# A STUDY OF FACTORS AFFECTING CLINICAL PERFORMANCE GRADES 

 OFNURSING STUDENTS
by Accompanying
Vidercassctite in MARY BOGHOS KRUGER

Special Collections
B.S.N., University of British Columbia, 1973

A THESIS SUBMITTED IN PARTIAL FULFILMENT
OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN NURSING in THE FACULTY OF GRADUATE STUDIES (School of Nursing)

We accept this thesis as conforming to the required standard

THE UNIVERSITY OF BRITISH COLUMBIA
February 1980
(c) Mary Boghos Kruger, 1980

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the Head of my Department or by his representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.



#### Abstract

A study of factors affecting clinical performance grades of nursing students


This study involved an investigation of factors affecting clinical performance grades of first and second year nursing students in a two-year Diploma program.

The purpose of the study was to increase knowledge of factors which may correlate with clinical performance of nursing students; related goals were to assist nurse educators in predicting student clinical performance and providing guidance accordingly.

The study was planned to test the following hypotheses:

1. Grades in Nursing, Biology, and Psychology courses during the first semester of the nursing program account for a significant proportion of the variance in clinical performance scores of students;
2. Complexity of the nursing situation accounts for an additional significant proportion of the variance in clinical performance scores of students. The study population consisted of 59 first and second year nursing students. The dependent variable was the average of the clinical performance scores of the students given by the clinical instructors (who wrote and scored the anecdotal records) and two nursing judges (who scored the written anecdotal records). The independent variables were the grades in first semester Nursing, Biology, and Psychology; and the complexity of the nursing situation.

Data for the study were collected over a three-week period. This phase included: (a) writing of anecdotal records of students' clinical performance by their instructors, and subsequently scoring of these records by the clinical instructor and two nursing judges using the scoring instrument; (b) assessing the level of complexity of the students' clinical assignments by their clinical instructors, using the complexity of the nursing situation instrument; and (c) obtaining the students' first semester Nursing theory, Psychology, and Biology grades from their permanent records. The clinical instructors and the nursing judges were trained in the writing and scoring of anecdotal records, and in determining the complexity of the students' assignments by utilizing a 10 minute videotape prepared by the investigator depicting the performance of a nursing student in a simulated clinical situation. Analyses of the data were carried out in two phases. Phase one included product moment correlational analysis and generalizability analysis to determine the reliability of the clinical scores. Phase two included stepwise multiple regression analysis to determine the predictors of clinical scores. The findings of the study showed that the inter-rater correlations among the clinical instructors and two nursing judges were moderately strong (. 58 to .84 ). Likewise generalizability analysis showed that the clinical scores were reliable. The data on complexity of the nursing situation showed very litfle variability in each of the four semesters and were therefore deleted from the regression analysis. The regression analysis showed that in Semesters II and III 51 and 76 percent of the variance in clinical performance scores could be accounted for by the grades
in first semester Nursing, Psychology, and Biology. However, in Semesters IV and $V$ the amount of variance accounted for by the independent variables was not significant. Thus, hypothesis 1 is accepted in the case of Semester II and III students, and rejected in the case of Semester IV and $V$ students.

## TABLE OF CONTENTS

Page
ABSTRACT ..... ii
LIST OF TABLES ..... x
LIST OF FIGURES ..... xii
ACKNOWLEDGEMENTS ..... xiii
CHAPTER

1. INTRODUCTION ..... 1
Purpose of the Study ..... 5
Statement of the Problem ..... 6
Hypotheses ..... 6
Variables ..... 6
Significance of the Study. ..... 7
Operational Definitions ..... 9
Assumptions. ..... 11
II. REVIEW OF RELATED LITERATURE. ..... 12
Introduction ..... 12
Educational Evaluation ..... 12
Clinical Performance Evaluation Versus Classroom Evaluation. ..... 15
Observational Techniques in Evaluation ..... 21
Anecdotal Record Technique of Evaluation ..... 22
CHAPTER Page
Prediction of Academic Performance Versus Clinical Performance ..... 25
The Criterion Problem ..... 25
Predictive Studies ..... 26
Summary ..... 29
III. RESEARCH METHODOLOGY ..... 31
Introduction ..... 31
Sampling ..... 31
Institution ..... 31
Semesters ..... 32
Weeks ..... 32
Instructors ..... 33
Students. ..... 34
Semester II sample ..... 34
Semester III sample ..... 35
Semester IV sample ..... 35
Semester V sample ..... 35
Confidentiality ..... 36
Instrumentation ..... 36
Complexity Scale ..... 36
Anecdotal Record ..... 37
Scoring system ..... 37
Procedure ..... 38
Development of Videotape ..... 38
The Training of Clinical Instructors and Judges ..... 39
Data Collection ..... 40
Method of Analysis ..... 41
Phase One. ..... 41
Phase Two ..... 42
IV. RESULTS OF ANALYSIS ..... 44
Complexity of the Nursing Situation ..... 44
Inter-Rater Reliabilities of Clinical Scores ..... 45
Correlational Analyses ..... 45
Semester II ..... 45
Semester III ..... 45
Semester IV ..... 45
Semester V ..... 45
Generalizability Analyses. ..... 50
Semester II ..... 50
Semester III ..... 50
Semester V ..... 54
Correlation Coefficients Versus Generalizability Coefficients ..... 54
Regression Analyses. ..... 57
Preliminary Analyses: Raw Clinical Scores Versus Standardized Clinical Scores ..... 57
Semester II ..... 57
Semester III ..... 57
Final Analyses: Raw Clinical Scores ..... 60
Semester II ..... 60
Semester III ..... 62
Semester IV ..... 64
Semester V . ..... 64
V. SUMMARY, FINDINGS AND CONCLUSIONS, LIMITATIONS, AND RECOMMENDATIONS ..... 68
Findings and Conclusions ..... 71
Limitations of the Study ..... 76
Recommendations ..... 77
REFERENCES ..... 79
APPENDICES ..... 83
APPENDIX A: Letter Requesting College Permission ..... 84
APPENDIX B: Memo Granting College Permission ..... 86
APPENDIX C: Explanations to Instructors. ..... 88
APPENDIX D: Consent Form (Instructors) ..... 91
APPENDIX E: Explanations to Students ..... 93
APPENDIX F: Consent Form (Students) ..... 95
CHAPTER Page
APPENDIX G: Complexity of the Nursing Situation ..... 97
APPENDIX H: Anecdotal Record ..... 102
APPENDIX I: Scoring System ..... 104
APPENDIX J: Generalizability Analysis ..... 112
APPENDIX K: Standardization Procedure ..... 116
APPENDIX L: Stepwise Multiple Regression Analysis ..... 119

## LIST OF TABLES

TABLE Page

$$
\begin{aligned}
& 1 \text { Means, Standard Deviations, and Intercorrelations Among } \\
& \text { Nursing Judges (J) in Semester II . . . . . . . . . . . . . . . } 46
\end{aligned}
$$

2 Means, Standard Deviations, and Intercorrelations Among Nursing Judges (J) in Semester III ..... 47
3 Means, Standard Deviations, and Intercorrelations Among Nursing Judges ( J ) in Semester IV ..... 48
4 Means, Standard Deviations, and Intercorrelations Among Nursing Judges (J) in Semester V ..... 49
5 Analysis of Variance of Clinical Scores of Semester II Students: Effects of Times on Judges ..... 51
6 Variance Components Contributing to Universe Score Variance and Expected Observed Score Variance of Judges and Weeks, and Generalizability Coefficients of Semester II. ..... 52
7. Analysis of Variance of Clinical Scores of Semester III Students: Effects of Times on Judges ..... 53
8 Analysis of Variance of Clinical Scores of Semester V Students: Effects of Times on Judges ..... 55
9 A Comparison of Generalizability Coefficients and Mean Correlation Coefficients Across Judges in Semesters II, III, and V ..... 56
10 Summary of Stepwise Multiple Regression Analysis of Week 3 of Semester II: A Comparison of Raw Clinical Scores to Standardized Clinical Scores. ..... 58
11 Summary of Stepwise Multiple Regression Analysis of Week 3 of Semester III: A Comparison of Raw Clinical Scores to Standardized Clinical Scores. ..... 59
12 Summary of Stepwise Multiple Regression Analyses: Semester II . ..... 61
13 Summary of Stepwise Multiple Regression Analyses: Semester III . . 63
14 Summary of Stepwise Multiple Regression Analyses: Semester IV . . ..... 65
15 Summary of Stepwise Multiple Regression Analyses: Semester V . . 66

## LIST OF FIGURES

FIGURE Page
1 Standardization Procedure. ..... 118

## ACKNOWLEDGEMENTS

I would like to express my gratitude to the administrators, nursing department chairperson, members of the audiovisual department, participating nursing instructors, and students of Semesters II, III, IV, and V of Vancouver Community College, Langara, for making this study possible.

