

By

Dianne Seaver Singleton

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Counseling, Personnel Services,
and Educational Psychology

1975

6-93/37

DEDICATION

To Michael Silverman
always my inspiration
to be true to myself
as he is true to himself
a gifted psychotherapist
an intense person

ACKNOWLEDGMENTS

Special thanks go to my dissertation committee members who were chosen because I like them. Each person has contributed to my learning. I thank Norm Stewart, Chairman, for facilitating my first therapeutic encounter. I also appreciate his willingness to share his delightful family with me. I thank Betty Fitzgerald for being the first professor to challenge me to think critically by asking insightful questions and providing unquestioning support. Her help to me in formulating my questions about the treatment of women by the medical profession was invaluable. I thank Al Rabin for being a true scholar, demanding yet gentle. My thanks also go to John Schneider whose creativity and brilliance never cease to amaze me.

I thank Nolan Singleton, my dear husband, for sharing my excitement during the research project. He gave me comfort, encouragement, and assistance. He remained my constant companion, playing with me when I needed to play and working with me when I needed to work.

I thank Lonny Rosen for assisting me in the development of the multiple sclerosis problem. I

appreciate the contributions of the expert raters: Alan Barnes, John Schneider, Sumi Verma, Arnie Berkman, Janice Lindstrom, Ed Lynn, and Tom Johnson. I had considerable confidence in Sandi Bolt's assistance with the research which included scoring, typing, and sharing enthusiasm. Evelyn Pummill handled the distribution and collection of the problems from the medical students with her own special grace. I thank the medical students and physicians who willingly responded to the patient management problems. I am grateful to Norbert Enzer for giving endorsement to my research.

I thank the Faculty Womens' Association for their excitement about my research and their financial assistance.

Noel Kaestner, my undergraduate psychology advisor, taught psychology in a way that kept alive my curiosity about human behavior.

I appreciate having Docent Corporation and the National Institute of Mental Health grant me permission to use respectively the asthma and suicide patient management problems.

I thank our lovable pets, Rascal, Sabrina, Elderberry, and Muffie for making me laugh.

TABLE OF CONTENTS

Chapter	Page
I. STATEMENT OF THE PROBLEM AND REVIEW OF LITERATURE	1
Purpose	1
Problem	1
Review of Literature.	4
Women as Patients	5
Women in Medical Careers.	12
Summary	18
II. DESIGN AND METHODOLOGY.	20
Design	20
Sample	21
Experimental Procedures.	23
Instrumentation	24
Validity	28
Reliability	34
Scoring	41
Analysis of the Data.	44
Hypotheses	52
Summary	54
III. FINDINGS AND DISCUSSION	56
Alpha Level.	56
Physicians	59
Measure 1.	59
Measures 2, 3, and 4	61
Measure 5.	62
Measure 6.	63

Chapter	Page
Medical Students	64
Measure 1.	64
Measures 2, 3, 4, 5	65
Measure 6.	66
Summary	67
IV. SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS FOR FURTHER STUDY.	69
Summary	69
The Problem	69
Design and Methodology	70
Results	73
Discussion	74
Medical Students	75
Physicians	81
Medical Students and Physicians	82
Implications	85
Limitations of the Study	86
Suggestions for Further Research.	90
In Retrospect	93
APPENDICES	
Appendix	
A. INSTRUCTIONS TO THE EXPERT RATERS	96
B. SCORING KEYS	105
C. LETTERS TO SUBJECTS.	108
REFERENCES	110

LIST OF TABLES

Table	Page
1. Interrater Reliability Coefficients (With All Items Included)	36
2. Interrater Reliability Coefficients (With Items Excluded on Which Experts Had Significant Disagreement)	37
3. Test-Retest Reliability Coefficients	38
4. Means of Physicians' Transformed Data.	57
5. Means of Medical Students' Transformed Data	58
6. Results for Physicians: Measure 1.	60
7. Results for Physicians: Measures 2, 3, 4	61
8. Results for Physicians: Measure 5.	63
9. Results for Physicians: Measure 6.	64
10. Results for Medical Students: Measure 1.	65
11. Results for Medical Students: Measures 2, 3, 4, 5	66
12. Results for Medical Students: Measure 6.	67
13. Medical Students: Rank Order by Groups	76
14. Physicians: Rank Order by Groups	77

LIST OF FIGURES

Figure	Page
1. Overall research design	22
2. Formulas for Proficiency, Efficiency, and Competence Indexes.	45
3. Main focus of the research: Measure 1 . . .	47
4. Additional exploration of the research: Measures 2, 3, 4, 5	49
5. Additional exploration of the research: Measure 6	51
6. Additional exploration of the research for physicians: Measure 5	53

CHAPTER I

STATEMENT OF THE PROBLEM AND REVIEW OF LITERATURE

Purpose

The purpose of this study was to assess the effect of gender upon management of medical patients by medical students and physicians. The effect of the gender of both the patients and the medical personnel upon responses to written simulations of patient management problems was examined. The components of patient management which specifically were investigated were the gathered information, the management plan, the patient diagnosis, the sequence of management actions, and the involvement of the patient in requested consultations. The study evaluated whether or not gender was a determinant of the quality and the percentage of agreement of the medical students' and physicians' responses to patient management problems as compared with experts' responses.

Problem

The health care issue of major concern in this study is whether or not women and men patients receive

inappropriately different medical treatment as a function of their gender or the gender of the medical personnel. Anecdotal and statistical evidence was accumulated to suggest that such differences may exist. Chesler (1972) found differences in mental health care along gender lines. Investigators have marshalled evidence to support their contention that health care for men has been less than optimal, but the literature more often suggests that the treatment of women has been inferior to the treatment accorded men. A reason why this study directs particular attention toward health care for women is contained in the following words of Ellen Frankfort (1972): "The woman-as-patient embodies much of the general anguish of women in the society, as well as the particular frustrations of any patient--female or male."

Statistically, women have had more contact with medical personnel than have men. This difference is explained by the woman's need for medical assistance with pregnancy, childbirth, gynecologic problems, and major responsibility for taking children for health care. Specifically, women have averaged 25% more annual visits to physicians than men, and more than 100% more visits if women's visits to take their children to the doctor were included (Male-feasance, 1970). If this study were to reveal that women received less adequate care

than men, persuasive evidence would be available to suggest that women were treated unjustly. Since women have more frequent contact with medical people, they would have more opportunity to challenge the current medical system and change the current status of care. Women more often than men seek medical attention for routine health maintenance rather than emergencies or illnesses. A major thrust of suggested changes in the medical system is to place more emphasis upon preventive and maintenance health care rather than crisis intervention. Thus, women would be an appropriate group to lobby for increased preventive health care. If women challenged the current medical system, they would have the existing organizations and political structure of the Women's Movement to provide financial and organizational support.

In 1971 the Women's Action Program was created by the Department of Health, Education, and Welfare. The premises which led to the formation of the group as stated by then Secretary of Health, Education, and Welfare, Elliot Richardson, were as follows: "(1) that many changes are needed in the status of women both within the Department and within American society, and (2) that this Department has a unique responsibility and opportunity to help bring about those changes." The first recommendation which came from the 1971

workshop of the Women's Action Program at the National Institute of Mental Health was that research issues concerning the role of women should be explored. Other recommendations made by the Women's Action Program which have had very direct bearing upon this study included the need for research regarding "social labeling, differential diagnoses in clinical treatment, and differential treatment under law" (Report of the Women's Action Program, 1972).

The need for well-designed research to assess health care for women has been well documented with statistical and anecdotal evidence along with reports on the role of women in the field of medicine. When assembled, these bits of data suggest that perhaps women receive medical treatment inferior to that received by men. However, no controlled study had been done to compare treatment of men and women with identical problems. Before considering the procedures which this study used to address the question of whether or not differential care exists for men and women, it is useful to review the literature in which authors suggest that health care for women has been inadequate.

Review of Literature

A considerable body of evidence exists which suggests women receive unequal treatment by the medical profession when compared with treatment given to men.

The evidence essentially is divided into two distinct areas. First, women as patients receive inadequate medical care. The literature regarding women as patients will be reviewed. Next, evidence is presented regarding the inferior role women play within medical careers when compared with men in medical careers. Unfortunately, most of the literature pertinent to gender and health care is conjectural or demographic in nature. Experimental evidence of differential treatment for men and women was not available.

Women as Patients

Feminists have actively made allegations that health care for women is inadequate. As a consequence of the presumed discriminatory medical treatment, women have grouped together to educate themselves about women's health. The Boston Women's Health Book Collective is the most well known of the self-education groups of women because of their book, Our Bodies Ourselves (1971). Their self-education efforts led them to write the book in order to share the knowledge they had acquired. As stated in the book, "We had all experienced similar feelings of frustration and anger toward specific doctors and the medical maze in general, and initially we wanted to do something about those doctors who were condescending, paternalistic, judgmental, and non-informative." The book contains numerous anecdotes of inferior medical

treatment for women. The allegations made by feminists regarding medicine include insufficient preventive medical care for women and children, notably those of low socio-economic status; inappropriate sexual behavior by physicians with their women patients; condescending attitudes held by physicians toward women; lack of women doctors; resistance by physicians to training paraprofessionals to perform simple medical procedures which could upgrade health care; unwarranted mastectomies and hysterectomies; and avoidance of training women in routine health care procedures and female physiology (Hole & Levine, 1971).

Statistical documentation of some of these allegations is available. In support of accusations that maintenance, preventive health care has been inadequate is the fact that only 20% of American women have an annual Pap smear test (Telinde & Mattingly, 1970). This test, the Pap smear, is a highly accurate procedure for the early detection of cervical cancer. There have been 12,000 annual deaths from cervical cancer though the means to detect and treat it in its early stages have been known for approximately 17 years (Boston Women's Health Book Collective, 1971). Though the citizens of the United States are affluent in comparison with people of other nations, in 1968 our infant mortality rate was sixteenth in the world (Myers, 1975). It has

been estimated that at least half of maternal deaths are avoidable. For example, nearly half the women who have delivered children in city and county hospitals have received no prenatal care (Wilson, 1970).

At the federal level, if not at the local level as well, health maintenance programs have not been a high priority for funding consideration. In 1973, maternal and child health, family planning, and neighborhood health centers received funding cuts. The Department of Health, Education, and Welfare announced in 1973 its intention eventually to terminate funding these direct health service programs. State and local funding sources, along with the private market, will have to maintain such programs if they are to survive (Kotelchuck & Levy, 1973).

Women have consumed more prescription drugs and been hospitalized more often than men. Scientists from the National Institute of Mental Health found that women have accounted for 60% of all prescription drug and 67% of psychotropic drug usage. In particular, stimulants and anti-depressants have more often been prescribed for women than men. The differences in psychotropic drug usage decreased with increasing age (Report of the Women's Action Program, 1972). Dr. Carl D. Chambers, former director of research for the New York State

Narcotics Addiction Control Commission, conducted a study in New York which revealed that women took 80% of diet pills, 76% of anti-depressants, and 72% of minor tranquilizers (Frankfort, 1972). Linn and Davis (1971) randomly sampled a group of women and found that 48% of the women 30 years or older said they had taken psychoactive drugs in the preceding year. In 1971 a San Francisco survey found that 45% of the women, and 33% of men had taken psychoactive drugs in the preceding year (Mellinger, Balter, & Manheimer, 1971). Medications for women may not be tested thoroughly before being marketed. The basis for determining that Enovid, the first oral contraceptive marketed in the United States, was safe was data collected on 132 women who had taken the pill for one to three years (Archer, 1970).

From a review of studies concerning the use of prescription drugs the author concluded that a higher percentage of women than men receive prescription drugs. This can be accounted for, in part, by their more frequent contact with physicians. However, when physician contacts for gynecologic or obstetric reasons were deleted, women made 54% of doctor visits. Composing 51.3% of the population and living longer than men, this percentage of visits was not much greater than would be expected. Other studies reported that women reported more symptoms when they saw physicians than did men

(Gurin, Veroff, & Feld, 1960), and whereas women presented more psychological symptoms than men on checklists, men presented more physiological symptoms (Phillips & Segal, 1969). Certainly these factors contributed to the disparate use of prescription drugs by women and men.

Several investigators suggested that sexual biases held by physicians may have partially accounted for this higher use of prescription drugs by women. Examples of their studies follow. One study involved asking physician subjects to describe the "typical complaining patient." In response, 72% of the physicians described a woman, 24% did not mention the sex of the patient, and 4% described a man (Cooperstock, 1971). There also was evidence to suggest that anger toward a patient or pessimism about the outcome of treatment increased the prescription of drugs (Shader, Binstock, & Scott, 1968). Linn (1971) found that 87% of physicians in a sample studied considered daily Librium an appropriate medication for housewife patients, but this figure dropped to 53% when students were the patients. Perhaps an explanation for this difference was represented by Brodsky's (1971) point of view that tranquilizers are an acceptable medication for housewives because they "can take naps and need not be mentally alert."

Further evidence of sexual biases operating in physicians' use of medications has been provided by research entailing content analyses of pharmaceutical advertisements in medical journals over a five-year span (Fidell, 1973; Prather & Fidell, 1972). The advertisements were taken from the New England Journal of Medicine, American Journal of Psychiatry, California Medicine, and Journal of the American Medical Association. The authors' interpretation of the content analyses revealed a strong tendency to link female patients with psychoactive drugs whereas the majority of nonpsychoactive drug advertisements were linked with male patients. The average male patient depicted was older than the average female patient, and the range of male patient ages depicted was greater than the range of women patients' ages. When females were presented in psychoactive drug advertisements, the symptoms mentioned most often were depression, diffuse anxiety, and tension. The male patients in the psychoactive drug advertisements, on the other hand, complained of specific symptoms which typically were associated with their jobs. The interpretation of the statistical analyses of the advertisements by sex and illness substantiated a significant relationship between the type of illness and sex of the patient. Other tendencies were the presentation of women patients as being irritating to significant others, and as

experiencing "socially embarrassing" symptoms. None of the advertisements had pictures of women as the physician, though all of the nurses pictured were women. Interestingly, the advertisements gave some indication that housewife roles can be bothersome, and consequently were reasonable problems to be treated by drugs.