To the members of my thes is committee--Mrs. Sheila Stanton, the chairperson; Dr. Margaret Campbell; Dr. Todd Rogers; and Dr. Robert Conry-I wish to express my sincerest gratitude for the support, guidance, and expertise that they provided to me throughout the conception and execution of the study.

Most particularly, I wish to thank my husband for his patience, understanding and encouragement.

## Chapter I

## INTRODUCTION

Evaluation of students in nursing education programs and attempts to identify factors which affect their clinical performance continue to be complex and challenging tasks. There are a number of reported nursing studies in which the objective was to identify factors associated with successful completion of a program of study. Fewer studies have focussed on identifying predictors of clinical performance. The latter is the focus for the present study. It will attempt to identify academic and situational factors associated with successful clinical performance.

Taylor, Nahm, Quinn, Harms, Mulaik, and Mulaik (1963), in a comprehensive review of research in the area of selection and recruitment of nursing students, made this statement:

Although it is useful to know how well a test works for groups in general, the more important question is how well it works for a particular group in the situation under consideration. However, the most important question may be: How well.will the graduate nurse perform on the job? The present test batteries were not designed or validated to answer this question. The few studies that have attempted to use certain test scores to predict success on the job have shown that there is no correlation between these test performances and performance on the job. Here, then, is a wide-open area that urgently needs to be investigated ( $p$. 17).

They further expounded that no one test or test battery can be specified as the best predictor for all nursing schools, and that any test will vary in usefulness from school to school and from year to year. Consequently, they suggested that each school should evaluate its own selection program.

In summarizing selection and recruitment measures, Taylor et al. (1963) stated that grade point average and other tests of cognitive abilities appeared to be the best predictors of academic success in nursing theory courses. However, grade point average, personality, and attitude tests have not been useful in predicting success in the practical aspects of nursing education or on-the-job performance. These authors have concluded that there is a great need for research utilizing multiple variables of achievement, aptitude, personality characteristics, motivation, interests, and other types of predictors.

In 1965, Taylor et al. reported research on the related problem of attrition, including this statement:

Further research in this area is sorely needed because of the high cost, both to the individuals and institutions of attrition in nursing schools and turnover among nursing personnel. Estimates of the cost of recruiting, selection, inducting, and training one individual for different occupations in the world of work range from hundreds to thousands of dollars to the organization, not counting the loss of a person in the training spot or the headaches of repeating the entire cycle to get a replacement into the status of a full-fledged worker. There is also an unestimated psychological and financial cost to the individual in terms of loss of time and effort and the personal disappointment and frustrations when one has an unsuccessful experience and perhaps even finds himself in the wrong field (pp. 52-53).

The findings of two more recent studies (Goza, 1970; Gore, 1973) have supported the statements quoted above and, more importantly, have provided new findings comparing dropouts and graduates with favorable performance evaluation.

Goza developed profiles of dropouts, graduates, graduates who passed the State Board Test Pool Examination (SBTPE) in nursing on the initial attempt, graduates who did not pass SBTPE, graduates who received a favorable work performance evaluation, and a sample of associate degree nursing students. He found that graduates had significantly higher means than the dropouts on the final grade point average and on the Responsibility scale of the Gordon Personal Profile. Likewise, the graduates who passed the SBTPE had means significantly higher than the SBTPE failures on American College Testing Program, California Test of Mental Maturity, final GPA, all 14 of the National League for Nursing Achievement tests, the Sociability scale of the Gordon Personal Profile, and the Original Thinking scale of Gordon Personal Inventory. He also found an interesting difference between means of graduates who received a favorable work performance evaluation and of those who did not on the Ascendency scale of the Gordon Personal Profile and the Vigor scale of Gordon Personal Inventory.

Goza concluded that the reason for minimum academic success of dropouts was due to the lack of responsible attitude, rather than a deficiency in ability; and that unfavorable work performance evaluation received by some of the graduates appeared to be the result of personality differences rather than differences in ability.

Research done by Gore (1973), who developed prediction equations for three associate degree nursing programs, showed that prediction equations could not be generalized from one institution to another. He therefore recommended that each nursing school should develop their own predictive criteria, rather than
using common ones. Goza's findings also revealed that students who were predicted to fail in one nursing program may well be predicted to pass in another nursing program. He concluded that this should be a crucial factor in the student's choice of nursing program.

The findings of Goza and Gore support Taylor et al. 's advice that each nursing school determine its own selection and recruitment criteria, so that they can adequately counsel their prospective applicants and provide guidance throughout their training program.

Another factor which may affect clinical performance is complexity of the student's assignment. Traditionally, patient assignments of graduate nurses were determined by the number of nursing activities/tasks to be performed by the nurse in a given period of time, and the nature or complexity of these activities. Of the many variables that affect quality patient care, the designers of quality assurance programs in nursing have identified complexity of nursing activities as a major factor contributing to the provision of quality nursing care to individual patients. The proponents of qual ity patient care programs believe that adequate staffing of nursing units--staffing based on the number of patients in a unit, number of nursing activities and the complexity of these activities--directly influences the type of care given to the patients in these units.

If the clinical performance of the graduate nurse is influenced by the number of patients she is assigned to, which in turn is based upon the number of activities she has to perform and their level of complexity, then the performance of the student nurse will also be influenced by the type of patient assignment she is given.

Of the studies done on predictors of clinical performance, there was no mention of complexity of the nursing assignment/situation or how it may influence performance. However, on numerous occasions, the investigator has noted that staff, instructors, and students suggest that complex assignments are the cause of "poor" performance. This hypothesis merits investigation, to ascertain whether, and to what extent, situational complexity affects clinical performance.

Many factors influence a nursing student's successful completion of a nursing program. In order to identify some of the factors which may affect a nursing student's clinical performance, this study will identify the extent to which academic grades in first semester Nursing, Biology, and Psychology can be used to predict the clinical performance of nursing students in a two-year community college nursing program. Also, the complexity of the student's assignment will be studied to determine the extent to which this variable affects clinical performance grades. The complexity of the nursing activities was added because recent studies have identified it as a major factor contributing to the nurse's performance and hence to quality patient care.

Purpose of the Study

The purpose of this study is to increase knowledge of the factors which may correlate with clinical performance of nursing students; related goals are to assist nurse educators in predicting student performance and providing guidance.

## Statement of the Problem

Specifically, this study will determine the degree to which (a) academic grades and (b) complexity of the nursing situation explain scores in clinical performance for first and second year nursing students.

Hypotheses

1. Grades in Nursing, Biology, and Psychology courses during the first semester of the nursing program account for a significant proportion of the variance in clinical performance scores of students.
2. Complexity of the nursing situation accounts for an additional significant proportion of the variance in clinical performance scores of students.

## Variables

The dependent variable employed in this research was the clinical performance score: the average of marks given by the clinical instructor (who wrote and scored the anecdotal records of the student's clinical performance during the weeks of January 30, and February 6 and 13, 1978), and the scores given by two nursing judges (instructors with a working knowledge of clinical instruction, who evaluated the student's performance using the anecdotal records written by the clinical instructors) from that semester.

The independent variables used in the prediction were:

1. the grades in the first semester Nursing, Biology, and Psychology courses;
2. the complexity of the nursing situation as assessed by the clinical instructor.

## Significance of the Study

Many of the prediction studies reviewed have focussed on success in nursing school as the variable of concern. However, only a few studies have investigated the relationship between theory scores and clinical performance, which is more relevant to the ultimate goal of preparing the nurse to perform successfully in a variety of clinical settings. Furthermore, although the few studies which have studied academic ability as a predictor of grades in clinical performance have found negative or weak positive relationships, they did not assess (1) the reliability of the clinical performance grades or (2) the relative contributions of the compulsory courses--Nursing theory, Biology, and Psychology-which are deemed to be theoretical knowledge prerequisite to the provision of the type of care that individual patients require or request.

A review of the literature on tools/methods used to evaluate clinical performance revealed one study by Rines (1963) that reported "77 percent of the nursing instructors used anecdotal records with varying degrees of success" (p. 26).

In summary, the need to identify predictors of clinical performance arose from the enduring problem of how best to prepare nurses. This study was initiated as a result of dissatisfaction, expressed by both students and instructors, with clinical evaluation in general and specifically with procedures for assigning clinical grades. Furthermore, an unpublished survey of nursing students' attitudes toward various methods of clinical evaluation (conducted by the investigator in 1976) showed that students generally prefer a combination of several evaluation methods: anecdotal records, simulation, clinical examination, and self evaluation. It was beyond the scope of the present study to examine several evaluation techniques in evaluating clinical performance, and then to use the method(s) which had the best reliability, validity, objectivity, and practicality. Therefore, the anecdotal record method of evaluation was chosen, since it was currently in use to some extent, and based on the belief that it would provide instructors with data to use in the course of clinical teaching and evaluation.

Finally, results of this study could help to identify those students in this school who may need considerable counselling and assistance during the course of their programs, and thus enable instructors to take some preventative action and possibly to decrease attrition and the cost of nursing education.

## Operational Definitions

The following terms are used:
nursing student - a student who is enrolled in a two-year college nursing program preparatory to nurse registration.
nursing program - a six-semester program in a community college leading to a diploma in nursing.
clinical area - a hospital ward.
clinical objectives - learning objectives to be achieved by the students in the clinical area.
performance - the behaviors of the nursing student in the clinical area, as related to relevant clinical objectives.
clinical instructor - a nursing instructor who teaches, observes, and evaluates the student's performance in the clinical area.
performance evaluation - an evaluation of the student's performance to determine whether clinical objectives have been met. nursing judge - a nursing instructor, with a working knowledge of instruction in any one of Semesters II, III, IV and V of the nursing program, who evaluates the performance of the nursing student using the anecdotal notes written by the clinical instructor.
clinical performance score - the average of the scores given by the clinical instructor and two nursing judges, each of whom evaluated anecdotal notes on the performance of the student during the weeks of January 30, February 6, and February 13, 1978.
$\underline{\mathrm{SCl}}$ - scores given by the clinical instructor.
SNJ - scores given by the nursing judge.
Grade Point Equivalent - the letter grades of students in first semester Nursing, Biology, and Psychology were transformed to numerical grades using the scale below.

Letter Grade*
A

B

C

P
F

Grade Point Equivalent
4 3 2 1 0
complexity of the nursing situation (CNS) - is determined by four variables:
(1) the activities required of the student in the clinical area, including
(a) administering treatments prescribed by the patient's physician;
(b) administering treatments prescribed by the nursing team;
(2) the condition of the patient;
(3) the availability to the student of appropriate patient experiences;
(4) the revision of the student's assignments based on the instructor's assessment of the student's abilities.

[^0]
## Assumptions

The design and execution of this study rests upon these assumptions:

1. the clinical objectives of a given semester are decided and agreed upon by the instructors responsible for that semester, and hence represent their expectations of the students enrolled in that semester.
2. the clinical instructor can validly and reliably judge the complexity of the nursing situation.

## Chapter II

## REVIEW OF RELATED LITERATURE

Introduction

Adequate evaluation and adequate prediction of the students ${ }^{2}$ clinical performance are problems that nursing education shares with other cognate disciplines such as medical education and dental education. The process of clinical evaluation requires observational techniques for collecting data in various clinical settings. This very process, although used extensively for decades, has created problems for the educators and the students alike.

## Educational Evaluation

Gronlund (1976) defined evaluation as a "systematic process of determining the extent to which educational objectives are achieved by pupils" (p. 6). He distinguished between evaluation and measurement, stating, "Evaluation is a much more comprehensive and inclusive term than measurement" (p. 6), which includes both quantitative and qulitative descriptions of pupil behavior, plus value judgments concerning the desirability of that behavior. Similarly Gordon (1960) described evaluation as". . . the measure of the extent to which the performance achieves certain desired goals" (p. 364).

In describing evaluation of students, MacKay (1974) stated that ". . . we must view evaluation as a means and not an end, as an ongoing process with the purpose being promotion of fuller development of the learner" (p. 4). Likewise, Gordon (1960) asserted that "The most important use of an evaluation is to assist the person to improve his performance" (p. 366).

In her effort to establish productive evaluation, MacKay (1974) has
outlined the following guidelines:

1. the purpose of evaluation as a means of facilitating learning should be known and accepted by both student and teacher;
2. the standards of evaluation should be known and shared, or goals for behavior based on assessed learning needs should be established;
3. the standards or goals should be realistic; the student should be potentially capable of meeting the objectives for performance normally expected for that level of preparation and his unique idiosyncrasies taken into consideration; and
4. behaviors within a specified time interval only should be evaluated. (p. 5)

Using a different perspective, Schweer and Gebbie (1976) stated that the trend in evaluation of clinical nursing is toward an ongoing evaluation program aimed at determining student growth in becoming a skillful practitioner. They also identified purposes that evaluation processes could serve. These are as follows:

1. to determine the background ability each student brings to the learning situation;
2. to determine the ability of students to use nursing theory as a whole while progressing in the development of clinical skills;
3. to determine the progress by each student at frequent intervals during the assignment;
4. to discover learning difficulties of individual students and the group in order to adjust the teaching, including remedial or advanced assignments, as needed;
5. to provide reinforcement of learning for students;
6. to foster development of self-evaluation;
7. to determine the readiness of students to become self-directive;
8. to determine the effectiveness of teaching techniques; and
9. to obtain data for conducting research studies relating to student achievement. (p. 166)

Evaluation as described by Gordon (1960), MacKay (1974), and Schweer and Gebbie (1976) suggests that they are using this term to mean both formative and summative evaluation of performance. Formative evaluation helps the student to learn by pointing out his strengths and weaknesses, and summative evaluation measures to what extent the student has met the course objectives.

Most students and teachers fail to make the distinction between evaluation and grading. However, the distinction was made by MacKay (1974), who stated that "the assignment of grades requires a further step of measuring the achieved levels of competency in respect to specified behavior characteristics of students in general in differing levels of expertise" (p.5).

To establish educationally sound evaluation programs, Gronlund (1976)
and Rines (1963) have proposed the following:

1. identify and define the objective of the program in terms of desired changes in the behavior of students;
2. determine evaluation tools;
3. construct tools which are valid and reliable;
4. decide on utilization of these tools;
5. decide what points in the course will be appropriate to evaluate the student's performance;
6. outline a program of evaluation;
7. plan and select situations to show desired change of behavior;
8. collect the responses; and
9. interpret the results.

The above program of evaluation indicates that the specific activities of the educator related to evaluation are also a part of the teaching process, since they help in defining the objectives, and in the planning and selecting of situations to show the desired behavior of the student. Although the interdependence of teaching and evaluation is recognized and accepted by nurse educators, Rines (1963) makes a strong point that

If evaluation is to be conducted with maximum efficiency and results, regular times should be designated for the purpose, separate from the teaching process . . . [the instructor] should plan her observations and evaluations of student behaviors just as she plans for paper-andpencil tests and examinations. These should be planned frequently enough . . . so that she [the instructor] cangetan estimate of behavior trends. (pp. 54-55)

Clinical Performance Evaluation<br>Versus Classroom Evaluation

Most nurse educators agree that the process of evaluation in the clinical area is difficult and complex (cf. NLN Report, 1972; Litwack, Sakata, and Wykle,
1972). In the clinical area, the student is learning to apply knowledge acquired in the classroom and laboratory. Moreover, as Wood (1972) has suggested, "The student is learning behavioral more than intellectual skills" (p. 336).

Most nurse educators will also agree that the achievement of clinical objectives is poorly measured by conventional classroom tests, whether teachermade or standardized. Furthermore, clinical experiences are not uniform--not all students will have the opportunity to care for the same patient with a given problem. Therefore, to the extent that each individual and his reactions are unique, the clinical experiences of the students will vary greatly, and thus make standard teaching and evaluation almost impossible.

Another difference between clinical and classroom evaluation is that various individuals in the clinical area add input and thus affect the student's behavior. It is therefore crucial that the input of these individuals--patients, nursing staff, and clinical instructors--be utilized for evaluation of the student's perfórmance.

Taylor et al. (1965) have pointed out that "the lack of relationship between performance in theory courses and clinical practice courses may point up a deficiency in traditional methods of evaluating students" (p.38). To explain the above, they have quoted the following shortcoming in nursing education initially described by McManus (1949):
many of those who teach nurses . . . have had no formal training as teachers and therefore tend to follow the patterns and teaching they themselves have experienced. They are not aware of the need to develop the higher mental abilities in this field as in other professions. Only exceptional teachers and students manage to free themselves from traditional patterns of teaching and learning.