Numerous allegations have been made that women's symptoms more often are linked with emotional than physiological factors than are men's symptoms (Physician's prejudice, 1973). Lennane and Lennane (1973) have given support to these allegations with quotes from medical texts. They addressed themselves to four women's medical problems which they believe are linked with psychogenic origins with insufficient scientific evidence. These four problems were: pain associated with menstruation, nausea during pregnancy, pain in child delivery, and infantile behavioral disorders. For example, numerous medical texts associated menstrual symptoms with nervous or neurotic women and avoidance of something disliked. Despite the finding that 75 to 88% of pregnant women experience nausea, one author blamed the nausea upon "resentment, ambivalence and inadequacy in women ill-prepared for motherhood." An editorial accompanying the Lennanes' article strongly recommended that physicians carefully evaluate themselves to ascertain when and if their sexual biases affected the

treatment they provided their patients (Kelly, 1973). These articles triggered numerous responses from physicians reacting to the likelihood of sexual prejudices having an impact upon health care (Women, pain, and prejudice, 1973). The illustrations of women's symptoms emanating from psychogenic factors have been widespread. One recent gynecologic article exemplified this distorted view of women. Easley (1971) viewed women as existing primarily to produce children and consequently suffering when childless. "Their specialty is built in anatomically, physiologically, and probably psychologically. . . . In many ways hysterectomy fits women's present needs. . . . A woman is a more reliable worker after she has had one."

Women in Medical Careers

Within the United States medical profession, the attitudes toward women have been reflected in the roles allocated to, or assumed by, women. Women comprise 70% of American health workers, although 93% of the physicians are men. In the past 40 years there has been a 300% increase in the number of women applying to medical school, but the percentage of their acceptance has dropped. During the same time span, applications to medical school have increased by 29%; but their percentage of acceptance has decreased (Norris, 1969).

Only three countries in the world, South Vietnam, Spain, and Madagascar, have fewer than our 7% women physicians (Boston Women's Health Book Collective, 1971). For example, a recent study reported that in England 24% and in Russia 75% of the physicians were women (Ehrenreich & English, 1973). Another report stated that 1% of dentists in the United States were women while in Europe the percentage ranged from 70% in Denmark to 24% in Sweden (Report of the Women's Action Program, 1972).

Throughout the health profession, men seem to have been granted higher status. Though not clearly documented, superficial evidence has suggested that male nurses, in most cases, have been treated with greater respect by most male physicians, regardless of ability, than have female nurses (Gilbert, 1970). Male dominance has even been detected in most medical school anatomy courses. Typically students have learned anatomy from a male sample except when they specifically are studying female reproductive functions. The reference points have customarily been male as in the instance of learning that the female skeleton is lighter and smaller than the male skeleton (Greer, 1970).

What positions have women health workers held? Women have comprised 98% of registered nurses, 96% of practical nurses, 95% of office workers, 94% of nutritionists and dieticians, 90% of medical technologists,

89% of medical social workers, 80% of physical therapists, 75% of x-ray technicians, 74% of aides and attendants, and 64% of cooks. Nearly all of dental hygienists and medical librarians have been women. Men, on the other hand, have been in the higher status positions, comprising 99% of dentists, 93% of physicians, 90% of chiropractors, and 80% of hospital administrators. A 1969 Department of Labor study revealed that women averaged 10% less pay in medical jobs than did men in identical jobs. Some dissatisfaction with pay and status has been reflected in the 60% turnover rate for nursing workers in 1967 (Reverby, 1972). The status of medical specialties has been determined primarily by income and esteem from medical colleagues. Within the medical specialties of high status, such as neurology and surgery, the percentage of women has been even less. Conversely, the areas of anesthesiology, psychiatry, and pediatrics, which have been lower status specialties, have a higher percentage of women (Frankfort, 1972). Efforts have been initiated more recently to improve conditions for women hospital workers by providing child care facilities within the hospitals (Barton, 1971).

The reluctance of the American Medical Association to allow paraprofessional medical workers has been noted as further indication of an unwillingness to allow women to assume responsible medical positions. Midwives,

typically women, were outlawed in this country early in this century, though in the United Kingdom, Scandinavia, and the Netherlands midwifery still thrives. In 1910 about 50% of babies in this country were delivered by midwives. However, rather than establishing training programs for the midwives, state after state outlawed midwifery largely because of pressure exerted by physicians, predominantly men, upon legislators (Ehrenreich & English, 1973).

A greater appreciation of the extent of subjugation by women by the American medical profession can be gained by a cross-cultural comparison. Post-revolutionary China has emphasized changing the status of women as in the marriage (and divorce) law of 1950 and the encouragement of late marriage. The specific health care policies regarding women in China demonstrate ways in which our system could be different. Forty days of training for midwives is provided covering such topics as sterilization of equipment, difficult deliveries, and insertion of I.U.D.'s. Women workers receive at least 56 days of paid maternity leave. Most factories have child and infant care centers. Women workers are given two 45-minute breaks for the purpose of visiting and nursing their babies. Women factory workers may take as many as three paid menstrual leave days per month. Medical research emphasizes common problems as

well as the most serious problems. For example, such widespread problems as cervical erosion, menstrual disorders, and vaginal infections are researched. Abortions are available freely and without consent from the husband. The vacuum aspirator method is the most widely used technique for abortions. Tubal ligations and vasectomies are available on demand, although the government prefers that people have had one or two children before such surgery. Birth control products are provided at no cost except for condoms which involve "waste" (Jervis, 1972).

In the past few years in the United States the number of women medical students has risen dramatically. In 1964 7.72%, in 1971 10.9%, and in 1974 15.4% of medical students were women (Shapley, 1974). This increase in enrollment has been cited as a consequence of recent efforts to equalize employment opportunities for men and women at all levels of expertise. For example, in 1970 Dr. Francis Norris presented documentation of discrimination against women applicants to medical school to a special education subcommittee of the House of Representatives (Discrimination against women, 1970).

Personality characteristics of women physicians have been the subject of numerous research projects. The upward shift in the percentage of women physicians

has been a potential catalyst for changing the stereotypical personality image of women physicians. "While many of the older women physicians took the ideal masculine behavior or the ideal sexlessness as givens, the younger physicians are no longer willing to accept these images" (Batt, 1972). Because discrimination against women was at a very high level when women became physicians five or more years ago, it is likely that they were forced to be "more conservative, more rigid, 'better' in every way than their male colleagues, or even renounce the mother-wife role altogether" (Boston Women's Health Book Collective, 1971).

Evidence gathered from several studies regarding the personalities of women physicians and medical students was conflicting. A study regarding the vocational interest patterns of men and women medical students over a four-year period revealed that in most ways the male and female students had similar interest patterns vocationally. The only variables of difference were that women students possessed a slightly higher "non-people" interest than did men students (Peterson, 1972). Cartwright (1972a, 1972b), on the other hand, compared personality characteristics of male and female medical students from 1960 to 1967 using adjective checklists and the California Psychological Inventory. On such characteristics as sense of well-being, dominance,

self-acceptance, and sociability the men and women were similar. The women, however, indicated that they were more willing than men to express affect, and admit weaknesses. The men indicated that they were more dominant and aggressive. The women expressed greater commitment to and contentment with their goals. Women also rated higher on tolerance, responsibility, achievement, and socialization. Women medical students valued their independence and individuality more than did their male counterparts (Cartwright, 1972b). The conscious factors which were cited by medical students as reasons for their selection of medicine as a career differed along gender lines. Women medical students mentioned altruism, support from other people, their interests, and self-development motives as reasons for choosing medicine. Men mentioned economic and prestige factors far more often than did the women (Cartwright, 1972a).

Summary

This study examined the role of medical students' and physicians' gender in responding to patient management problems. Did female and male medical students and physicians respond to patient simulations differently as a function of their own sex or the sex of the patient? The literature which was reviewed supported the notion that women as patients and women as health workers have been given inferior treatment to that accorded men.

The effect of this presumed inferior treatment needed to be tested. One means of testing this effect of sex upon responses to patient problems was through the use of a tight research design in which sex was the only variable. This study was addressed toward answering questions regarding whether or not the sex of patients and medical personnel influenced components of patient management including information gathering, diagnosis, treatment formulation, involvement of the patient in consultation arrangements, and the sequence of management actions.

CHAPTER II

DESIGN AND METHODOLOGY

Design

The design of this study was two by two across three patient management problems. The independent variables in the study were the sex of the medical personnel, the sex of the patients, and the type of patient problem. The study was conducted twice, once with men and women physicians and once with men and women medical students. The studies were analyzed separately, though the design and measures were the same for both studies.

Six measures, dependent variables, were used in this study as follows: the Overall Competence Index, the Information Gathering Competence Index, the Pathway Competence Index, the Management Competence Index, the Diagnosis Competence Index, and the percentage of consultations requested which included a discussion by the medical person with the patient about the consultation. These measures, with the exception of the Diagnosis Competence Index, were repeated across the three patient management problems. The Diagnosis Competence Index only

occurred in the third patient management problem. The Overall Competence Index was the major emphasis in this study, the other measures being included for additional exploration. The Competence Indexes measuring Information Gathering, Pathway, Management, and Diagnosis were components of the Overall Competence Index. The overall research design is diagrammed in Figure 1.

Sample

Two separate samples were used in the research, one a group of 40 physicians and the other a group of 40 medical students. The Michigan Medical Directory was used to obtain a random sample of 120 women physicians and 120 men physicians around the state. A cover letter from Dr. Norbert Enzer, the Chairman of the Michigan State University Department of Psychiatry, requesting participation in the study accompanied a set of either Form A or Form B of the three patient management problems along with a latent access pen. A follow-up letter from Dr. Enzer was sent to nonrespondents one month following the initial request. Copies of the two letters are in Appendix C. Twenty women and 20 men responded to these requests, a response rate of 16.67%. No data regarding these physicians were obtained.

The medical student sample was obtained by posting notices on the osteopathic and allopathic student

	Patient Management Problem I-Suicide						Patient Management Problem II-Asthma						Patient Management Problem III-Multiple Sclerosis					
	M ₁	M ₂	M ₃	M ₄	M ₅		M ₁	M ₂	M ₃	M ₄	M ₅		M ₁	M ₂	M ₃	M ₄	M ₅	M ₆
Female Physicians or Female Medical Students	Form A of Problems- Female Patients S ₁ S ₁₀																	
	Form B of Problems- Male Patients S ₁₁ S ₂₀																	
Male Physicians or Male Medical Students	Form A of Problems- Female Patients S ₂₁ S ₃₀																	
	Form B of Problems- Male Patients S ₃₁ S ₄₀																	

Code: M₁ = Overall Competence Index
M₂ = Pathway Competence Index
M₃ = Information Gathering Competence Index
M₄ = Management Competence Index
M₅ = Percentage of Consultations Including Discussion with Patient
M₆ = Diagnosis Competence Index

Fig. 1. Overall research design

bulletin boards at a midwestern university. The notices informed 2nd- and 3rd-year students that they would be paid \$5.00 for responding to three patient management problems. Feedback regarding their performance also was offered. Third-year students were not available. Two women in the second-year allopathic medical class responded to the notice along with 14 second-year women osteopathic medical students and 24 second-year men osteopathic medical students. No additional data regarding these students were obtained.

Experimental Procedures

The physicians responded to the patient management problems after receiving them in the mail. The directions were included with the problems. Half of the male and half of the female physicians were randomly sent Form A of the problems in which the patients were women. The other physicians were sent Form B of the problems on which the patients were men.

The medical students were told via bulletin board notice where they could obtain the patient management problems, directions, and latent access pen. They were given one week to complete the problems after they obtained them. Whether they received Form A or Form B of the problems was randomly determined. A secretary distributed the problems and paid the students when the problems were returned.

Instrumentation

The instrument used to assess responses to patient problems consisted of three patient management simulations. The simulations approximated actual patient problems as realistically as possible in that they required the respondent to make judgments and gather data as if he or she were in a clinical setting. The respondents were asked to assume that they were the physician described at the beginning of each problem. After reading a brief description of the presenting problem, the physician or medical student was asked to make a decision as to what he or she would like to do next. For example, the potential choices might have included conducting a physical exam on the patient, obtaining additional information about the patient's medical and personal history, running laboratory tests, or seeking consultation with someone else. Once the physician selected one of the possible decisions, he or she was directed to another section of the problem or provided with additional data. This information, either directions or data, was provided by the respondent rubbing with a special chemical pen the area in the answer booklet which corresponded to the number of the selection made. For example, if the respondent selected number three, he or she then rubbed the area adjacent to number three in the answer booklet. The

process, called latent access, entailed an invisible ink which was made visible by rubbing over it with a special chemical pen. In other words, the physician proceeded through each of the three patient problem simulations by making choices as to what he or she would like to do or obtain. As the respondents made their selections, they rubbed over the corresponding sections in the answer booklet and thus were provided information as to the answer to their question, the outcome of their interventions, and directions as to where they were to proceed in the instrument.

The format for this instrument was developed at the Office of Research in Medical Education at the University of Illinois. In 1958 the faculty of the University of Illinois College of Medicine requested that efforts be focused upon developing assessment techniques which would measure accurately the outcome of medical education in a manner consistent with the goals of medical education. The Committee on Examinations which coordinated the development of the assessment techniques attempted to emphasize the assessment of important rather than trivial skills, and integrated rather than fragmented skills. The outcome of their work in assessment of clinical judgment was the development of "patient management problems" (McGuire & Solomon, 1971).

Several considerations were involved in the selection and development of the problems for the instrument to be used in this study. First and foremost, every effort was made to resemble reality as closely as possible. The problems, for example, were not presented in classical textbook fashion, but rather included the usual bits of confounding data and problems encountered by physicians. Second, the problems represented a variety of types of problems. Patient problems can be organized according to a variety of schemes such as by degree of acuteness or chronicity, the causative agent, or the body systems affected (White, 1973). Third, the two forms of the instrument were designed to be as parallel as possible. For example, if the education of the female patient in Problem I on Form A was that of a B.A., the male patient on Form B also had a B.A. Finally, the three problems were selected because they included cultural and behavioral elements to suggest that the male and the female patient might be responded to in different ways. The patient management problems are not included in the dissertation because the problems continue to be used by medical personnel and mental health workers and their developers wish to keep them from public access.

The problems in the instrument included a case of multiple sclerosis in a 21-year-old patient, a

32-year-old asthmatic patient, and a 26-year-old patient who has been hospitalized for a suicide attempt.

The suicide problem was developed by the St. Elizabeth's Hospital Division of the National Center for Mental Health Services, Training and Research, National Institute of Mental Health in collaboration with the Center for Studies of Suicide Prevention, National Institute of Mental Health with funds provided under Public Health Service Grant Number MH 12557. The problem entailed the patient and the patient's spouse demanding that the patient be released from the hospital despite the fact that the patient's response to the six weeks of treatment had been equivocal.