Among the nursing educators who have had some preparation for teaching are many who still erroneously use instructional methods which demand the exercise of little more than the lower mental processes of memorization and recall. Frequently facts are taught and facts are learned in catechismal fashion. Testing and grading systems tend to put a premium upon fact-getting and rote-learning of subject matter. Skills are frequently taught exclusively in the classroom, with drill on specific procedures. Although the students spend a considerable portion of their time in the clinical nursing practice, the theory is likely to remain sterile and compartmentalized and will rarely, and only accidentally transfer unaided to give the practice meaning. . . . (pp. 38-39)

Lancaster (1972) in her Opinion Survey of nurse teachers in Scotland has addressed the question of
whether the present tutors' courses give the kind of preparation which will enable future nurse teachers to cope effectively with the professional demands made upon them. The kind of work done in the past . . . in the traditional type of nursing school hardly seems to be a realistic basis for decisions regarding the preparation and qualifications appropriate to nurse tutors in the future. (p. 5)

Lancaster stated that "the answer to this question depends on the definition of the tutor's role" (p. 5). She has suggested that the role of the teacher depends on many things, such as the type of students, the type of subjects to be taught, the attitudes and values of the educational system, and the teaching-learning environment. Furthermore, she has concluded that the preparation of teachers is crucial in the development and maintenance of effective nursing practice. But, the role of the teacher must first be explicitly defined so that criteria may be developed to assess the teacher training programs.

A quarter century later, Infante (1975) and Miller (1976) expressed
the very same concerns about the shortcomings in nursing education as McManus
(1949). Commenting on the preparation of nurse educators, Sister Miller (1976)
stated that "most graduates of master's programs become teachers immediately after graduation, but has their education prepared them for this role?" (p. 752). She found out that only 21 of the 64 graduate nursing programs offer preparation in teaching.

Infante (1975) stated emphatically that nurse educators should be educated in strategies of clinical teaching:

Teachers in graduate schools who are preparing teachers for undergraduate programs in nursing might place more emphasis on strategies of teaching in the clinical laboratory. Strategies for teaching in the clinical laboratory, the heart of professional education, have been noticeably absent from course outlines. Perhaps increased opportunity for future teachers to apply theory to practice under guidance is also needed. Just as changes in the service of a professional practicum must start with the educational program of which he is a product, so changes in the type of teaching done by teachers . . . must start with changes in the educational programs that prepare these teachers. (p. 60)

Similarly, Miller (1976) stated:
The best method to help [graduate] students to internalize the principles of learning, planning, and evaluation, along with the philosophy of education and the act of teaching, so that they may function effectively as teachers of nursing . . . is for those teaching graduate students . . . to be role models who facilitate the learning process for the student. (p. 752)

This lack of educational programs in clinical teaching and evaluation has made nurse educators vulnerable, and has encouraged the adoption of evaluation tools and practices to which they themselves were subjected as students. These evaluation practices, although not educationally sound, could prove to be very efficient and effective. However, unless clinical evaluation research endorses such practices, nurse educators must not assume that clinical grades reflect the students' true performance. This viewpoint has been asserted
by Champney, Chatfield, and Sims (1975) who stated that it is dangerous to adopt scales and objectives unless the team of instructors who have to use them are in complete agreement. Tate (1962) stated, "the more complex the behavior to be evaluated, the more difficult the task of building an evaluation instrument" (p. 36).

Another area which has influenced clinical performance grades of nursing students is team teaching. Various approaches to this technique were described by Schweer and Gebbie (1976) as having (1) all clinical teachers work together as a team in supervising the student, (2) certain portions of the program taught by a master clinical teacher, with other clinical teachers supervising groups of students, and (3) specialists in the field of nursing contribute to the clinical teaching and evaluation. These various approaches seem to have created inconsistencies among evaluators of clinical performance, because of (1) lack of communication among clinical instructors, (2) lack of training of these instructors relative to evaluation and methodology, and (3) lack of adequate time for instructors in the team to discuss clinical objectives, and the behaviors that the student is expected to demonstrate to successfully meet the objectives.

Hayter (1973) demonstrated that the effects of the difficulties created by team teaching, although they cannot be avoided entirely, could be diminished if each instructor in the team understood the objectives and the specific behaviors expected of a given level of student. She found that the agreement among 31 evaluators as to the grades of three students was $76 \%$ after discussing the objectives, as opposed to $44 \%$ before the objectives were discussed.

In her extensive review of the literature, Hart (1974) found that nursing schools have been concerned with the evaluation of clinical performance of students, and in some instances, have employed ameliorative measures. These measures include involving the faculty in the development of a devised rating scale or check list, the employment of student evaluations, and the identification of behaviors in the clinical setting which are regarded as satisfactory for a given level of student in the nursing curriculum.

Finally, the issue of assigning either a letter or a numerical value for clinical performance has perturbed many nurse educators. Rines (1963) wrote on this issue:

> Human behavior of any description is much too complex to permit fine discriminations . . . we do not know enough about the way in which students learn or what influences learning to say more that [sic] a behavior is acceptable or unacceptable. (p.17)

Hart (1974), in a survey of 105 nurse educators to determine their perception of the significance of criteria and tools employed in the evaluation of clinical performance of baccalaureate students, found that some of the least significant criteria and tools were the use of a devised rating scale and the percentage/letter grade for the course. Historically, says Hart, the letter/ percentage grade was widely used in the evaluation of the students' performance. However, she concluded that educators in baccalaureate programs are cognizant of the many variables in the evaluation of the students' clinical performance which cannot be reduced to a single variable representing a percentage or a letter grade. This finding supports the statement by Rines (1963), that clinical performance is too complex to permit fine discrimination.

It is interesting to note the discrepancy between the ameliorative measure--involving the faculty in the development of a devised rating scale or check list as stated by Hart--and her finding of a devised rating scale as the least significant criterion and tool.

Observational Techniques in Evaluation

Some aspects of learning and development can be evaluated only by direct observation (Ahmann \& Glock, 1971; Gronlund, 1976); however, evaluations based on observations have serious limitations.

Thorndike and Hagen (1961) have stated that "Evaluation based on naturally occurring situations of life is likely to vary from person to person" (p. 17). Clinical nurse educators continue to receive numerous complaints from the students about discrepancies between instructors or between clinical settings in clinical performance evaluations and grades assigned. To minimize these discrepancies Schweer and Gebbie (1976) concluded that "observation continues to be regarded as an effective means of evaluating nursing performance, but teachers must have a thorough knowledge of the exact behaviors to be accomplished during a given experience, along with an analysis of the level of each performance to be rated according to a standardized scale" (p. 172).

Other problems in performance evaluation have been discussed by Champney, Chatfield, and Sims (1975). These authors have suggested that "raters do better if carefully trained with respect to the distribution of abilities and the nature of the scale, and cautioned against common errors" (p. 168).

Four rater idiosyncrasies identified by them were: (1) leniency; (2) halo effect; (3) logical error; and (4) central tendency (p. 168).

The limitations of direct observation as identified by Ahmann and Glock (1971), Fuerst and Wolff (1969), Gronlund (1976), Rines (1963), and Thorndike and Hagen (1961) were as follows:

1. the difficulty in determining the significance of an isolated item of behavior;
2. the complexity involved in identifying the meaningful and productive set of behavior categories to observe;
3. the problems arising from the biases of the observer, who may bring his own needs and feelings to the evaluation;
4. the decrease of cogency in observations which arises as the student numbers increase; and
5. the high cost of direct observations.

## Anecdotal Record Technique of Evaluation

The anecdotal record technique is an informal, direct method of recording behavior observed in a natural setting (Thorndike \& Hagen, 1961), which provides a large amount of information about the learning and development of the student. Gronlund (1976) stated that "Anecdotal records are factual descriptions of the meaningful incidents and events which the teacher has observed in the lives of his pupils" (p. 429). Similarly, Fivars and Gosnell (1966), and Fuerst and Wolff (1969) have described anecdotal records as informal descriptions of the actual observed behavior and activities of students.

The advantages of such anecdotal records, described by Ahmann and Glock (1971), Gronlund (1976), Rines (1963), and Thorndik and Hagen (1961), are as follows:

1. lend themselves well to frequent observations of short duration;
2. make it possible to see trends in behavior over a period of time;
3. provide a description of actual behavior in natural situation;
4. provide a check on other evaluation methods;
5. make it possible to gather evidence on events that are exceptional but significant;
6. increase the observer's awareness; and
7. make instructors more diligent in observations.

Like any evaluation tool, the anecdotal record technique has its disadvantages, which are similar to the limitations of direct observational techniques used to evaluate behavior in naturally occurring situations of life (Ahmann \& Glock, 1971; Gronlund, 1976; Rines, 1963; and Thorndike \& Hagen, 1961).