The asthma problem was developed as part of a series of patient management problems by the Docent Corporation expressly for Hoffman - La Roche Incorporated. The problem entailed a medical emergency with the patient having status asthmaticus. Swift diagnosis and treatment were required.

The multiple sclerosis problem was not in the patient management format and thus was changed to this style by the author with assistance from Dr. Lionel Rosen. Originally the problem was used in the Medical Inquiry Project at Michigan State University (Elstein, Kagan, Shulman, Jason, & Loupe, 1972). The Medical Inquiry Project was for the purpose of investigating the way

in which expert physicians think and make decisions when they see patients. Actors and actresses were programmed to simulate actual patients. One of these simulations was the multiple sclerosis patient whose presenting complaint was paralysis below the waist.

Validity

To insure having patient cases which were identical in every way except gender, written patient simulations were used in this study rather than live patients. However, one may question whether or not the written simulations measured how a medical student or physician actually would perform in a real clinical setting with a live patient. In other words, were the written simulations valid measures of actual clinical performance? Validity can be defined as "an estimate of the extent to which a test (or, more properly, the scores on a test and the inferences made therefrom) measures the variable(s) it is designed to, or purports to, measure" (McGuire, Solomon, & Bashook, 1972).

The validity of patient management problems, in general, cannot be estimated since they are only an examination technique. One must consider validity in reference to something else. For example, one could question whether patient management problems are a valid measure of how a physician would respond to such patients in actual clinical practice. The validity of patient

management problems is affected by many variables. For example, the validity of a problem would be increased by careful test construction. Careful test construction would entail such things as only selecting items for inclusion which measured what the test maker wanted to measure. In this case, the test would include only items which measured how a physician or medical student would work with a patient in clinical practice. Inaccuracy in test scoring would jeopardize the validity of a test. How could a score on a test be a valid measure if it were not obtained with accuracy? The skills assessed by the test must be important. In summary, then, the validity of problems varies depending upon such variables as quality of test construction, the accuracy of test scoring, the importance of the skills assessed by the instrument, the degree to which the questions are representative of the area being assessed, and how the scores on the test relate to some other criterion such as actual clinical practice. For this research, it was especially important that the patient management instruments measure how the physicians and medical students would perform in actual clinical practice.

Although validation data were not collected on the three patient management problems included in this instrument, there have been studies regarding other

medical patient management problems constructed and scored in the same manner as the three used in this study. With regard to content validity, the content of the problems, that is, information gathering, management, and diagnostic skills, has been perceived by most physicians to represent important skills for medical students as well as practicing physicians to have.

Respondents' reactions to the patient management format of examinations have been very positive. For example, when students were polled regarding different types of questions on the American Board of Orthopaedic Surgery exams, both the students who failed and the students who passed the exams reacted more favorably to the patient management questions than the multiple choice questions. They evaluated the patient management problems as relevant and good ways to allow them to demonstrate their abilities. Furthermore, most physicians have responded positively to the content validity of the patient management problems, stating that they approximate reality and the "thinking which one must go through in making decisions about the workup and management of patients who present themselves in the office, clinic, or hospital setting" (McGuire, Solomon, & Bashook, 1972). With regard to the face validity of the instrument used in this research, caution was employed in the construction of the patient management problems.

The concurrent validity of various medical patient management problems has been studied by comparing performance on the tests with supervisors' ratings, with performance on other types of written or oral tests, and with performance in clinical settings with real patients having problems similar to the written simulated patients. For this research, it was most important that the patient management problems have high concurrent validity in order for the results of the research to be generalizable to actual clinical practice. Studies in which performance on patient management problems was compared to actual clinical practice suggest that the performances are "almost identical for the relevant specific items of diagnostic workup that one can plausibly assume would most likely have been recorded." One difference which was not consistent, however, was that fewer bits of data were recorded on actual patients' charts than were requested on patient management problems. This could probably be accounted for by incompleteness in recording information on hospital charts (McGuire, Solomon, & Bashook, 1972). Correlations of supervisors' ratings with performance on patient management problems ranged from .20 downwards. These low correlations, while statistically significant, were not meaningful. It is difficult to interpret these because correlations of supervisor's ratings with other supervisors' ratings

ranged from .30 to .70. Of course, it also must be recognized that supervisory ratings are global evaluations, whereas patient management problems assess primarily clinical judgment. The correlations between multiple choice test scores and patient management problem scores testing similar content areas ranged from .20 to .40. These findings can be attributable to the two types of tests being addressed to slightly different components of competency (McGuire, Solomon, & Bashook, 1972).

In summary, although the validity of the three problems used in this study is not known, the validities of other medical patient management problems have been acceptable. The educational importance of the areas assessed by patient management problems has been judged to be high. Similarly, the content or face validity was evaluated by both test takers and medical experts as being high for most medical patient management problems. The concurrent validity of medical patient management problems, while statistically significant, has not been found to be very high. The concurrent measures which have been compared with scores on patient management problems include scores on other types of tests, supervisory ratings, and actual clinical performance. Unfortunately, it is difficult to say which of these assessments are the most valid measure of a

medical person's performance. An analogy regarding spelling might further elucidate this confusing issue. If a person correctly spelled a word on a spelling test, but later misspelled the word when writing a letter, who is to say which is the more valid assessment of whether or not the person knows how to spell the word accurately? An exception to this, of course, occurs when one is interested in knowing only if the respondent has the capability to respond correctly, in which case the spelling test would be considered the valid measure. The person did have the capacity to spell the word correctly. Finally, the construct validity of medical patient management problems has been substantiated by factor analytic studies of a group's performance on written simulations and on other measures. That is, the components of information gathering, management, pathway, and diagnosis have been substantiated by factor analytic studies. These components emerged as separate factors within patients management problems. Additionally, different groups' performances, such as expert physicians and medical trainees, on written simulations have yielded results which suggest that medical patient management problems possess reasonable construct validity. That is, it appeared that performances on medical patient management problems did correspond to plausible hypotheses about how various groups of people should perform (McGuire, Solomon, & Bashook, 1972).

Reliability

Reliability refers to the consistency of a measuring instrument. Several kinds of reliability were relevant to consider with regard to the three patient management problems used in this study. The scoring keys for each of the three problems were developed by different experts. One group of three experts developed the scoring key for Problem I and three other experts developed the scoring key for Problem II. Two additional experts developed the scoring key for Problem III. The experts weighted the importance or value of each item on a given problem. Thus, the degree to which the experts agreed on the value of each item was one type of reliability evaluated for this study. This reliability, interrater reliability, was calculated for each of the problems and its subparts. When experts had significant disagreement on an item, the item was eliminated from the scoring of the problem. "Significant disagreement" was operationally defined as a range of +2 to -2 ratings by the experts on a given item. One item on Problem I (item 71) was eliminated because of significant disagreement. Nine items on Problem II (numbers 97, 98, 102, 115, 116, 120, 122, 123, and 124) were eliminated because of significant disagreement among the experts. The experts on Problem III did not have significant

disagreement on any items. Reliability coefficients using a Pearson product moment correlation coefficient were calculated for the problems with all the items included and for the problems with items eliminated on which there was significant disagreement. It should be clearly noted that the elimination of items elevated the reliability coefficients. These reliability coefficients are presented in Table 1 and Table 2. It is difficult to understand why most of the reliability coefficients for Problem III were higher than Problems I and II. However, according to physicians with whom the author spoke, more consensus exists on how to treat multiple sclerosis patients than about how to treat asthma patients. The management of suicidal patients, according to physicians, is more clear-cut than the management of asthma patients and less clear-cut than the management of multiple sclerosis patients.

Another type of reliability relevant to this study was test-retest reliability. The 40 students who initially responded to the patient management problems were contacted five weeks later and asked to respond to them again. Twenty-seven students completed Problems II and III and 26 students completed Problem I a second time. Pearson product moment correlation coefficients were calculated for Problems I, II, and III and for a

Table 1

Interrater Reliability Coefficients (With All Items Included)

Measures	Raters	Problem I	Problem II	Problem III
Overall Competence Index	Raters A X B Raters B X C Raters A X C	.56 .75 .75	.40 .47 .62	.80 X X
Pathway Competence Index	Raters A X B Raters B X C Raters A X C	.78 .83 .61	.87 .88 .73	.25 X X
Information Gathering Competence Index	Raters A X B Raters B X C Raters A X C	.62 .63 .79	.53 .69 .48	.66 X X
Management Competence Index	Raters A X B Raters B X C Raters A X C	.76 .89 .70	.22 .23 .76	.84 X X
Diagnosis Competence Index	Raters A X B Raters B X C Raters A X C	X X X	X X X	.75 X X

Table 2
Interrater Reliability Coefficients (With Items Excluded on Which Experts
Had Significant Disagreement)

Measures	Raters	Problem I	Problem II	Problem III
Overall Competence Index	Raters A X B Raters B X C Raters A X C	.62 .74 .80	.83 .74 .74	.80 X X
Pathway Competence Index	Raters A X B Raters B X C Raters A X C	.78 .83 .61	.87 .88 .73	.25 X X
Information Gathering Competence Index	Raters A X B Raters B X C Raters A X C	.62 .63 .79	.83 .63 .71	.66 X X
Management Competence Index	Raters A X B Raters B X C Raters A X C	.90 .87 .86	.85 .85 .78	.84 X X
Diagnosis Competence Index	Raters A X B Raters B X C Raters A X C	X X X	X X X	.75 X X

total of the three problems. These measures of test-retest reliability are presented in Table 3.

Table 3
Test-Retest Reliability Coefficients

	r
Problem I	.62
Problem II	.51
Problem III	.36
Total	.61

During the five weeks between administrations of the problems, the students were on vacation. Since they were not in classes during that time, the learning which took place came about primarily from independent reading. However, the researcher learned that the students discussed the problems with each other between administrations. The students indicated that they discussed Problem III more than the others because it involved diagnosis and they compared their diagnoses with one another. This may, in part, account for the low test-retest reliability coefficients of the three problems. Another possible contributor to the low test-retest reliability coefficients was the use of medical students rather than practicing physicians. Perhaps physicians would have been more definite in the way they managed patients, whereas the students were

more flexible in their management and thus tried different ways of responding the second time they performed the task. A third possible contributor to the low test-retest reliabilities was suggested by several of the student subjects. They said they wanted to try a different way of managing the patient "just to see what would happen." It seems reasonable to guess that the instrument which was used actually had higher test-retest reliability than is indicated with the medical student sample. However, if the reliability were as low as reported, these errors in measurement would reduce the precision of the instrument. That is, it would be more difficult for the instrument to detect differences if differences existed.

Coefficients of equivalence were not calculated for Forms A and B of the instrument since it was expected that they would not be equivalent because the sex of the patients was different.

Measures of internal consistency within the problems were not calculated because an assumption involved in such measures could not be met. Specifically, the number of respondents on each item varied. The sample size across items was not constant. In other words, all students did not respond to every item. The only indication regarding the internal consistency of patient management problems came from a study conducted by Donnelly (1972) who analyzed the performance on ten patient management problems by third-year medical students

at Wayne State University. He estimated the reliability of the problems by comparing each student's performance across all ten problems. The reliability of the problems was .84. In other words, students' scores on the ten problems correlated .84 with each other. The subscores of information gathering correlated .90 with each other, whereas the ten subscores of management correlated .35 with each other and the subscores of diagnostic ability correlated .25 with each other. The ten pathway subscores correlated .36 with each other. These data provided estimates of the internal consistency of that examination.

McGuire, Solomon, and Bashook (1972) have evaluated the reliability of various scoring schemes for the 1965 Clinical Comprehensive Examination administered to 200 third-year and 200 fourth-year medical students at the University of Illinois, College of Medicine. The exam was composed of 12 patient management problems from six medical disciplines. When different scoring schemes were used, the reliability was .97. The reliability coefficient obtained when different experts composed the criterion groups was .95. The correlation of the students' performance on this exam with their performance on an exam containing different problems was .56. The correlation between students' performance on the test with performance on problems containing similar problems which emphasized different skills and abilities was .86. The correlations between the scores on the test and

scores on a test containing identical problems with alternate options ranged from .67 to .94.

Scoring

The problems were handscored by a clerical assistant. Each item was scored between +2 and -2 except in the pathway and diagnosis sections on which the items were scored between +1 and -1. The weights for each item were determined by a group of experts. More specifically, each problem was given weights by two or three people having expertise in the particular cases. One member of the suicide expert group is a clinical psychologist who teaches a seminar on suicide and has been involved in suicide research. The second and third members of the criterion group, a psychiatrist and a psychologist, actively participated in a federally funded project to develop materials for a packaged training program on suicide for personnel in mental health and health care. The multiple sclerosis criterion group consisted of a psychiatrist who actively worked with this case in the Medical Inquiry project and one neurologist. The asthma criterion group consisted of two internists and a psychiatrist who does considerable liaison work with other physicians and thus is very familiar with psychophysiological illnesses.

The detailed directions which were given to each criterion group member are presented in Appendix A. The rater first was asked to review the entire problem. The rater scored the information gathering and management

sections of the problem according to the following scheme: +2 if the item was clearly indicated and important, +1 if the item was clearly indicated but of a more routine nature, 0 if the item had debatable or remote utility, -1 if the item clearly was not indicated though not harmful, and -2 if the item clearly was contraindicated. The pathway sections of the problems were rated by the criterion group members using a +1, 0, and -1 scheme. The item in each pathway section which the criterion group members would have chosen was rated a +1, and the other items were rated 0 if they were not harmful and -1 if they were clearly contraindicated. On Problem III there was a diagnostic section on which items were rated +1 if clearly correct, 0 if plausible but not quite correct, and -1 if clearly incorrect. Differences in weightings by criterion group members first were checked for inattention or carelessness on the part of the experts. The next thing which was checked was possible lack of clarity in defining the setting of the problem, the role to be assumed by the examinee, or the potential courses of action in the problem. If the discrepancy in the experts' ratings persisted, it was highly probable that the differences were ones which would exist always between experts. If on one item an expert felt that the option was never helpful and another expert felt it was always helpful

to do, the item was eliminated from the scoring process. Item 71 on Problem I was eliminated, and items 97, 98, 102, 115, 116, 120, 122, 123, and 124 on Problem II were eliminated for these reasons.

The process of scoring each exam entailed several steps. First, the Proficiency Index was calculated for the overall exam and for the pathway, management, diagnostic, and information-gathering sections by summing the positive and negative scores. Each sum was divided by the larger of either the total number of questions asked by the student or physician on that section or the total number of questions which were positively weighted by the expert group and then multiplied by 100 (Donnelly, 1972). This Proficiency Index represents the respondent's percentage agreement with the experts. The second step was the calculation of the Efficiency Index which is a measure of how good, on the average, the questions were which were asked by the students and physicians. This score does not consider questions which should have been asked. The Efficiency Index was calculated for the overall problem and for the information gathering, pathway, management, and diagnostic sections by dividing the total number of positively weighted items on each section by the total number of possible choices on each section and then multiplied by 100 (McGuire, Solomon, & Bashook, 1972). The Competence Index, the

measure being used in this study, is a combination of the Efficiency Index and Proficiency Index with more emphasis placed upon the Proficiency Index. It was calculated for the overall problem and for the management, diagnostic, information-gathering, and pathway sections by multiplying the Proficiency Index by the Efficiency Index and dividing the product by 100, adding the Proficiency Index to this figure, and then dividing the sum by two (McGuire, Solomon, & Bashook, 1972). The formulas for the three calculations are presented in Figure 2. The scoring key based upon the mean of the experts' weightings for each item is presented in Appendix C.

Analysis of the Data

The research hypotheses in this study were tested on two different samples. One sample consisted of medical students and the other sample consisted of practicing physicians. The data from these two samples were analyzed separately. The analyses for both samples were comparable.

The major focus of this study was finding out whether or not the gender of the medical personnel or patients affected the management of the patient. This basic question was best reflected in the Overall Competence Index measure. This measure indicated the quality of how the patients were treated. This measure

Proficiency Index:	$\frac{\Sigma [(+) + (-)]}{\text{Larger of either number of questions asked by students or physicians, or number of questions weighted positively by the criterion group}}$	x 100
Efficiency Index:	$\frac{\text{Number of (+) choices}}{\text{Number of all choices}}$	x 100
Competence Index:	$\frac{(\text{PI} \times \text{EI}/100) + \text{PI}}{2}$	

Fig. 2. Formulas for Proficiency, Efficiency, and Competence Indexes.

was calculated for each of the three problems. That is, an Overall Competence Index was computed for the suicide problem, the asthma problem, and the multiple sclerosis problem. Since this measure was repeated, the data analysis involved a repeated measures design. It was expected that a medical person's performance on one problem would correlate with the performance on the other two problems. In other words, the three measures of the Overall Competence Index were expected to be dependent upon each other. The expectation that measures would be dependent upon each other suggested that a multivariate analysis be used. Therefore, the data were analyzed using a multivariate approach to repeated measures. A diagram of this major focus of the research is presented in Figure 3. It should be noted that Figure 3 is exactly like Figure 1 except that it includes only Measure 1, the Overall Competence Index. The alpha level for the multivariate approach to repeated measures was set at .05 for a two-tailed test. The more analyses one conducts the greater the chance that one of the analyses will be statistically significant. In order to control for this chance factor in this research since seven hypotheses were tested, the alpha level for each of these hypotheses was set at .05 divided by seven, or .007.

		Patient Management Problem I-Suicide	Patient Management Problem II-Asthma	Patient Management Problem III-Multiple Sclerosis
		M ₁	M ₂	M ₃
Female Physicians or Female Medical Students	Form A of Problems- Female Patients S ₁₀			
	Form B of Problems- Male Patients S ₂₀			
Male Physicians or Male Medical Students	Form A of Problems- Female Patients S ₃₀			
	Form B of Problems- Male Patients S ₄₀			

Code: M1 = Overall Competence Index
M2 = Pathway Competence Index
M3 = Information Gathering Competence Index

Fig. 3. Main focus of the research: Measure 1

As stated earlier, the Overall Competence Index was broken down into parts to provide more detailed information. These parts were the Pathway Competence Index, the Information Gathering Competence Index, the Management Competence Index, the Diagnosis Competence Index, and the percentage of consultations requested which included a discussion about it with the patient. These five measures were expected to be dependent upon each other so this phase of the data analysis also required a multivariate approach. Four of these five measures were repeated across the three patient management problems and thus also necessitated a repeated measures analysis. One of these measures, the Diagnosis Competence Index, occurred only in the multiple sclerosis problem. Consequently, it did not require a repeated measures analysis. In summary then, a multivariate approach to repeated measures analysis was used to analyze the data obtained from the Pathway Competence Index, the Information Gathering Competence Index, the Management Competence Index, and the requested consultations which involved a discussion with the patient. This additional exploration of the research including measures 2, 3, 4, and 5 is diagrammed in Figure 4. Because the Diagnosis Competence Index was measured only once, it had to be analyzed separately. A two-way

	Patient Management Problem I-Suicide					Patient Management Problem II-Asthma					Patient Management Problem III-Multiple Sclerosis				
	M ₂	M ₃	M ₄	M ₅		M ₂	M ₃	M ₄	M ₅		M ₂	M ₃	M ₄	M ₅	
Female Physicians OR Female Medical Students	Form A of Problems- Female Patients S ₁₀														
	Form B of Problems- Male Patients S ₂₀														
Male Physicians OR Male Medical Students	Form A of Problems- Female Patients S ₃₀														
	Form B of Problems- Male Patients S ₄₀														

Code: M₂ = Pathway Competence Index
M₃ = Information Gathering Competence Index
M₄ = Management Competence Index
M₅ = Percentage of Consultations Including Discussion with Patient

Fig. 4. Additional exploration of the research: Measures 2, 3, 4, 5

univariate analysis of variance was used for this analysis. This phase of the research on Measure 6 is diagrammed in Figure 5.

A conversion of all raw scores to z scores was necessary because the three patient management problems had different numbers of possible points. Since the rating scales were therefore different, z transformations were done so that the means and standard deviations of the three problems would be equal.

A problem occurred when conducting the analysis of the data which required a slight change in the intended analysis. The problem occurred within the physician sample but not the medical student sample. More specifically, Measure 5, the percentage of consultations which included a discussion with the patient, did not measure anything in the asthma problem, Problem II, for the physicians. In other words, physicians when requesting consultations for the asthma patient, never mentioned discussing the consultation with the patient. This measure therefore was zero for Problem II. The computer program which was used in conducting the data analysis, the Finn program, could not handle zeros in the data. It then became necessary to eliminate Measure 5 for the doctor sample and analyze it separately. Thus, for the doctor sample Measures 2, 3, and 4 were analyzed and then Measure 5 was analyzed. When Measure 5 was

Patient Management Problem III-Multiple Sclerosis		M ₆	
Female Physicians or Female Medical Students	Form A of Problems- Female Patients S ₁₀	Form A of Problems- Female Patients S ₁₀	
	Form B of Problems- Male Patients S ₂₀	Form B of Problems- Male Patients S ₂₀	
Male Physicians or Male Medical Students	Form A of Problems- Female Patients S ₃₀	Form A of Problems- Female Patients S ₃₀	
	Form B of Problems- Male Patients S ₄₀	Form B of Problems- Male Patients S ₄₀	

Code: M6 = Diagnosis Competence Index

Fig. 5. Additional exploration of the research: Measure 6

analyzed for doctors, it was repeated in Problems I and III but not Problem II. A multivariate approach to repeated measures remained the type of data analysis which was used. The analysis of Measure 5 for physicians is presented in Figure 6.

Hypotheses

Seven hypotheses were tested in this study. These hypotheses were tested on a sample of physicians and also a sample of medical students. The hypotheses which were tested, as stated in the null form, are as follows:

Ho₁:

There are no significant differences between the responses of male and female medical students/physicians to patient management problems.

Ho₂:

There are no significant differences between the responses of medical students/physicians to male patients and to female patients on patient management problems.

Ho₃:

There is no significant interaction between the gender of patients and the gender of medical students/physicians on responses to patient management problems.

		Patient Management Problem I-Suicide	Patient Management Problem II-Asthma
		M ₅	M ₅
Female Physicians Or Female Medical Students	Form A of Problems- Female Patients S ₁₀		
	Form B of Problems- Male Patients S ₂₀		
Male Physicians Or Male Medical Students	Form A of Problems- Female Patients S ₃₀		
	Form B of Problems- Male Patients S ₄₀		

Code: M5 = Percentage of Consultations Including Discussion with Patient

Fig. 6. Additional exploration of the research for physicians: Measure 5

Ho₄:

There are no significant differences between the responses of medical students/physicians to three different patient management problems (suicide, asthma, and multiple sclerosis).

Ho₅:

There is no significant interaction between the gender of the medical students/physicians and three different patient management problems (suicide, asthma, and multiple sclerosis) on responses to patient management problems.

Ho₆:

There is no significant interaction between the gender of the patients and three different patient management problems (suicide, asthma, and multiple sclerosis) on responses to patient management problems.

Ho₇:

There is no significant interaction between the gender of the patients, the gender of the medical students/physicians, and three different patient management problems (suicide, asthma, and multiple sclerosis) on responses to patient management problems.

Summary

This study assessed the effect of gender upon management of medical patients in written simulated problems by physicians and medical students. The effect of the gender of both the patients and the medical personnel was examined. The study with physician subjects was analyzed separately from the study with medical

student subjects. The subjects, all volunteers, responded to three patient management problems pertaining to suicide, asthma, and multiple sclerosis. The Competence Index was calculated for their overall performance, management of the patients, diagnosis of the patients, information gathering regarding the patients, and pathway of management actions. A sixth measure which was examined was the percentage of consultations requested which included a discussion with the patient about the consultation. The Competence Index for their overall performance was the major focus of the study. The other five measures provided additional information. A multivariate approach to repeated measures was used to analyze the data.

CHAPTER III

FINDINGS AND DISCUSSION

No statistically significant differences were found in the analysis of the research data. Differences of statistical significance were not found for either the physician sample or the medical student sample on any of the variables being studied. The variables studied were the gender of the medical students or physicians, the gender of the patients in the patient management problems, the type of problem in each of the three patient management problems, and the four possible combinations of these three variables. The findings for the physicians and the medical students will be presented in this chapter. The mean scores of the transformed data for the physician subjects are presented in Table 4. The mean scores of the transformed data for the medical student subjects are presented in Table 5.

Alpha Level

The overall alpha level for the study was set at .05. First, the main focus of the study for both medical

Table 4
Means of Physicians' Transformed Data

	Problem I Suicide					Problem II Asthma					Problem III Multiple Sclerosis					Total						
	M ₁	M ₂	M ₃	M ₄	M ₅	M ₁	M ₂	M ₃	M ₄	M ₅	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₁	M ₂	M ₃	M ₄	M ₅	
Physicians	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Male physicians	.8158	.1525	.1527	-.1605	.1054	-.7445	.1520	-.9945	-.7760	.0000	-.9718	.6939	-.1151	.1227	1.0797	.1044	-.6410	.3991	.2857	-.1211	.1144	
Male patients with male physicians	.1207	.7979	.1696	-.3062	-.3689	-.1470	.2530	-.0021	-.2614	.0000	-.0329	-.0049	-.0400	.0748	-.2697	.2262	.0462	1.3584	.1316	-.6746	.0519	
Female patients with male physicians	.0424	-.4928	.1359	-.0147	.1581	-.0019	.0510	-.1968	.2458	.0000	-.1615	.1437	-.1901	.1706	.1618	-.0174	-.0475	-.5601	-.1245	.4324	.1768	
Female physicians	-.8158	-.1526	-.1527	.1605	-.1054	.7442	-.1520	.9947	.7783	.0000	.9724	-.6930	.1151	-.1227	-1.0782	-.1044	.8266	-.3995	-.2860	.1207	-.1146	
Male patients with female physicians	-.1692	-.0684	-.3089	.2425	.2635	.4819	-.0257	.2745	.2605	.0000	.1552	.0152	.1408	.0831	-.0539	-.1566	.1166	-.1128	-.0557	.7806	.2155	
Female patients with female physicians	.0061	-.2367	.0034	.0785	-.0527	-.2700	-.2783	-.0755	-.2449	.0000	.0392	-.1538	.0893	-.3285	.1618	-.0522	-.1149	-.6862	-.0014	-.5391	-.4477	
Male patients	-.2425	.3648	-.6963	-.3187	.0527	.1359	.1137	.1362	-.4428	.0000	.6117	.5150	.5042	.7896	3.2376	.3481	.8141	.6228	.6294	.5299	.1352	
Female patients	.2425	-.3648	.6965	.3189	-.0527	-.1360	-.1137	-.1362	.4461	.0000	-.6112	-.5060	-.5044	-.7894	-3.2361	-.3481	-.8122	-.6232	-.6296	-.5334	-.1354	

Code: M1 = Overall Competence Index
M2 = Pathway Competence Index
M3 = Information Gathering Competence Index
M4 = Management Competence Index
M5 = Percentage of Consultations Including Discussion with Patient
M6 = Diagnosis Competence Index

Table 5
Means of Medical Students' Transformed Data

	Problem I Suicide					Problem II Asthma					Problem III Multiple Sclerosis						Total					
	M ₁	M ₂	M ₃	M ₄	M ₅	M ₁	M ₂	M ₃	M ₄	M ₅	M ₁	M ₂	M ₃	M ₄	M ₅	M ₆	M ₁	M ₂	M ₃	M ₄	M ₅	
Medical Students	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Male medical students	-.4562	.9683	-.2694	-.1383	.1590	-.6581	-.1324	-.1990	.7936	.0167	-.5917	.4404	-.9941	.2484	.0229	.1237	-.9899	.1712	-.1420	-.0478	.1235	.1235
Male patients with male medical students	.3809	.1731	.3807	.1090	.8566	.2468	.1284	.0304	.3681	.0329	-.0785	-.0106	-.1578	-.1477	.2565	.2494	.4919	.6936	.3750	.1715	.7161	.7161
Female patients with female medical students	-.5497	-.1834	-.5087	-.4743	-.6656	-.4353	-.1806	-.4701	-.2619	.0025	-.0364	.0222	-.0304	.2387	-.2532	-.2678	-.7973	-.7824	-.7529	-.3070	-.5768	-.5768
Female medical students	.6830	-.1441	.4065	.2376	-.2384	.9947	.2362	.2990	-.1189	-.0256	.8925	-.8511	.1497	-.3761	-.0344	-.3332	.1489	-.2606	.2136	.0718	-.1853	-.1853
Male patients with female medical students	.0491	.1170	.0330	.1085	-.4488	.2225	-.1761	.6221	-.0969	-.2789	-.4835	-.2835	-.3576	-.5315	-.0344	-.3460	-.0148	.0750	.1471	-.2309	-.4168	-.4168
Female patients with female medical students	.0875	-.1459	.0483	.3668	-.0280	-.0235	.2233	-.0241	-.1408	.2278	.6621	.2665	.6570	.4563	-.0344	-.3204	.3125	-.1271	.2801	.3744	.0463	.0463
Male patients	.2545	.1517	.2482	.1088	.3593	.2375	.1239	.2558	.1910	-.0859	-.2328	-.1146	-.2339	-.2939	.1456	.2261	.2989	.4579	.2882	.8541	.2845	.2845
Female patients	-.2814	-.1676	-.2742	-.1201	-.3972	-.2619	-.1053	-.2823	-.2110	.0944	.2577	.1251	.2590	.3245	-.1611	-.2899	-.3300	-.5065	-.3180	-.7705	-.3145	-.3145

Code: M1 = Overall Competence Index
M2 = Pathway Competence Index
M3 = Information Gathering Competence Index
M4 = Management Competence Index
M5 = Percentage of Consultations Including Discussion with Patient
M6 = Diagnosis Competence Index

students and physicians was upon Measure 1, the Overall Competence Index. Seven hypotheses were tested on Measure 1 and the alpha level for each analysis was set at .007. As mentioned in Chapter II, this reduction in alpha level was necessary because seven analyses were made. The figure .007 was obtained when .05 was divided by seven. Secondly, the additional exploration consisted of several parts. For the physician subjects, Measures 2, 3, and 4 were analyzed, and then Measure 5 was analyzed, and finally Measure 6 was analyzed. Measures 2, 3, and 4 and Measure 5 were tested for seven hypotheses each time and Measure 6 was tested for three hypotheses. Thus, the additional exploration studies for physicians entailed 17 analyses. Therefore, the alpha level for these additional exploration tests was set at .05 divided by 17 or .003. The additional exploration for the medical students entailed 10 analyses rather than 17 because Measure 5 was analyzed with Measures 2, 3, and 4 for the medical students. Consequently, the additional exploration alpha level of .05 for medical students was divided by 10, equalling .005.