The disadvantages include:

1. the time required to maintain the records;
2. the difficulty in being objective when observing and reporting student behavior;
3. the difficulty in obtaining an adequate sample of behavior;
4. the tendency to accumulate unfavorable anecdotes;
5. the difficulty of periodically summarizing the anecdotes into some kind of framework;
6. the tendency to report behavior without describing the setting;
7. the tendency to evaluate rather than report behavior; and
8. the tendency to describe behavior in general terms.

The suggestions made by Gronlund (1976) and Thorndike and Hagen (1961) for writing proper anecdotal records are:

1. determine in advance what you want to observe, but be alert for unusual behavior;
2. observe and record enough of the situation to make the behavior meaningful;
3. make a record of the incident as soon after the observation as possible;
4. limit each anecdote to a brief description of a single specific incident;
5. keep the factual description of the incident and the interpretation of it separate;
6. record both positive and negative behaviors;
7. collect a number of anecdotes on a student before drawing inferences concerning typical behavior;
8. obtain practice in writing anecdotal records;
9. teach the observers what to look for; and
10. train the observers in standards of judgment.

In summary, anecdotal records provide a systematic procedure for recording observation, but Blood and Budd (1972) state that it does not guarantee that these observations will be systematically made nor directed toward specific relevant behaviors. Two authors, Blood and Budd (1972) and Rines (1963) have suggested that perhaps time sampling systems might provide an opportunity to systematically observe student behavior.

## Prediction of Academic Performance

Versus Clinical Performance

## The Criterion Problem

In a review of research studies and practices dealing with selection, recruitment and prediction, Taylor et al. (1963) found that when academic grades in nursing were used as the criterion, the best predictors--in terms of single highest correlations--were pre-nursing College Grade Point Average (.63), the Otis (.48), and High School Grade Point Average (.43) (p. 33). Another review of research studies by the same authors in 1965, revealed that the best predictors were Otis (.65), American College Test (ACT) total score (.64), and High School Grade Point Average (.51).

In 1963, Taylor et al., when using clinical grades as the criterion, found that

The three best predictors with highest single correlations again are pre-nursing College Grade Point Average (.59), the Otis (.59), and the High School Grade Point Average (.37). In general, the correlations drop from academic or theory grades to clinical or practice grades indicating that the best predictors of the former criterion are not necessarily the best predictors of the latter criterion and that we have not yet learned how to predict the latter criterion. (p. 33)

In their discussion of the tremendous variability in the results with various predictors and criteria for success in nursing schools, Taylor et al. (1963) suggested that variability in results can in part be attributed to the many different ways of measuring the criteria from one study or location to the next. Furthermore, they indicated that the prediction problem may be confounded by the many complex
behaviors that are neither recognized in training programs nor in job evaluations as crucial for success at this point in time. These authors state that much more work is needed on the measurement problems involved in developing meaningful, reliable and valid clinical criterion measures for success in nursing schools, and that these criteria once developed should then be carefully related to "on-the-job" criteria for success.

## Predictive Studies

There are very few reported predictive studies on clinical performance in nursing. Ryden (1977) studied the predictive value of a clinical examination of interpersonal relationship skills and subsequent interpersonal competence in clinical situations as measured by ratings of clinical instructors, and obtained a coefficient of correlation. 45 (significant at $p \leq .05$ ). Although this is a positive relationship, it is not high, and therefore suggests that there perhaps are other factors which influence performance in the clinical area.

In their study of "Measurement and Prediction of Nursing Performance", Taylor et al. (1965) found that the graded performance of students in nursing theory and general education courses was not closely related to their performance in clinical nursing courses. These authors have suggested that perhaps the clinical grades represent measures of additional abilities and behaviors from those measured by academic courses.

The shortcomings in nursing education described earlier by Taylor et al. (1965, pp. 38-39) may indicate one reason for nursing theory and clinical practice grades not correlating highly. They have suggested that the lack of relationship
between performance in theory and clinical practice courses may point out the undesirable utilization of the deficient traditional methods of evaluating students, especially because of the lack of relationship between classroom and clinical grades.

These authors have postulated that this lack of correlation between theory and practice has occurred because:

1. theory courses have not stressed basic nursing processes which are relevant across patient problems, service settings, and functions;
2. instead of being expected to assume responsibility for acquiring the information needed to deal with a particular problem, students were given the answers;
3. until very recently these courses have been built around clinical areas and/or disease entities rather than patient problems;
4. lack of specificity in criteria by which students have been evaluated as they practice nursing in the clinical units;
5. objectives have not been stated in terms of behaviors that could be easily evaluated; and
6. learning how to practice as a professional is a much more complex learning process than learning of a theory per se or the learning of a patient care procedure. (p. 39)

In a study of prediction of success of community college nursing students, Powers (1974) studied the interpersonal competencies of 80 students in their second semester using selected items from the Minnesota Multiphasic Personality Inventory (MMPI), the California Psychological Inventory (CPI), the Strong Vocation Interest Blank (SVIB), and the American College Test (ACT). The results indicated that only two subtests from the four predictor variables produced $F$ values significant at .05 level when correlated with interpersonal competencies of nursing students.

These items were "Social Introversion" from the MMPI, which had a multiple correlation coefficient of . 1021; and "Good Impression" from the CPI which increased the multiple correlation coefficient to .2182 (p. 56).

A review of medical literature revealed results similar to those found in nursing--that is, low correlations between clinical performance grades and academic grades. Korman and Stubblefield (1971) administered a rating scale assessing internship performance and then correlated the results with medical school grades. They found that GPA bore no significant relationship to any of the evaluative components of internship performance.

Similarly, Wiener, Koran, Mitchell, Schattner, Fierstein, and Hotchkiss (1976) correlated the scores of medical students for interviewing, physical examinations, and total clinical skills with the scores on the final multiple-choice exam given at the end of clerkship. The respective product moment correlations were $.36, .58$, and .58 . Using a broader sampling of physician performance, Gonnella, Goran, Williamson, and Cotsonas, N., Jr. (1970) and Leaper, Gill, Staniland, Horrocks, and De Dombal (1973) found little relationship between written examination scores and clinical performance scores.

These findings suggest that there is a weak relationship between students' knowledge and the application of that knowledge. However, these investigators did not take into account all the other factors which affect evaluation of clinical performance.

The review of the literature did not reveal any information relevant to the alleged relationship of complexity of the nursing situation and poor clinical performance. The strong feelings of students, staff, and instructors on this issue need attention.

## Summary

Research has been sporadic and meager on what constitutes competent clinical performance, on approaches to grading it, on tools to measure it, and on corresponding issues of validity and objectivity. In the few studies which have been completed, findings have been discouraging--correlations between clinical performance grades and academic or psychological tests have ranged from 0 to .65. Furthermore, the correlations have varied from one institution to another and from one sample to another, making it difficult to generalize to the population of nursing students at large.

Some of the numerous reasons given by various researchers for these low correlations are: more subjectivity is involved in assigning clinical grades than academic grades; components of satisfactory clinical performance have not yet been identified; training of nurse educators is inadequate in the case of clinical teaching-learning-evaluation processes; there are substantial differences in the grading methods used; there is real variability in the quality of students; there are substantial differences in the curricula; many problems are inherent in the use of observational techniques in the evaluation and the grading of
clinical performance; there are no standardized evaluation procedures; problems of communication and consistency among instructors are inherent in team teaching; and traditional methods of clinical teaching and evaluation are outmoded.

Only a few of the studies indicated clearly whether the clinical performance grade did indeed comprise the clinical grade or was just one component, along with theory, of the clinical grade. Likewise, very few studies attempted to ascertain the reliability of the so-called clinical grades before studying variables which would predict it.

The limited and controversial research on clinical performance evaluation --and specifically the weak relationship between knowledge and its application in real life situations--requires vigorous efforts to identify influential factors and to analyze how grading processes could be improved so as to benefit the individual nurse, her patient, and the nursing profession.

Chapter III

## RESEARCH METHODOLOGY

Introduction

In order to assess the extent to which first semester grades in Nursing theory, Psychology, and Biology courses, and complexity of the nursing situation were valid predictors of clinical performance of nursing students, a sample of associate degree nursing students was selected. Likewise, a sample of clinical instructors was selected.

## Sampling

Institution
Requests to conduct the study at Vancouver Community College, Langara, British Columbia were made to the Dean of Administrative and Student Services (Appendix A), and to the chairperson of the Nursing Department. Each was approached personally and given the research proposal. Written permission to proceed was subsequently received from the Dean of Administrative and Student Services (Appendix B), followed by oral permission from the Nursing Department chairperson.