Physicians

Measure 1

The data regarding physicians' responses to patient management problems were analyzed using a

multivariate approach to repeated measures. The data first were run to analyze whether or not there were statistically significant differences on the Overall Competence Index, Measure 1. The findings are presented in Table 6.

Table 6
Results for Physicians: Measure 1

Sources of Variation	Degrees of Freedom	F Ratio	Probability
D	1,36	.0872	.7695
P	1,36	.3213	.5744
C	2,35	.0000	1.0000
D X P	1,36	.2168	.6444
D X C	2,35	.3908	.6794
P X C	2,35	.2843	.7543
D X P X C	2,35	.8181	.4496

Code: D = Doctor Gender
P = Patient Gender
C = Condition, i.e., Problem I, Problem II, Problem III
Alpha = .007

The findings suggest that there were no significant differences on Measure 1 for any of the variables being tested. That is to say, none of the seven null hypotheses stated at the end of Chapter II were rejected. The gender of the physicians, the gender of the patients, which patient management problem it was, and the interaction of these things did not significantly affect the way in which the physicians performed on the Overall Competence Index, Measure 1. Based upon these findings,

one could state that men and women physicians responded to men and women suicide, asthma, and multiple sclerosis patients with comparable quality of care.

Measures 2, 3, and 4

Measure 1 was broken into five parts, noted as Measures 2, 3, 4, 5, and 6. Measures 2, 3, and 4 occurred in all three of the patient management problems so they were analyzed simultaneously using a multi-variate approach to repeated measures. The findings are presented in Table 7.

Table 7
Results for Physicians: Measures 2, 3, 4

Sources of Variation	Degrees of Freedom	F Ratio	Probability
D	3,34	.9096	.4466
P	3,34	1.4569	.2436
C	6,31	.0000	1.0000
D X P	3,34	1.1816	.3313
D X C	6,31	.5786	.7000
P X C	6,31	.8727	.5260
D X P X C	6,31	.6651	.6783

Code: D = Doctor Gender
P = Patient Gender
C = Condition, i.e., Problem I, Problem II, Problem III
Alpha = .003

The results suggest that physicians obtained similar results on Measures 2, 3, and 4 regardless of their gender, the gender of the patients, and which patient management problem it was. That is, there were no statistically significant differences found on the Pathway Competence Index, the Information Gathering Competence Index, or the Management Competence Index for men and women physicians, or for men and women patients, or for the suicide, asthma, and multiple sclerosis problems. The seven null hypotheses stated in Chapter II were not rejected.

Measure 5

Measure 5, the percentage of consultations requested which included a discussion with the patient, was repeated in the suicide problem and the multiple sclerosis problem. These data were analyzed using a multivariate approach to repeated measures. The results are presented in Table 8.

The results indicate that there were no statistically significant differences for Measure 5 on any of the variables tested. The seven null hypotheses stated at the end of Chapter II failed to be rejected. The gender of the physician, the gender of the patient, and the type of problem did not significantly affect the physicians' performance on Measure 5.

Table 8
Results for Physicians: Measure 5

Source of Variation	Degrees of Freedom	F Ratio	Probability
D	1,36	.4687	.4980
P	1,36	.8496	.3629
C	1,36	.0000	1.0000
D X P	1,36	1.2938	.2629
D X C	1,36	.0539	.8178
P X C	1,36	.2425	.6255
D X P X C	1,36	.5009	.4837

Code: D = Doctor Gender
P = Patient Gender
C = Condition, i.e., Problem I, Problem II,
Problem III
Alpha = .003

Measure 6

Measure 6, the Diagnosis Competence Index, occurred only for the multiple sclerosis problem. The data on this measure were analyzed using a two-way univariate analysis of variance. The findings are presented in Table 9.

The results suggest that there were no statistically significant differences for Measure 6 with regard to the gender of the physicians, gender of the patients, or interaction between the physicians' gender and patients' gender. The percentage of consultations which the doctors requested which included a discussion with the patient was not significantly affected by the sex of

the doctors or patients. The seven null hypotheses stated at the end of Chapter II failed to be rejected.

Table 9
Results for Physicians: Measure 6

Sources of Variation	Degrees of Freedom	F Ratio	Probability
D	1,36	.9998	.3241
P	1,36	.9997	.3241
D X P	1,36	1.0004	.3239

Code: D = Doctor Gender
P = Patient Gender
Alpha = .003

Medical Students

Measure 1

As noted previously, Measure 1, the Overall Competence Index, was the major focus of this study. The hypotheses which were tested using the Overall Competence Index were accepted as stated in the null form. That is, no significant differences or interactions were found on the medical students' overall performance on the three patient management problems with regard to the gender of the patients, students, or which problem it was. These results were obtained using a multivariate approach to repeated measures. These statistically insignificant results are presented in Table 10. Although statistically insignificant,

special note should be made of the low probability (.0477) of chance accounting for the interaction of the doctor and patient gender and the .0381 probability of chance accounting for the interaction of the patient gender with the different problems.

Table 10

Results for Medical Students: Measure 1

Sources of Variation	Degrees of Freedom	F Ratio	Probability
P	1,36	.4902	.4884
P	1,36	.8815	.3541
C	2,35	.0000	1.0000
D X P	1,36	4.2057	.0477
D X C	2,35	.0081	.9920
P X C	2,35	3.5943	.0381
D X P X C	2,35	.4615	.6342

Code: D = Medical Student Gender
P = Patient Gender
C = Condition, i.e., Problem I, Problem II,
Problem III
Alpha = .007

Measures 2, 3, 4, 5

The findings regarding Measures 2, 3, 4, and 5 for medical student subjects are presented in Table 11. No statistically significant differences were found with this multivariate approach to repeated measures. The gender of the patients, the gender of the medical students, and the type of patient management problem did not affect the performance of the medical students on

the Pathway Competence Index, Information Gathering Competence Index, the Management Competence Index, and the percentage of consultations requested which included a discussion with the patients. As in Measure 1, special note should be made of the low probabilities of chance accounting for the interaction between doctor gender and patient gender and the interaction between the patient gender and the type of patient problem.

Table 11

Results for Medical Students: Measures 2, 3, 4, 5

Sources of Variation	Degrees of Freedom	F Ratio	Probability
D	4,33	.7516	.5642
P	4,33	1.0000	.4215
C	8,29	.2654	.9723
D X P	4,33	2.4760	.0633
D X C	8,29	.6805	.7048
P X C	8,29	2.5611	.0303
D X P X C	8,29	1.0215	.4424

Code: D = Medical Student Gender
P = Patient Gender
C = Condition, i.e., Problem I, Problem II,
Problem III
Alpha = .005

Measure 6

The analysis of Measure 6, the Diagnosis Competence Index, which occurred only for the multiple sclerosis problem, did not detect statistically significant differences between men and women medical students or between men and women patients. Neither

the gender of the medical student nor the gender of the patient significantly influenced the accuracy of the diagnosis for the multiple sclerosis patient. The male patient and the female patient with multiple sclerosis were diagnosed with about the same degree of accuracy regardless of the patient's gender or the medical student respondent's gender. These results are presented in Table 12.

Table 12

Results for Medical Students: Measure 6

Sources of Variation	Degrees of Freedom	F Ratio	Probability
D	1,36	3.9740	.0539
P	1,36	.2692	.6071
D X P	1,36	5.2761	.0276

Code: D = Medical Student Gender
P = Patient Gender
Alpha = .005

Summary

The physician and medical student data were separately analyzed using a multivariate approach to repeated measures. No statistically significant differences were found on any of the hypotheses. None of the null hypotheses stated at the end of Chapter II were rejected. No support was found for the assumption that either the gender of the physicians or medical students or the gender of the patients influence the

treatment of the patients in the problem simulations. Furthermore, the type of problem in the simulation did not significantly affect the treatment of the patients. That is, the suicide patient, the asthma patient, and the multiple sclerosis patient were not treated significantly better or worse as a function of their illness. Finally, no significant interactions were found between the gender of the medical person, the gender of the patient, and the type of problem the patient had.

CHAPTER IV

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS FOR FURTHER STUDY

Summary

The Problem

The purpose of the research was to assess the effect of gender upon management of medical patients by medical students and physicians. The effect of the gender of both the patients and the medical personnel upon responses to written simulations of patient management problems was examined. The instrument which was used included three types of patient management problems. Therefore, the type of problem was also examined to see if it affected the treatment of the patients.

The issue of major concern in this study was whether or not women and men patients received inappropriately different medical treatment as a function of their gender or the gender of the medical personnel. A review of the health care literature provided some suggestion that gender might affect the quality of care given patients. The literature which had implications

for the effect of gender upon health care focused around two central tenets. One tenet was that women have received inadequate health care. The other tenet was that women have been relegated to inferior status positions within the medical profession. Unfortunately, most of the literature pertinent to gender and health care was conjectural or demographic in nature. Experimental evidence of differential treatment for men and women was not available in either the literature regarding treatment of women as patients or women as medical personnel. However, there was sufficient anecdotal and demographic evidence to suggest that perhaps women receive unequal treatment by the medical profession when compared with treatment given to men. The existence of this presumed inferior treatment needed to be tested. One means of testing this effect of sex upon responses to patient problems was through the use of a tight research design in which sex was the only variable.

Design and Methodology

The research design was two by two across three patient management problems. That is, the effect of gender of the medical personnel, male or female, was assessed along with the effect of the gender of the patients treated, males or females, across three different types of patient problems. The independent

variables in the study were the sex of the medical personnel, the sex of the patients, and the type of patient problem. There were three types of patient problems, namely suicide, asthma, and multiple sclerosis. The study was conducted twice, once with men and women physicians and once with men and women medical students. The studies were analyzed separately, though the design and measures were the same for both studies.

There were six measures, dependent variables, used in this study as follows: the Overall Competence Index, the Information Gathering Index, the Pathway Competence Index, the Management Competence Index, the Diagnosis Competence Index, and the percentage of consultations requested which included a discussion by the medical person with the patient about the consultation. These measures, with the exception of the Diagnosis Competence Index, were repeated across the three patient management problems. The Diagnosis Competence Index only occurred in the third patient management problem. The Overall Competence Index was the major emphasis in this study, the other measures being included for additional exploration. The Competence Indexes were calculated using a series of three formulas which included measures of both the quality and efficiency of the respondents' treatment of the patient. Experts on suicide, asthma, and multiple sclerosis participated

in developing the scoring key by assigning positive or negative weights to the various options in each problem which determined levels of quality and efficiency.

The three patient management problems were divided into two parallel forms. On Form A the three patients were women and on Form B the three patients were men. Form A and Form B were randomly assigned to the subjects. The problems approximated actual clinical reality as closely as possible. They required the respondent to make judgments and gather data as if he or she were in a clinical setting. The respondents were asked to assume that they were the physician described at the beginning of each problem. After reading a brief description of the presenting problem, the physician or medical student was asked to make a decision as to what he or she would like to do next. For example, the potential choices might have included conducting a physical exam on the patient, obtaining additional information about the patient's medical and personal history, running laboratory tests, or seeking consultation with someone else. Once the physician selected one of the possible options, he or she was directed to another section of the problem or provided with additional data. This information, either directions or data, was obtained by the respondent

rubbing with a special chemical pen the area in the answer booklet which corresponded to the number of the selection made. The process, called latent access, entailed an invisible ink which was made visible by rubbing over it with the special chemical pen.

There were two separate samples used in the research, one a group of 40 physicians and the other a group of 40 medical students. A random sample of 120 women and 120 men physicians living in a midwestern state were asked to contribute approximately one hour of their time to respond to Form A or B of the problems. Twenty men and 20 women voluntarily completed the problems. The medical student sample was obtained by posting notices requesting response to three patient management problems on the medical student bulletin boards at a midwestern university. Forty second-year students, 16 women and 24 men, responded to the request and were paid \$5.00 each for their participation.

Results

The results of the study were that no statistically significant differences were found on any of the variables tested. In other words, no significant differences were found in the ways that men and women physicians or medical students treated the patients in the problems. Furthermore, the women patients on

Form A were not treated differently from the men patients on Form B. The suicide patient, the asthma patient, and the multiple sclerosis patient received qualitatively similar treatment. No significant differences were found in the quality of the treatment received by the three types of patients. Finally, no significant differences were found in the interaction between the patients' gender, the type of problem, or the gender of the physicians or medical students.

These results were obtained by using a multivariate approach to repeated measures. The one exception to this was that the Diagnosis Competence Index occurred in only one of the three problems and thus was not a repeated measure. Therefore, a univariate analysis of variance was used for the Diagnosis Competence Index. Before the data were analyzed, the scores of the three problems were transformed to z scores so that the problems could be compared with each other.

Discussion

The researcher intends this discussion of the research data to be speculative rather than rigorous. All other sections of the dissertation have been based upon statistical rigor. The data from the research from a statistical standpoint failed to disprove the null hypotheses which were stated at the end of Chapter II.

However, the data were fascinating to the researcher from an intuitive perspective.

The speculations are derived from material presented in Tables 13 and 14 which are listings of the rank order of quality of treatment for men treating male patients, men treating female patients, women treating female patients, and women treating male patients. The rankings go from the highest mean scores of these groups at the top to the lowest mean scores of these groups at the bottom. The mean scores of the groups are presented in Tables 4 and 5. In other words, if male medical students treated male patients the best according to a particular measure, that group was listed at the top, and if male medical students treated female patients the worst on that measure, that group was placed fourth, at the bottom, on the list. Table 13 is based upon the medical student sample and Table 14 is based upon the physician sample.

Medical Students

Let us first examine the medical students' data. For the 16 measures on the three problems, the top rankings were in 15 instances held by like-sexed groups. More specifically, on eight measures the best treatment was given by male medical students to male patients and on seven measures the best treatment was given by female medical students to female patients. On only one of the

Table 13

Medical Students: Rank Order by Groups

Problem I						Problem II						Problem III						Total					
Rank	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5	M6		M1	M2	M3	M4	M5	
Highest	Mm	Mm	Mm	Ff	Mm	Mm	Ff	Fm	Mm	Ff	Mm	Ff	Ff	Ff	Mm	Mm		Mm	Mm	Mm	Ff	Mm	
2nd	Ff	Fm	Ff	Mm	Ff	Fm	Mm	Mm	Fm	Mm	Mm	Mf	Mf	Mf	Fm*	Mf		Ff	Fm	Ff	Mm	Ff	
3rd	Fm	Ff	Fm	Fm	Fm	Ff	Fm	Ff	Ff	Mf	Mf	Mm	Mm	Mm	Ff*	Ff		Fm	Ff	Fm	Fm	Fm	
Lowest	Mf	Mf	Mf	Mf	Mf	Mf	Mf	Mf	Mf	Fm	Fm	Fm	Fm	Fm	Mf	Fm		Mf	Mf	Mf	Mf	Mf	

Code: M = Male Medical Students

m = Male Patients

F = Female Medical Students

f = Female Patients

M1 = Overall Competence Index

M2 = Pathway Competence Index

M3 = Information Gathering Competence Index

M4 = Management Competence Index

M5 = Percentage of Consultations Including Discussion with Patient

M6 = Diagnosis Competence Index

* These two mean scores were equal and tied with each other.

Table 14

Physicians: Rank Order by Groups

Problem I					Problem II					Problem III					Total						
Rank	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5	M1	M2	M3	M4	M5	M6	M1	M2	M3	M4	M5
Highest	Mm	Mm	Mm	Fm	Fm	Fm	Mm	Fm	Fm	Mf*	Fm	Mf	Fm	Mf	Mf*	Mm	Fm	Mm	Mm	Fm	Fm
2nd	Mf	Fm	Mf	Ff	Mf	Mf	Mf	Mm	Mf	Ff*	Ff	Fm	Ff	Fm	Ff	Mf	Mm	Fm	Ff	Mf	Mf
3rd	Ff	Ff	Ff	Mf	Ff	Mm	Fm	Ff	Ff	Fm	Mm	Mm	Mm	Mm	Fm	Ff	Mf	Mf	Fm	Ff	Mm
Lowest	Fm	Mf	Fm	Mm	Mm	Ff	Ff	Mf	Mm	Fm	Mf	Ff	Mf	Ff	Mm	Fm	Ff	Ff	Mf	Mm	Ff

Code: M = Male Medical Students

m = Male Patients

F = Female Medical Students

f = Female Patients

M1 = Overall Competence Index

M2 = Pathway Competence Index

M3 = Information Gathering Competence Index

M4 = Management Competence Index

M5 = Percentage of Consultations Including Discussion with Patient

M6 = Diagnosis Competence Index

* These two mean scores were equal and tied with each other.

16 measures, the Information Gathering Competence Index for Problem II, did a mixed-sex group receive the highest score. This evidence suggests that for medical students, there was a consistent interaction of the gender of the patient with the gender of the medical student.

This observation was further substantiated when the lowest rankings were examined. Which groups fared worst according to the 16 measures? Interestingly, on all 16 measures the lowest scores were obtained by mixed-sex groups. On 10 measures, the worst treatment was given by male medical students to female patients and on the other six measures, the worst treatment was given by female medical students to male patients. Again, the evidence suggests that some kind of interaction occurred between the gender of the patients and the gender of the medical students.

When examining the total scores for the medical students on the five repeated measures, the pattern of gender interaction with quality of treatment was again witnessed. All five of the top performances were by like-sexed groups, four of them men and one woman. The lowest scores were all obtained by mixed-sex groups, namely male medical students treating female patients.

Male medical students treated male patients better than female patients on 13 measures and female

patients better than male patients on three measures. Female medical students treated female patients better than male patients on 11 measures, male patients better than female patients on four measures, and equally on one measure. Again, it appeared that there was an interaction between the genders of patients and medical students because like-sexed groups were treated more favorably than mixed-sex groups.

Referring back to the statistical research results presented in Chapter III, it was noted in Tables 10, 11, and 12 that the probability of the results for the interaction between gender of the medical students and gender of the patients being what they were was quite low. While not statistically significant, it is interesting to observe on Table 10 that the probability of the interaction between the gender of the medical students and the gender of the patients was .0477 for Measure 1. On Table 11 it was noted that for Measures 2, 3, 4, and 5 the probability of the interaction between the gender of the patients and the gender of the medical students was .0633. On Table 12 it was noted that for Measure 6 the probability of the interaction of the gender of the medical student with the gender of the patient was .0276. These low probabilities suggest that perhaps there was some interaction between the gender of the patients and

the gender of the medical students, though not statistically significant. One additional observation on Tables 10 and 11 was interesting. The probability of the interaction between which problem it was and the gender of the patient was .0381 on Measure 1 and .0303 on Measures 2, 3, 4, and 5.

Let us for the moment drop the examination of the interaction of the medical student gender with the gender of the patients and move on to look only at the gender of the patients. Reference back to the bottom two rows of Table 5 was useful for exploring how the gender of the patients affected the scores on three problems. Problem I data on all five measures were interpreted as indicating that male patients received better treatment than female patients. Problem III data, further, were interpreted as suggesting that the female patient received better treatment on four of the measures and male patients received better treatment on two of the measures. When the scores on the three problems were combined for the Total scores, male patients received better treatment than the female patients on all five measures. The reasons for these differences could not be explained by differential incidence of the problems in men and women. Asthma and multiple sclerosis for adults occur about equally often in men and women. Whereas suicide is more often

committed by men, it is more often attempted by women (Freedman, Kaplan, & Sadock, 1972). Since there were no logical explanations for these differences, it was accepted that it was mere happenstance that on two problems men received better treatment and on one problem women received better treatment.

Physicians

Turning now to the physician data, more interesting observations were made. The listing of the rank order of quality of treatment for men treating male patients, men treating female patients, women treating female patients, and women treating male patients is reported in Table 14. For the 15 measures on the three problems, the top rankings were in 12 instances held by groups in which male patients were treated better than female patients. Five of these 12 instances were of male physicians treating male patients, and the other seven were of female physicians treating male patients. On the other three measures, male physicians treated female patients the best. When the scores were combined for Total Measures, on all five measures male patients were treated best, twice by male physicians and three times by female physicians.

On eight measures male physicians treated male patients better than female patients and on seven measures the male physicians treated female patients

better than the male patients. This suggested that male physicians treat men and women patients without being particularly influenced by the gender of the patient. The data for the female physicians were quite different. Women physicians tended to treat male patients better than female patients. Specifically, the female physicians treated male patients better than female patients on 11 of the measures, and female patients better than male patients on four of the measures.

Summarizing the physician data, it appeared that male physicians were not affected by the gender of the patients, whereas female physicians treated male patients more favorably than female patients.

One final observation of the physicians' responses pertained to the differential return of latent access pens. The researcher requested that the pens be returned. Eight of 20 male physicians returned the pens, whereas 17 of 20 female physicians returned the pens. Furthermore, it appeared that female physicians followed the directions in the problems more closely than did the male physicians. This observation was intuitive more than numerical.

Medical Students and Physicians

Examination of the medical student data in combination with the physician data was fascinating. A number of intriguing questions emerged from these findings. Why was it that female medical students gave superior treatment

to female patients while female physicians gave superior treatment to male patients? Why was it that male physicians treated men and women patients similarly but male medical students gave better treatment to male patients? Why did male patients receive better treatment overall by the physicians whereas gender did not affect the overall treatment by the medical students?

These findings and questions make sense to the researcher from an intuitive sense of the current upheaval of men's and women's roles in society. The Women's Liberation Movement may well account for some of the findings. How so? The researcher based the following speculations on the assumption that the physician sample was older than the medical student sample and thus probably less open to change in basic attitudes about women and men.

The fact that women physicians treated male patients better whereas women medical students treated female patients better can be attributed to younger women being more aware of their own sex and developing more positive feelings about others of their same sex. The women's movement has facilitated the increasing enrollment of women in medical school as discussed in the review of the literature earlier in the dissertation. Formerly, women who became physicians had few women physicians from whom to garner support. As discussed previously, women physicians may have had to "out-man the men." Just as the Black Movement increased Black pride, the Women's

Movement has increased women's pride. Women medical students have support from other women medical students. Perhaps the increased attention to gender brought about by the Women's Movement has also facilitated greater unity and pride among men. It also seems possible to hypothesize that male medical students gave preferential treatment to male patients because they feel some resentment toward women for upsetting the status quo of sex roles. Some of the male medical student subjects told the researcher that they have felt "defensive" because of women's increasing assertiveness which both the men and women attributed to the Women's Movement. Because the roles of women and men are in a state of upheaval, it may be that men are more cautious of women, fearing a potential loss of the status they have previously had.

Extra insight into why male physicians were less affected by the gender of the patients than the female physicians or the medical students might stem from their having held the majority of the power and thus not feeling threatened by women. Furthermore, in the past it has been shown that men are more favorable toward women than women are toward women. Another possibility is that physicians in practice, presumably older than the medical students, are more fixed in their sexual attitudes and less susceptible to the influence of the Women's Movement.

The same dynamics may also account for why male patients received better treatment overall by the physicians

whereas the gender of the patients did not affect the overall treatment by the medical students. Men are no longer as readily perceived as "better" than women. Thanks to the Women's Liberation Movement, women may be recognized more favorably than they formerly were. The Women's Liberation Movement may also have facilitated in allowing people to respond to other people with more spontaneity and comfort. It would seem natural for women to have a more empathic regard for women and men to have more empathy for men. This emerged in the results of the medical students' performance with regard to quality. Male medical students gave better treatment to patients of their same sex and the same held true for female medical students.

Implications

The implications of the research, if one looks only at whether or not statistically significant differences were found, are minimal. The results were such that none of the null hypotheses were rejected. These findings would suggest that the gender of physicians and patients has no effect upon the quality of treatment given to patients. Consequently, no special attention to attitudes about men and women would be necessary in either medical school or in continuing education programs for practicing physicians.

However, if one looked at the pattern of the results for the medical students and physician sample,

it seems less reasonable to dismiss the notion that sex attitudes affect the quality of medical care. Continuing education programs for practicing physicians might sponsor workshops for women physicians directed toward improving their self-esteem as women, and consequently their attitudes about other women. Similarly, men and women medical students could perhaps benefit from focusing attention on their attitudes about the opposite sex in an effort to reduce the likelihood that less adequate medical care be given to patients of the opposite sex.

Based upon the findings from the statistical analyses and the intuitive speculations, it seems that the major implication of this research is that further research should be conducted before final conclusions and recommendations are made. Suggestions for future research are discussed later in this chapter.

Limitations of the Study

The statistical findings of this research project suggest that men and women patients are not treated differently by men and women physicians and medical students. However, several things must be kept in mind when interpreting the results of this research. The way in which the research subjects were obtained, the size of the samples, the technique for assessing potential differences in treatment, and the number of variables

in the research must be considered. The precision of the study was lessened because of a small sample size and probable errors of measurement due to the instrument used.

First of all, the research subjects were volunteers. The physicians were volunteers from a random sample in the state of Michigan. Only 16% of the practitioners asked to participate in the study responded to the three patient management problems. Can one assume that this small percentage of volunteers responded the same as would nonvolunteers? Are physicians in Michigan the same as physicians elsewhere? The medical student volunteers all were from one university, and predominantly one medical college. Did these medical student volunteers respond the same as students who did not volunteer would have? Additionally, they all were second-year students. It would be incorrect to assume automatically that second-year medical students from two medical schools and one university are representative of other medical students without conducting further investigation. Most of the medical student subjects were studying osteopathic medicine, whereas the physician subjects had studied allopathic medicine. Are osteopathic and allopathic physicians different in ways that might have influenced the research data? The two sample sizes, each 40, were smaller than would be desirable. The small sample size had a decidedly negative

impact on this research because one would predict that differences in treatment of men and women patients would be slight. Therefore, if differences did actually exist and if the differences were small, the instruments to detect these differences would have to be powerful. Power for detecting differences is substantially increased by larger sample sizes.

The validity of the patient management problems used was not determined. Did they in fact measure what a medical person would do in actual clinical practice? More valid measures of clinical performance would be obtained with testing procedures which more closely approximated actual clinical practice. Instruments with higher reliability, particularly test-retest reliability, would have been desirable. With the test-retest reliabilities as low as they were, it would not be acceptable to assume that the subjects' performance on the instrument was indicative of how they would perform at another time. It is unfortunate that a medical student sample was used to gather test-retest reliability data rather than a physician sample, since physicians presumably would have more stable response patterns. The asthma problem had low interrater reliability among the experts before some of the items were eliminated. This strongly suggested that the management of asthma patients is not clear-cut. It would have been preferable to have had problems on which experts clearly agreed so that results of the

research could be accounted for by gender rather than having the confounding variable that chance differences in the management of the patient based upon differences in the acceptable modes of treatment may have accounted for the results.

After examining the research data, it seems reasonable to suggest that the research design was unnecessarily complex. More specifically, nothing seemed to have been gained by doing the study with repeated measures. It would have been preferable to combine the results on different problems to one set of scores rather than scores for each problem. The streamlining of the design which would have been preferable would entail deleting four of the measures leaving only the Overall Competence Index and the percentage of consultations requested which included discussion with the patient. The Overall Competence Index appears to have been an adequate measure, not needing to be broken down into parts.