The two-year Nursing Program at Vancouver Community College, Langara is divided into six semesters, each 13 to 14 weeks long. Because students in Semesters I and VI had, respectively, very short and very long clinical experiences, they were deleted from the study; thus, students in the remaining semesters were included. These groups had clinical experience throughout the semester ranging from 8 to 16 hours per week.

## Weeks

Each semester was comprised of two rotations, each six to seven weeks in length. The grading system used by the nursing faculty was such that at the mid-term (end of first rotation), the students received a formative evaluation, and at the end of the semester they received a final clinical grade. Only the anecdotal records written during the first rotation were used in order to eliminate the possible influence of this research on students' grades at the end of the semester.

The clinical instructors were asked to write anecdotes about the performance of the ir students for the fourth, fifth, and sixth weeks of the first rotation, which extended from January 30 to February 13, 1978. These weeks were chosen because (1) the first week of the rotation was orientation, (2) the second week would give the students a chance to establish themselves and allow them to begin to feel comfortable in the clinical area, and (3) the third week would give the instructors a chance to practise writing anecdotal records. Furthermore, the anecdotal records of only three weeks were chosen to keep the task of writing and scoring them reasonable and manageable for the nursing instructors.

## Instructors

Approximately two months prior to the initiation of the study, permission was requested from the Nursing Department chairperson to present the highlights of the study to faculty members in a general meeting. The purposes of this presentation were to ascertain whether interest in the study was sufficient, and to answer questions. The investigator (one of the instructors who was not included in the study) met with the instructors. The proposed study was explained, questions were answered, and feedback was obtained indicating that the majority of faculty were in favor of the study and interested in its results.

During the week of January 9, 1978, each instructor from Semesters II, III, IV, and V, was given a sheet explaining the study (Appendix C) and a consent form (Appendix D); they were asked to return the consent form to the investigator by January 13, 1978. Some of the instructors asked for additional information or clarification, which was provided as requested.

There were four instructors in Semester II, three of whom consented to participate in the study. Similarly, there were four instructors in Semester III, three of whom consented to participate in the study. One of these instructors consented only to write anecdotes and judge the performance of her own students. Another instructor who was not teaching in Semester III but had taught in it several times, consented to be a judge. Semester IV had four instructors, all of whom consented to participate fully in the study.

Finally, Semester $V$ had five instructors, all of whom consented to participate. Also, another instructor who had taught in that semester and was at that time teaching in Semester 1, wanted to participate, and consented to be a judge in Semester $V$.

In summary, of 17 instructors in Semesters II, III, IV, and V, 15 consented to write anecdotal records and be judges. Two other instructors who were not teaching in those semesters but had taught in them before, consented to be judges.

Students
After obtaining the signed consents from the instructors indicating their willingness to participate in the study, the investigator asked the leaders of Semesters II, III, IV, and $V$ respectively for 15 to 20 minutes of their class time during the week of January 16, 1978. In these sessions, the investigator explained the study to the students, and distributed both the explanation (Appendix E) and the consent forms (Appendix F). Students were asked to complete the consent form and return it to the investigator by the end of that week. Furthermore, they were told that if they had any questions regarding any part of the study to contact the investigator.

Semester II sample. There were 35 students registered in the clinical nursing course, 30 of whom agreed to participate. The group of students who consented to participate, but whose instructor wished not to participate, were eliminated from the study. Also, to equalize the workload of the three
participating instructors, five students were randomly chosen from each group. And of these 15 students, one withdrew from the semester, leaving a final sample of 14 students.

Semester III sample. There were 25 students registered in Semester III clinical course, 21 of whom consented to participate. Those students who consented to participate, but whose instructor wished not to participate, were eliminated from the study. Of the remaining 14 students, two were in a specialized clinical area and, therefore, were eliminated from the study. The final sample of Semester III was comprised of 12 students.

Semester IV sample. There were 32 students registered in the clinical nursing course of Semester IV, 24 of whom agreed to participate. To equalize the workload of the four instructors, four students were randomly chosen for each instructor, yielding a final sample of 16 students.

Semester $V$ sample. Finally, Semester $V$ had 28 students registered in the clinical nursing course, 22 of whom agreed to participate. Two of these students had advanced credits in the independent variables from countries other than Canada and United States. However, the assessed transcripts did not indicate the equivalent college grades of these students. Consequently, it was decided to eliminate them from the study. Also, three students withdrew from the semester. The final sample was comprised of 17 students.

In summary, the total sample of students at the end of data collection period was 59: 14 from Semester II, 12 from Semester III, 16 from Semester IV, and 17 from Semester $V$.

## Confidentiality

To ensure confidentiality, each instructor and student was assigned a number by the investigator. Each anecdotal note was typed before being given to the judges to eliminate effects of handwriting and the identities of instructor and student. Also, the investigator proofread all the anecdotal notes and, where the need arose, names of students were deleted and substituted with appropriate terms such as, "the student", "she", "he", et cetera.

## Instrumentation

## Complexity Scale

The decision as to who could best determine the complexity of the nursing situation was complicated by the many constraints on time and resources. It was felt that complexity could be most reliably and objectively assessed by a group of observers trained to follow up each student and record minute-to-minute changes in their patients' health status and to observe the students' abilities. A panel of these observers could then judge the complexity of the nursing situation as recorded. However, considering the constraints mentioned above, the investigator believed that the best persons to determine the complexity of the students' assignments were their clinical instructors, because they were in a position to be most knowledgeable about the students' assignments and the factors affecting them.

A form was then developed (Appendix G) to be used by the clinical instructors for rating the complexity of the nursing situation. This form delineated
three levels of complexity, and the instructor was asked to circle one level at the end of each clinical experience. Each of the three levels of complexity was respectively described by four criteria. These were (1) activities of the student; (2) condition of the patient; (3) patient experiences; and (4) revisions of student's assignments based on the instructor's assessments of the student's abilities/performance. Furthermore, the form was accompanied by guidelines for its utilization--these were specific examples which the instructor had to follow to decide on the level of complexity.

Definitions of the three levels of complexity were based on clinical experience and on numerous discussions with both instructors and students about the difficulty of clinical assignments. Resources were not sufficient to allow a complete pilot test of the instrument. However, four instructors were consulted; each read it independently and reacted, and the form was refined accordingly.

## Anecdotal Record

The anecdotal record form (Appendix H) was adapted from Rines (1963) following a review of the literature and the assessment of the needs of this study. The following elements were noted: identification data (student and instructor numbers, date and semester); description of the situation; description of the student's behavior; and the performance score given.

Scoring system. The purpose of the scoring system was to provide the judges with a uniform measuring device, thereby minimizing scorer subjectivity and increasing reliability of the scores given.

The scoring system (Appendix 1) developed by the investigator is a Likert-type scale with a range of 1 to 9, representing unsatisfactory performance to superior performance. This scale was further divided into five major categories which were more familiar to the instructors so that there would be an association between the scoring system and the grading system used at the college. This assisted in utilizing the scoring procedure. The five categories were $A, B, C$, P, and F. Each of these grading categories was further described by a set of six criteria which were thought to be critical in the evaluation of the students' attainment of each of the clinical objectives. Nebulous terminology was defined, so as to provide the judges with a common understanding of the crucial terms and thereby increase objectivity and reliability of the scores.

## Procedure

Development of the Videotape
In order to assist clinical instructors in the writing of anecdotal records and the use of these records to judge the clinical performance of the students, training procedures were developed to sharpen observational skills, and to allow practice in the above activities in a simulated setting.

The investigator wrote a script for a nine-minute videotape, titled "Evaluation of a Student's Clinical Performance", which has five sections designed to depict the various activities of a nursing student in the clinical laboratory.

The five vignettes and the ir respective lengths were:

Vignette

1. The student doing a brief assessment of the patient before breakfast.
2. The student giving subcutaneous injection to patient.
3. The student reporting to the team leader.
4. The student and the instructor discussing the student's nursing care plan on the patient.
5. The instructor checking the student's charting on one of the patients.

Length
$2 \frac{1}{2} \mathrm{~min}$.
$2 \frac{1}{2} \mathrm{~min}$. $\frac{1}{2} \min$. $2 \min$. $1 \frac{1}{2} \mathrm{~min}$ 。

## The Training of Clinical Instructors and Judges

Separate workshops were scheduled for the instructors of each of the Semesters II, III, IV, and V. The instructors who had consented to participate in the study were notified in writing to attend these workshops on the evaluation of clinical performance. Other instructors who had indicated an interest in the topic were also invited to attend.

The purpose of the workshops was to teach instructors how to (l) write anecdotal records according to the sample description of anecdotal records adapted from Rines (1963); (2) score these, employing the scoring system (Appendix I); and (3) determine the complexity of the nursing situation.

Each workshop was conducted over a period of three to four hours, during which time the following activities were carried out: (1) reviewing the advantages, disadvantages, and characteristics of effective anecdotal records;
(2) viewing each of the five sections of the videotape, followed by writing and discussing the performance of that student based on the clinical objectives and the assignment; (3) scoring this student's performance using the scoring system; and (4) determining the complexity level of the student's assignment.

## Data Collection

The instructors wrote anecdotal records for the last three weeks of the first rotation of their semesters. They were asked to write three anecdotes on each student, one anecdote for each week of clinical experience. The written anecdotal records and the complexity of the nursing situation were collected weekly from the clinical instructors. These anecdotal records were then typed, duplicated, and distributed to three judges--the clinical instructor and two other instructors--who scored them independently using the scoring system (Appendix I), and the clinical objectives of the student. These scores were then analyzed to determine their reliability using conventional correlational and generalizability analyses.

During the week of January 16, 1978, the investigator obtained from the registrar's office the participating students' grades in their first semester Nursing theory, Biology, and Psychology courses. High school grade point averages (GPA's) were not used because the college did not require that information from applicants. These students were admitted into the nursing program on the condition of high school graduation and/or their ability to successfully complete appropriate "make up" courses.

## Method of Analysis

The analyses of the data were executed in two phases. First, it was necessary to assess the reliability of the clinical scores before proceeding to the stepwise multiple regression analyses. The following is a description of the two phases of the data analyses.

## Phase One

During phase one, preliminary analyses of the data were conducted to determine (1) inter-rater reliability by correlating the scores given by the clinical instructors and the two nursing judges; and (2) generalizability coefficients among judges. The scores for these analyses which ranged from 1 to 9 were derived as shown in Appendix I. The latter was believed to be a more sensitive and informative measure of the reliability of the clinical scores. Thus, both methods of analyses were carried out and the results compared.

A program from the Statistical Package for the Social Sciences (SPSS.V701) was used to compute the Pearson Product Moment Correlation Coefficients for the data of each of the four semesters separately and within each of the four semesters, each week separately.

Each correlation coefficient ( $r$ ) in turn was converted to Fisher's $Z$ employing an $r$-to- $Z$ transformation table. Next the $Z$ values obtained for each of the three weeks of each of the fours semesters were added up separately and then the mean $Z$ values were calculated. The mean $Z$ values obtained for each of the weeks were afterwards converted to Pearson Product Moment Correlation Coefficients utilizing Fisher's procedure.

Then, the mean $Z$ values for the three weeks of each of the semesters were added up and the total semester mean $Z$ value was calculated and subsequently converted to Pearson Product Moment Correlation Coefficients employing the r-to-Z transformation procedure. Finally, the significance of the $r$ values of each of the Semesters II to $V$, and each of the three weeks of those semesters were assessed against an $\alpha$ level of. 05 .

The UCLA Biomedical Program BMD08V (Appendix J) was used to compute the generalizability coefficients across the nursing judges and weeks. These analyses were carried out with the data from Semesters II, III, and part of V. The fact that, within Semesters IV and $V$, the same three judges did not score the anecdotal records of all corresponding students--all Semester IV students and some Semester V students--resulted in a not fully crossed design. Therefore, for these two semesters, it was not possible to complete the generalizability analyses.

## Phase Two

Before conducting the regression analyses called for in phase two, it was first necessary to examine the potential influence of judges. This was completed by comparing the regression results using observed ratings with the regression results obtained when scores were standardized within judges. Regression analyses on a sample of the data were compared, utilizing standardized scores in one analysis and raw clinical scores in another. The clinical scores of Week 3 of both Semesters II and III were then standardized across the three judges following the procedure elaborated in Appendix K.

The significance level for the regression analyses was set at $\alpha=.10$ at each step (Green \& Tull, 1965, p. 73). Given the exploratory nature of the study
and the available small sample sizes, a more relaxed level of significance was adopted so as to ensure a reasonable power. The independent variables--first semester grades in Nursing theory, Biology, and Psychology--were quantified by equating the letter grades to numerical grade point equivalents according to the college calendar (refer to operational definitions in Chapter I). Upon inspection of the scores on complexity of the nursing situation, it was noted that they were mainly level 2 across all the students throughout the three weeks. It was subsequently decided to delete this variable because it would not yield enough variance to warrant its addition to the multiple regression equation. Finally, using the stepwise multiple regression program (Appendix L), the criterion scores--mean pooled raw clinical scores and mean standardized clinical scores--were analyzed and their prediction equations were compared to determine whether there was a significant difference between standardized scores and raw scores. Results of this analysis revealed that there were no significant differences between the prediction equations using standardized or raw clinical scores. Consequently, the rest of the data were analyzed using the mean pooled raw clinical scores and employing the linear model:

$$
\left[\frac{\mathrm{Y}_{\mathrm{SCl}}+\text { SNJ's (2) }}{3}=\beta_{0}+\beta_{1} \text { Nursing }+\beta_{2} \text { Biology }+\beta_{3} \text { Psychology }+\varepsilon\right]
$$

The stepwise multiple regression program was used, and the clinical scores from each of the three weeks, and from the three weeks combined, were analyzed for each of Semesters II, III, IV, and V. 1

[^1]
## Chapter IV

## RESULTS OF ANALYSIS

In this chapter are presented the results of the analyses on the complexity of the nursing situation; the correlations of the inter-rater reliabilities and generalizability coefficients; and the correlations among the clinical performance and Nursing, Biology, and Psychology grades.

## Complexity of the Nursing Situation

The independent variable, complexity of the nursing situation; was deleted from the regression analyses of Semesters II, III, IV, and V because complexity scores within each of the three weeks, and between the three weeks, revealed that there was essentially no difference in the complexity of the nursing situation in which these students were asked to work. Possible reasons for this are:
(a) the tendency of clinical instructors to perceive that the students' assignments were almost always within their abilities;
(b) the phenomenon of central tendency. Because the instrument had only three levels of complexity, it may have encouraged instructors to avoid extreme ratings in assessing the complexity level of their students' situations;
(c) although the levels of complexity were not labeled as such, it appeared that the clinical instructors unconsciously thought of level 2 representing the desired and ideal complexity of students' assignments.

The consequence of having to delete the variable, complexity of the nursing situation, limits the study to only academic predictors. Furthermore, it prevents an examination of possible relationships or interactions that may have existed between the academic predictors and complexity of nursing situation.
Inter-Rater Reliabilities of Clinical Scores

## Correlational Analyses

Semester II. Table 1 shows the mean and the standard deviation for each of the three judges, and the correlations between the judges for each of the three weeks. Also, it shows the mean correlations of each week and the three weeks combined. The mean correlation for Semester II was . 75.

Semester III. Table 2 gives the correlational results for Semester III. The overall semester correlation was .58 .

Semester IV. Table 3 gives the correlational results for Semester IV. In Weeks 2 and 3, the r's between Judges 2 and 4 could not be computed because there was no variance in the scores given by Judge 2. The overall semester correlation was. 84.

Semester V. Table 4 gives the correlational results for Semester $V$. The overall semester correlation was . 73 .

TABLE 1
MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS AMONG NURSING JUDGES (J) IN SEMESTER II

| Week | Means |  |  | SD |  |  | Correlations between Judges ${ }^{\text {a }}$ |  |  | Mean |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{J}_{1}$ | $\mathrm{J}_{2}$ | $\mathrm{J}_{3}$ | $\mathrm{J}_{1}$ | $\mathrm{J}_{2}$ | $\mathrm{J}_{3}$ | $\bar{J} \mathrm{~J}_{2}$ | $\mathrm{J}_{1} \mathrm{~J}_{3}$ | $\mathrm{J}_{2} \mathrm{~J}_{3}$ |  |
| 1 | 5.00 | 5.29 | 4.50 | 0.88 | 0.91 | 0.92 | . 58 | . 72 | . 82 | . 72 |
| 2 | 4.75 | 4.93 | 4.32 | 0.83 | 1.39 | 0.85 | . 62 | . 48 | . 88 | . 70 |
| 3 | 4.61 | 4.79 | 4.29 | 1.13 | 1.31 | 0.83 | . 80 | . 79 | . 84 | . 81 |
| Combin |  |  |  |  |  |  |  |  |  | . 75 |

${ }^{\text {a }}$ Correlations are based on sample size of 14 .
Mean correlations were obtained by converting r 's between judges to Fisher $\mathrm{Z}^{1} \mathrm{~s}$, averaging these, and reconverting them to correlation coefficients.