A preferable research design would have been one that maintained the gender of the medical personnel and the patients as independent variables but reduced the dependent variables to only the Overall Competence Index on the total instrument. By this simplification, it appears that nothing would have been lost, yet the alpha level would have been increased to .05. A two-way analysis of variance would have been used had the design been simplified. For the physician sample, no statistically

significant differences would have been found, but statistically significant interaction between the gender of the patients and the gender of the medical students would have been found. Of course, in the original development of the research design it was not known that the measures in addition to the Overall Competence Index would not reveal additional or different information.

An additional limitation of the study was the difficulty encountered in the consultation measure. Because the occurrence of a subject requesting a consultation was low, the likelihood of detecting differences was very small. Differences must be extremely large to detect significance among events occurring either very frequently or very infrequently. Therefore, in this research instrument the problems should have included more need for the respondent to request consultations. If more consultations were requested, more opportunities would have existed to measure how often the medical personnel discussed the consultation with the patient.

Suggestions for Further Research

Taking into consideration the limitations of this present study as well as the intriguing, though statistically insignificant, results, this researcher believes that it would be appropriate and useful to further explore the effect of the gender of patients and medical personnel upon the medical treatment patients receive.

This researcher suggests that one possible way to further examine the effect of gender upon medical treatment would be to repeat this study, but with correction of the shortcomings in this research discussed in the preceding section. More specifically, it would be desirable to test the hypotheses that the gender of the medical person, the gender of the patient, and the interaction of their gender affects medical treatment on a larger sample. It would be highly desirable to establish the sample size based upon the alpha level (.05 would be reasonable), the power of the test at .80, and the recognition that the probable magnitude of differences in medical treatment based upon gender would be small. In order to do this, the sample size would have to be at least doubled (Cohen, 1965). Furthermore, it would be useful to conduct the research using a non-volunteer practitioner sample from a more varied geographical area. It would be interesting to conduct the research on medical students at different levels of training, such as on second-, third-, and fourth-year students. Improvements in the instrumentation are highly recommended by this researcher. Notably, instruments on which concurrent validities are known to be high with regard to actual clinical practice would be the most desirable. Furthermore, an instrument having higher (at least .80) test-retest and interrater

reliability coefficients should be used. It would be useful to have the instrument include more illnesses than only three so that the results of the research could be generalized to more disease entities. Finally, the research design should be simplified as suggested in the preceding section.

Another interesting research project directly stemming from this one would focus upon the consultation aspect of medical management. More specifically, the study would be designed to test the effect of the gender of medical personnel and patients upon how often the medical personnel discuss the consultation with the patient. In order to examine this, the instrument used would have to include numerous instances in which acceptable treatment of the patients would necessitate requesting consultations. This would eliminate the difficulties encountered with the consultation measure in this study which were discussed in the preceding section.

If it proved to be impossible to develop patient management problems which accurately reflect how medical personnel perform in actual clinical practice, there are several other possible ways to test the effect of gender upon patient management. It might be possible to program actors and actresses to play the parts of the patients in this study. Physicians and medical

students might then perform more closely to reality than on a paper-and-pencil simulation. The most ideal way to test the hypotheses in this study would be to have access to patient charts. If charts were maintained in a reliable, comprehensive way, it is conceivable that men and women patients with comparable problems and histories could be matched. Then comparisons could be made on the quality and efficiency of their care. Another interesting dimension which could be examined is the cost of the medical treatment. Are men and women patients charged the same amount for identical services? Are men and women patients equally asked to go through unnecessary medical tests?

In sum, this researcher believes that it would be appropriate and useful to further explore the effect gender of patients has upon the treatment they receive. The findings of this study suggest that there are no significant differences between the treatment of men and women patients. However, the evidence from this study is not sufficient to state conclusively that such differences do not exist. More powerful, focused studies on this topic would be necessary before final conclusions can be made.

In Retrospect

In retrospect the researcher would like to reiterate what was discussed in Chapter I. It would

be valuable to know if the gender of physicians and the gender of patients has an effect upon the quality of medical treatment. If it were revealed that male physicians treated male patients better and female physicians treated female patients better, patients could be assigned physicians of their own gender, or the training of physicians could attempt to reduce the effect of gender upon their treatment. On the other hand, if it were revealed that either women or men were given inferior medical treatment as a function of their gender, corrective efforts would be necessary in training physicians. The results of this study, though not statistically significant, do suggest that it would be inappropriate at this point to dismiss the possibility that gender affects medical care.

If this researcher were to have the knowledge that she now has prior to conducting this research, the changes she would make center around instrumentation. She would have attempted to perfect the instrument prior to running the actual study in terms of obtaining more information regarding the reliability and validity of the instrument. The asthma problem probably would have been deleted or, in any event, the items upon which the experts significantly disagreed would have been deleted prior to distributing the problems to the subjects. Idealistically, it would be nice to think

that more representative samples would have been used, but due to the financial limitations of this researcher, it is more realistic to assume that volunteers would again have been used. Her previous efforts to obtain nonvolunteer physicians to respond to the problems were thwarted. The research design would have been simplified according to the scheme discussed in the preceding section. Had more subjects and money been available to the researcher, a pilot study would probably have been run and hopefully would have revealed that the repeated measures design and the use of more specific measures such as the Information Gathering, Competence Index and the Management Competence Index provided no additional useful information.

The study has been fascinating to the researcher. It is her hope that this topic is one which will receive further attention and improved health care will be available for everyone as a result.

APPENDICES

APPENDIX A

INSTRUCTIONS TO THE EXPERT RATERS

RATER: _____

APPENDIX A

INSTRUCTIONS TO THE EXPERT RATERS SCORING THE PATIENT MANAGEMENT PROBLEM PROBLEM I

- I. Before assigning weights to the items in the attached Patient Management Problem (PMP), carefully review the problem. That is, look over each of the sections of the problem. In a general way formulate what information you would obtain to correctly identify the patient's problem, what the appropriate management is and when you would initiate it. Examine those pathways in the problem which would have detrimental effects on the patient.
- II. The problem will be scored along three basic dimensions:
(1) Information Gathering (i.e., history, physical, laboratory tests, and radiographs); (2) Management; (3) Path through Problem (this dimension represents the degree to which a student has made the correct choice at each decision point in the problem. For example, did he take a history when some immediate management should have been instituted?
- III. Scoring the information gathering ability of the student:
Section B has items related to the patient's history, physical, and laboratory tests. Go to Section B and rate the utility of each of the items as follows:
 - (1) Rate the item +2 if it is clearly indicated and important in identifying the patient's problem(s).
 - (2) Rate the item +1 if it is clearly useful but of a more routine nature.
 - (3) Rate the item as 0 if it is not useful in identifying the patient's condition or if it has a remote or debatable utility.
 - (4) Rate the item -1 if it is clearly not useful though not harmful.

- (5) Rate the item -2 if it is clearly contraindicated (i.e., in terms of risk, pain, or unjustifiably high cost).

Indicate your ratings by writing the weight next to the item.

Your task will be simplified if you first go through the entire Section and assign positive weights to the "ideal" choices; then go through the Section again and assign zero and negative weights to the items not chosen on the first PMP scoring.

IV. Scoring the management ability of the student: Sections C and D contain management decisions which can be initiated at various stages in the problem. You are to rate the appropriateness of each of these decisions given the condition of the patient. Go to each of these Sections and rate them as follows:

- (1) Rate the item +2 if the item is clearly indicated and important in the case of this patient.
- (2) Rate the item +1 if the item is clearly indicated but of a more routine nature, i.e., should be selected but not of special significance in the case of this patient.
- (3) Rate the item 0 if the item has debatable or remote utility in the case of this patient.
- (4) Rate the item -1 if it is clearly not indicated though not harmful in the case of this patient.
- (5) Rate the item -2 if it is clearly contraindicated (i.e., is definitely harmful or carries an unjustifiably high cost in terms of risk, pain or money) in the case of this patient.

V. Scoring the student's pathway through the problem (path score): At the end of Sections B and C and in Section A, there are a list of choices (decision points) concerned with what part in the patient management process the student wants to initiate next (i.e., take a history, initiate management, do a physical, etc.). The purpose of this score is to evaluate the degree to which a student has made the correct choice at each decision point in the problem. For example, did he take a history when he should have instituted some immediate management? Rate each of the items at the decision points of Sections A, B, and C as follows:

- (1) Rate the choice as +1 if it is clearly indicated and important in the case of this patient at this stage of the problem.
- (2) Rate the choice as 0 if it is not harmful to the case of the patient at this stage but is not clearly indicated or necessary to be done at this point.
- (3) Rate the choice as -1 if it is clearly contra-indicated given the condition of the patient at this stage in the problem.

The rating of the path score can be accomplished quite simply if you read the chief complaint and decide what you would do next. Mark a +1 for this choice. Go to the section of your choice and obtain all relevant information. Then, mark a +1 for your choice for the next step in the problem. Repeat this process until you reach the end of the problem. You have defined your ideal path through the problem. Then, go back and rate all the other choices at the decision points.

RATER: _____

SCORING THE PATIENT MANAGEMENT PROBLEM PROBLEM II

- I. Before assigning weights to the items in the attached Patient Management Problem (PMP), carefully review the problem. That is, look over each of the sections of the problem. In a general way formulate what information you would obtain to correctly identify the patient's problem, what the appropriate management is and when you would initiate it. Examine those pathways in the problem which would have detrimental effects on the patient.
- II. The problem will be scored along three basic dimensions: (1) Information Gathering (i.e., history, physical, laboratory tests, and radiographs); (2) Management; (3) Path through Problem (this dimension represents the degree to which a student has made the correct choice at each decision point in the problem. For example, did he take a history when some immediate management should have been instituted?)
- III. Scoring the information gathering ability of the student: Sections C and D have items related to the patient's history, physical, and laboratory tests. Go to each of these sections and rate the utility of each of the items as follows:
 - (1) Rate the item +2 if it is clearly indicated and important in identifying the patient's problem(s).
 - (2) Rate the item +1 if it is clearly useful but of a more routine nature.
 - (3) Rate the item as 0 if it is not useful in identifying the patient's condition or if it has a remote or debatable utility.
 - (4) Rate the item -1 if it is clearly not useful though not harmful.
 - (5) Rate the item -2 if it is clearly contraindicated (i.e., in terms of risk, pain, or unjustifiably high cost).

Indicate your ratings by writing the weight next to the item.

Your task will be simplified if you first go through each of the Sections and assign positive weights to the "ideal" choices; then go through the Sections again and assign zero and negative weights to the items not chosen on the first PMP scoring.

IV. Scoring the management ability of the student:

Sections E and F contain management decisions which can be initiated at various stages in the problem. You are to rate the appropriateness of each of these decisions given the condition of the patient. Go to each of these Sections and rate them as follows:

- (1) Rate the item +2 if the item is clearly indicated and important in the case of this patient.
- (2) Rate the item +1 if the item is clearly indicated but of a more routine nature, i.e., should be selected but not of special significance in the case of this patient.
- (3) Rate the item 0 if the item has debatable or remote utility in the case of this patient.
- (4) Rate the item -1 if it is clearly not indicated though not harmful in the case of this patient.
- (5) Rate the item -2 if it is clearly contraindicated (i.e., is definitely harmful or carries an unjustifiably high cost in terms of risk, pain or money) in the case of this patient.

V. Scoring the student's pathway through the problem (path score): In Sections A and B there are a list of choices (decision points) concerned with what part in the patient management process the student wants to initiate next (i.e., take a history, initiate management, do a physical, etc.). The purpose of this score is to evaluate the degree to which a student has made the correct choice at each decision point in the problem. For example, did he take a history when he should have instituted some immediate management? Rate each of the items at the decision points of Sections A and B as follows:

- (1) Rate the choice as +1 if it is clearly indicated and important in the case of this patient at this stage of the problem.

- (2) Rate the choice as 0 if it is not harmful to the case of the patient at this stage but is not clearly indicated or necessary to be done at this point.
- (3) Rate the choice as -1 if it is clearly contraindicated given the condition of the patient at this stage in the problem.

The rating of the path score can be accomplished quite simply if you read the chief complaint and decide what you would do next. Mark a +1 for this choice. Go to the section of your choice and obtain all relevant information. Then, mark a +1 for your choice for the next step in the problem. Repeat this process until you reach the end of the problem. You have defined your ideal path through the problem. Then, go back and rate all the other choices at the decision points.

RATER: _____

SCORING THE PATIENT MANAGEMENT PROBLEM
PROBLEM III

- I. Before assigning weights to the items in the attached Patient Management Problem (PMP), carefully review the problem. That is, look over each of the sections in the problem. In a general way formulate what information you would obtain to correctly identify the patient's problem, what the appropriate management is and when would you initiate it. Examine those pathways in the problem which would have detrimental effects on the patient.
- II. The problem will be scored along four basic dimensions: (1) Information Gathering (i.e., history, physical, laboratory tests, and radiographs); (2) Diagnostic Ability; (3) Management; (4) Path through Problem (this dimension represents the degree to which a student has made the correct choice at each decision point in the problem. For example, did he take a history when some immediate management should have been instituted?)
- III. Scoring the information gathering ability of the student: Sections B, C, D, and E have items related to the patient's history, physical, and laboratory tests. Go to each of these sections and rate the utility of each of the items as follows:
 - (1) Rate the item +2 if it is clearly indicated and important in identifying the patient's problem(s).
 - (2) Rate the item +1 if it is clearly useful but of a more routine nature.
 - (3) Rate the item as 0 if it is not useful in identifying the patient's condition or if it has a remote or debatable utility.
 - (4) Rate the item -1 if it is clearly not useful though not harmful.
 - (5) Rate the item -2 if it is clearly contraindicated (i.e., in terms of risk, pain, or unjustifiably high cost).

Indicate your ratings by writing the weight next to the item.

Your task will be simplified if you first go through each section and assign positive weights to the "ideal" choices; then go through each section again and assign zero and negative weights to the items not chosen on the first PMP scoring.

IV. Scoring the diagnostic ability of the student:

Section F contains a list of possible diagnoses for this problem. Notice that for a given problem it may be that more than one diagnosis is correct or plausible. Go to the diagnostic section and rate the correctness of each diagnosis listed. Rate them as follows:

- (1) Rate the item +1 if the diagnosis is clearly correct.
- (2) Rate the item 0 if the diagnosis is very plausible but not quite correct.
- (3) Rate the item -1 if the diagnosis is clearly incorrect.

V. Scoring the management ability of the student:

Section G contains management decisions which can be initiated at various stages in the problem. You are to rate the appropriateness of each of these decisions given the condition of the patient. Go to this section and rate the items as follows:

- (1) Rate the item +2 if the item is clearly indicated and important in the case of this patient.
- (2) Rate the item +1 if the item is clearly indicated but of a more routine nature, i.e., should be selected but not of special significance in the case of this patient.
- (3) Rate the item 0 if the item has debatable or remote utility in the case of this patient.
- (4) Rate the item -1 if it is clearly not indicated though not harmful in the case of this patient.
- (5) Rate the item -2 if it is clearly contraindicated (i.e., is definitely harmful or carries an unjustifiably high cost in terms of risk, pain or money) in the case of this patient.

VI. Scoring the student's pathway through the problem (path score): At the end of each Section and in Section A, there are a list of choices (decision points) concerned with what part in the patient management process the student wants to initiate next (i.e., take a history, initiate management, do a physical, etc.). The purpose of this score is to evaluate the degree to which a student has made the correct choice at each decision point in the problem. For example, did he take a history when he should have instituted some immediate management? Rate each of the items at the decision point of each section as follows:

- (1) Rate the choice as +1 if it is clearly indicated and important in the case of this patient at this stage of the problem.
- (2) Rate the choice as 0 if it is not harmful to the case of the patient at this stage but is not clearly indicated or necessary to be done at this point.
- (3) Rate the choice as -1 if it is clearly contra-indicated given the condition of the patient at this stage in the problem.

The rating of the path score can be accomplished quite simply if you read the chief complaint and decide what you would do next. Mark a +1 for this choice. Go to the section of your choice and obtain all relevant information. Then, mark a +1 for your choice for the next step in the problem. Repeat this process until you reach the end of the problem. You have defined your ideal path through the problem. Then, go back and rate all the other choices at the decision points.

APPENDIX B

SCORING KEYS

APPENDIX B

SCORING KEY

Problem I

Total number of items weighted positively: 51.

Number of pathway items (1-8, 45-47, 55-59) weighted positively: 6.

Number of information gathering items (9-44) weighted positively: 32.

Number of management items (48-54, 60-71) weighted positively: 13.

1.	+1	24.	+2	47.	-1
2.	+1	25.	-0.33	48.	+2
3.	-1	26.	+2	49.	+1
4.	+1	27.	+0.66	50.	+1
5.	-0.66	28.	+2	51.	-2
6.	+0.66	29.	+2	52.	-2
7.	0	30.	+2	53.	+1
8.	-1	31.	+2	54.	0
9.	+0.33	32.	+2	55.	+0.33
10.	+2	33.	+0.66	56.	-1
11.	+2	34.	+1	57.	-0.33
12.	+0.33	35.	+1.33	58.	-1
13.	+1.66	36.	+1.66	59.	-0.33
14.	+1	37.	+2	60.	+2
15.	+2	38.	+2	61.	+2
16.	+1.66	39.	-0.33	62.	+2
17.	-0.33	40.	+0.33	63.	+2
18.	+1	41.	+2	64.	+2
19.	+0.66	42.	+0.33	65.	+1.66
20.	+2	43.	+1	66.	+0.66
21.	0	44.	+1.66	67.	-1.33
22.	+1	45.	-1	68.	-2
23.	+0.33	46.	+1	69.	+0.33
				70.	+1.66
				71.	Eliminate

SCORING KEY

Problem II

Total number of items weighted positively: 28.

Number of pathway items (72-83) weighted positively: 6.

Number of information gathering items (86-110) weighted positively: 13.

Number of management items (114-133) weighted positively: 9.

72.	-1	93.	+1.33	114.	+0.66
73.	-1	94.	+0.33	115.	Eliminate
74.	+1	95.	-1	116.	Eliminate
75.	+1	96.	-1.66	117.	+2
76.	+0.33	97.	Eliminate	118.	+2
77.	-0.33	98.	Eliminate	119.	+1
78.	-0.66	99.	+2	120.	Eliminate
79.	+1	100.	+1.66	121.	-1.66
80.	-1	101.	0	122.	Eliminate
81.	0	102.	Eliminate	123.	Eliminate
82.	+1	103.	-1	124.	Eliminate
83.	+0.66	104.	-0.66	125.	-2
84.	Not Scored	105.	+1.33	126.	+0.66
85.	Not Scored	106.	-0.66	127.	-0.33
86.	+1.66	107.	-0.66	128.	+2
87.	+1.66	108.	+1	129.	+2
88.	+2	109.	-0.33	130.	-0.66
89.	+2	110.	+2	131.	+1.33
90.	+1.66	111.	Not Scored	132.	-2
91.	-0.33	112.	Not Scored	133.	+2
92.	+1.66	113.	Not Scored		

SCORING KEY

Problem III

Total number of items weighted positively: 70.

Number of pathway items (134-138, 151-155, 185-189, 195-199, 208-212, 221-225) weighted positively: 19.

Number of information gathering items (139-150, 156-184, 190-194, 200-207) weighted positively: 47.

Number of management items (226-233) weighted positively: 3.

Number of diagnosis items (213-220) weighted positively: 1.

134.	-1	168.	+0.5	202.	+1.5
135.	+0.5	169.	+2	203.	+1.5
136.	+1	170.	+0.5	204.	+1.5
137.	0	171.	+0.5	205.	+0.5
138.	-0.5	172.	+0.5	206.	+0.5
139.	+1	173.	+0.5	207.	-1.5
140.	+1.5	174.	+0.5	208.	+1
141.	+0.5	175.	+1	209.	+0.5
142.	+0.5	176.	+1	210.	+1
143.	+1.5	177.	+0.5	211.	+1
144.	+0.5	178.	+1.5	212.	+0.5
145.	+2	179.	+1.5	213.	-1
146.	+2	180.	+1.5	214.	-1
147.	+2	181.	+1	215.	-0.5
148.	+2	182.	+1	216.	-1
149.	+1.5	183.	+1	217.	-1
150.	+0.5	184.	+2	218.	-0.5
151.	+0.5	185.	+1	219.	+1
152.	+0.5	186.	+0.5	220.	-0.5
153.	+1	187.	+0.5	221.	0
154.	0	188.	0	222.	+0.5
155.	+0.5	189.	0	223.	-0.5
156.	+0.5	190.	0	224.	+1
157.	+0.5	191.	-2	225.	0
158.	+0.5	192.	+0.5	226.	+1.5
159.	0	193.	+2	227.	-0.5
160.	0	194.	0	228.	-2
161.	0	195.	0	229.	+1.5
162.	+0.5	196.	0	230.	-1.5
163.	+0.5	197.	+0.5	231.	0
164.	+2	198.	+1	232.	+1
165.	+2	199.	+0.5	233.	-1
166.	+0.5	200.	+1		
167.	+2	201.	+1.5		

APPENDIX C

LETTERS TO SUBJECTS

APPENDIX C

LETTERS TO SUBJECTS

Dear

Despite the fact that I recognize you receive a great many requests for your attention, I hope you will be able to devote some time to this request. A new faculty member of our Department is conducting research regarding physicians' decision-making processes. This project has my full endorsement.

I would greatly appreciate your response to the enclosed materials. There are three patient management problems, an answer booklet, and a latent access pen enclosed. Your responses to these written simulated patient problems will provide some indication of how you would respond to these patients in actual practice. The purpose of these problems is to further explore the process by which physicians manage patients. The intent of the research is not directed toward who selected "right" or "wrong" options, but rather the pattern of the management responses. The time it would take you to respond to the three problems approximates 60 minutes. I think you may find this process of responding to patient problems rather enjoyable because you will obtain new data about the patients as you progress through the problems.

We shall be happy to provide you with feedback regarding your responses to the problems. Please read the enclosed directions and progress through the problems as directed. A stamped, addressed envelope is enclosed for the return of your answer booklet and latent access pen. You could expect feedback on your responses within two months. Your earliest response to the problems is appreciated. Thank you for your attention to this project.

Sincerely,

Norbert B. Enzer, M.D.
Chairman, Department of Psychiatry

ewp
enclosures

Dear

As you may recall from the letter I wrote to you one month ago, I recognize that you are a busy person. However, I would be most grateful if you could find time to respond to the three patient management problems which were sent with my preceding letter. A special pen and stamped addressed envelope were also enclosed for your use. The purpose of asking you to respond to the patient management problems is to provide additional data on how physicians make decisions in treating patients. The research being conducted is examining the pattern of management responses, rather than their "accuracy." The time it will take to respond to the problems approximates one hour.

If you are able to respond to this request within the next six weeks, we shall be happy to provide you with feedback regarding your responses. Thank you for your consideration of this project.

Sincerely,

Norbert B. Enzer, M.D.
Chairman, Department of Psychiatry

NBE/soa

REFERENCES

REFERENCES

- Archer, E. The birth of conspiracy. Health/PAC Bulletin, 1970, 17, 10-13.
- Barton, C. C. Can women's liberation movement help hospitals? Modern Hospital, 1971, 116, 65.
- Batt, R. Creating a professional identity. American Journal of Psychoanalysis, 1972, 32, 156-162.
- Boston Women's Health Book Collective. Our bodies ourselves. New York: Simon and Schuster, 1971.
- Brodsky, C. M. The pharmacotherapy system. Psychosomatics, 1971, 11, 24-30.
- Cartwright, L. K. Conscious factors entering into decisions of women to study medicine. Journal of Social Issues, 1972, 28, 201-215 (a).
- Cartwright, L. K. Personality differences in male and female medical students. Psychiatry in Medicine, 1972, 3, 213-218 (b).
- Chesler, P. Women and madness. New York: Doubleday, 1972.
- Cohen, J. Some statistical issues in psychological research. In B. B. Wolman (Ed.), Handbook of clinical psychology. New York: McGraw-Hill, 1965.
- Cooperstock, R. Sex differences in the use of mood-modifying drugs: An explanatory model. Journal of Health and Social Behavior, 1971, 12, 238-244.
- Discrimination against women. Hearings before the special subcommittee on education of the Committee on Education and Labor, House of Representatives. Ninety-first Congress, second session on Section 505 of H.R. 16098. Hearings held in Washington, D.C., July 1 and July 31, 1970.

- Donnelly, M. B. Analysis of the 1971-72 year III PMP exam: Report #1, Unpublished manuscript, Wayne State University, 1972.
- Easley, E. B. The dilemma of women in our culture: Gynecologic repercussions. Part II. American Journal of Obstetricians and Gynecologists, 1971, 110, 858-864.
- Ehrenreich, B. & English, D. Witches, midwives, and nurses: A history of women healers. Old Westbury: Feminist Press, 1973.
- Elstein, A. S., Kagan, N., Shulman, L. S., Jason, H., & Loupe, M. J. Methods and theory in the study of medical inquiry. Journal of Medical Education, 47, 85-92.
- Fidell, L. Put her down on drugs: Prescribed drug usage in women. Paper presented at Western Psychological Association Meeting, Anaheim, April, 1973.
- Frankfort, E. Vaginal politics. New York: Bantam Books, 1972.
- Freedman, A. M., Kaplan, H. I., & Sadock, B. J. Modern synopsis of psychiatry. Baltimore: Williams & Wilkins, 1972.
- Gilbert, M. Women in medicine. In R. Morgan (Ed.), Sisterhood is powerful. New York: Vintage Books, 1970.
- Greer, G. The female eunuch. New York: Bantam Books, 1970.
- Gurin, G., Veroff, J., & Feld, S. Americans view their mental health. New York: Basic Books, 1960.
- Hole, J., & Levine, E. Rebirth of feminism. New York: Quadrangle Books, 1971.
- Jervis, N. Health care for women. Health/PAC Bulletin, 1972, 47, 14-15.
- Kelly, M. J. Who me--a sexist? The New England Journal of Medicine, 1973, 288, 317-318.
- Kotelchuck, R., & Levy, H. Federal health cutbacks. Health/PAC Bulletin, 1973, 52, 3-18.

- Lennane, K. J., & Lennane, R. J. Alleged psychogenic disorders in women--a possible manifestation of sexual prejudice. The New England Journal of Medicine, 1973, 288, 288-292.
- Linn, L., & Davis, M. The use of psychotherapeutic drugs by middle aged women. Journal of Health and Social Behavior, 1971, 12, 331-340.
- Linn, L. Physician characteristics and attitudes toward legitimate use of psychotherapeutic drugs. Journal of Health and Social Behavior, 1971, 12, 132-140.
- The male-feasance of health. Health/PAC Bulletin, 1970, 17, 1-2.
- McGuire, C. H., & Solomon, L. M. (Eds.) Clinical simulations. New York: Appleton Century Crofts, 1971.
- McGuire, C. H., Solomon, L. M., & Bashook, P. G. Handbook of written simulations: Their construction and analysis. Chicago: University of Illinois, 1972.
- Mellinger, G. D., Balter, M. B., & Manheimer, D. I. Patterns of psychotherapeutic drug use among adults in San Francisco. Archives of General Psychiatry, 1971, 25, 385-394.
- Myers, P. F. (Ed.) 1975 world population data sheet. Washington, D.C.: Population Reference Bureau, 1975.
- Norris, F. S. Medical womanpower. The New England Journal of Medicine, 1969, 281, 910-911.
- Peterson, R. A. Vocational interest patterns of male and female medical students over a four-year period. Journal of Counseling Psychology, 1972, 19, 21-25.
- Phillips, D. L., & Segal, B. E. Sexual status and psychiatric symptoms. American Sociological Review, 1969, 34, 58-72.
- Physician's prejudice. Human Behavior, 1973, 3, 50-51.
- Prather, J., & Fidell, L. Put her down and drug her up. Paper presented at American Sociological Association meeting, New Orleans, August, 1972.

Report of the women's action program. Washington, D.C.:
U.S. Department of Health, Education, and Welfare,
1972.

Reverby, S. Health: Women's work. Health/PAC Bulletin,
1972, 40, 15-20.

Shader, R., Binstock, B., & Scott, A. Subjective determinants of drug prescription: A study of therapists' attitudes. Hospital and Community Psychiatry,
1968, 19, 384-387.

Shapley, D. Medical education: Those sexist putdowns may be illegal. Science, 1974, 184, 449-451.

Telinde, R. W., & Mattingly, R. F. Operative gynecology.
(4th ed.) Philadelphia: Lippincott, 1970.

White, K. L. Life and death and medicine. Scientific American, 1973, 229(3), 23-33.

Wilson, R. J. Health care for women: Present deficiencies and future needs. Obstetrics and Gynecology, 1970, 36(2), 18.

Women, pain, and prejudice. The New England Journal of Medicine, 1973, 288, 968-970.

MICHIGAN STATE UNIV. LIBRARIES



31293103482034