## TABLE 2

MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS AMONG NURSING JUDGES (J) IN SEMESTER III

| Week | Means |  |  | SD |  |  | Correlations between Judges ${ }^{\text {a }}$ |  |  | Mean/ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | ${ }_{1}$ | $\mathrm{J}_{2}$ | $\mathrm{J}_{3}$ | $\mathrm{J}_{1}$ | $\mathrm{J}_{2}$ | $J_{3}$ | $\overline{J_{1} J_{2} / N}$ | $\mathrm{J}_{1} \mathrm{~J}_{3} / \mathrm{N}$ | $\mathrm{J}_{2} \mathrm{~J}_{3} / \mathrm{N}$ |  |
| 1 | 5.96 | 5.82 | 6.41 | 1.35 | 0.75 | 1.14 | $.53 / 11$ | .49/11 | . $57 / 11$ | .53/11 |
| 2 | 5.75 | 5.63 | 5.54 | 1.06 | 0.93 | 1.03 | $.66 / 12$ | $.55 / 12$ | $.28 / 12$ | $.51 / 12$ |
| 3 | 5.75 | 5.50 | 5.38 | 1.36 | 0.71 | 0.88 | $.88 / 12$ | $.62 / 12$ | $.55 / 12$ | $.72 / 12$ |
| Combined |  |  |  |  |  |  |  |  |  | . $58 / 12$ |

${ }^{a}$ The numbers $/ N$ indicate sample size.
Mean correlations were obtained by converting r's between judges to Fisher $Z$ 's, averaging these, and reconverting them to correlation coefficients.

## TABLE 3

MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS AMONG NURSING JUDGES (J) IN SEMESTER IV

| Week | Means |  |  |  | SD |  |  |  | Correlations between Judges ${ }^{\text {a }}$ |  |  |  |  |  | Mean/ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{J}_{1}$ | $\mathrm{J}_{2}$ | $\mathrm{J}_{3}$ | $\mathrm{J}_{4}$ | ${ }^{1}$ | $\mathrm{J}_{2}$ | $J_{3}$ | $\mathrm{J}_{4}$ | $J_{1} J_{2} / N$ | ${ }^{J_{1} J_{3} / \mathrm{N}}$ | $\mathrm{J}_{1} \mathrm{~J}_{4} / \mathrm{N}$ | $\mathrm{J}_{2} \mathrm{~J}_{3} / \mathrm{N}$ | $J_{2} J_{4} / \mathrm{N}$ | ${ }^{J_{3}{ }^{\frac{1}{4} / \mathrm{N}}}$ |  |
| 1 | 4.55 | 5.25 | 6.09 | 5.36 | 1.51 | 0.62 | 1.38 | 2.46 | . $63 / 8$ | .81/7 | .91/7 | .04/8 | .76/8 | .88/7 | .75/15 |
| 2 | 4.46 | 5.00 | 5.50 | 6.10 | 1.06 | 0.45 | 1.08 | 1.60 | . $92 / 8$ | .93/7 | . $97 / 7$ | .81/7 | */7 | $.59 / 6$ | .89/14 |
| 3 | 4.68 | 5.00 | 5.40 | 6.30 | 1.10 | 0.47 | 1.58 | 1.77 | . $81 / 7$ | .84/8 | .89/7 | . $92 / 6$ | */6 | .76/6 | .85/14 |
| Combine |  |  |  |  |  |  |  |  |  |  |  |  |  |  | .84/16 |

${ }^{\text {a }}$ The numbers $/ \mathrm{N}$ indicate sample size.
Mean r's were obtained by converting r's between judges to Fisher $Z$ ' $s$, averaging these, and reconverting them to correlation coefficients.
*Coefficient could not be computed because there was no variance in the scores given by Judge 2 for those students co-rated with Judge 4.

## TABLE 4

MEANS, STANDARD DEVIATIONS, AND INTERCORRELATIONS AMONG NURSING JUDGES (J) IN SEMESTER V

| Week | Means SD |  |  |  |  |  |  |  |  |  |  |  | Correlations between Judges ${ }^{\text {a }}$ |  |  |  |  |  |  |  |  |  | Meary |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{J}_{1}$ | $\mathrm{J}_{2}$ | $J_{3}$ | $\mathrm{J}_{4}$ | $J_{5}$ | $J_{6}$ | $J_{1}$ | $\mathrm{J}_{2}$ | $J_{3}$ | $\mathrm{J}_{4}$ | $J_{5}$ | $J_{6}$ | $J_{1} J_{3}$ | $J_{1} J_{4} / N$ | $\mathrm{J}_{1} \mathrm{~J}_{5} / \mathrm{N}$ | $J_{1} J_{6}$ | $\mathrm{J}_{2} \mathrm{~J}_{3} / \mathrm{N}$ | $J_{2} J_{4} / N$ | $\mathrm{J}_{2} \mathrm{~J}_{5} / \mathrm{N}$ | $\mathrm{J}_{2} \mathrm{~J}_{6} / \mathrm{N}$ | $J_{3} J_{6} / N$ | $J_{4} J_{5} / N$ |  |
| 1 | 5.38 | 5.67 | 5.85 | 5.00 | 4.93 | 5.05 | 0.79 | 1.80 | 1.11 | 1.16 | 1.57 | 1.38 | $.95 / 4$ | . $97 / 4$ | . $82 / 4$ | . $87 / 4$ | .69/6 | 1.00/3 | . $58 / 3$ | . $70 \%$ | .73/10 | . $32 / 7$ | . $87 / 17$ |
| 2 | 4.75 | 4.63 | 4.78 | 5.14 | 5.07 | 4.17 | 0.66 | 1.41 | 1.64 | 1.35 | 1.37 | 0.66 | .79/4 | . $13 / 4$ | . $07 / 4$ | . $49 / 4$ | . $57 / 5$ | .00/3 | -. $50 / 3$ | .85/5 | . $56 / 9$ | . $22 / 7$ | . $38 / 16$ |
| 3 | 5.13 | 5.44 | 5.65 | 5.14 | 5.29 | 5.00 | 0.64 | 1.24 | 1.25 | 1.07 | 1.73 | 0.75 | . $58 / 4$ | . $94 / 4$ | . $88 / 4$ | . $58 / 4$ | . $96 / 6$ | . $50 / 3$ | . $50 / 3$ | .83/6 | . $69 / 10$ | . $74 / 7$ | . $78 / 17$ |
| Combined |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  | .73/17 |

aThe numbers $/ \mathrm{N}$ indicate sample size.
Mean correlations were obtained by converting r's between judges to Fisher $Z^{\prime}$ 's, averaging these, and reconverting them to correlation coefficients.

In summary, the correlation coefficients obtained in Semesters II, III, IV, and $V$ show that the inter-rater reliabilities of ratings of clinical performance of nursing students were moderate to moderately strong, varying between . 58 and .84 .

## Generalizability Analyses

Semester II. Table 5 reports a summary of the generalizability analysis for Semester II. Estimates of the variance components for instructors (I), weeks (W), and weeks-by-judges (WJ) were negative. The negative estimates were substituted with zero values as recommended by Cronbach et al. (1972). Consequently, the expected variances--E $\sigma^{2}(x)$--of the above sources were set to zero.

In examining the remaining eight expected variances, it can be seen that the two sources of variance--judges and students $(S)$ were relatively large, and therefore contribute most to the universe score variance, and hence to the generalizability coefficients. Table 6 provides an example of how the universe score variance, the observed score variance, and the generalizability coefficients of both judges and weeks were computed. The generalizability coefficients (Ep ${ }^{2}$ ) for the judges and weeks were .75 and .79 , respectively, indicating that the clinical scores were stable both across judges and over time.

Semester III. Table 7 reports a summary of the ANOVA which supports the generalizability analysis for Semester III. Estimates of the variance components for instructors, weeks, judges, and instructors-by-weeks (IW) were negative, and therefore given a value of zero.

The largest variance source was the one by students, followed by the source students-by-weeks nested within instructor (SW:I). The generalizability

## TABLE 5

## ANALYSIS OF VARIANCE OF CLINICAL SCORES OF SEMESTER II STUDENTS: EFFECTS OF TIMES ON JUDGES

| Source <br> of <br> Variance | Sum <br> of <br> Squares | Degrees <br> of <br> Freedom | Mean <br> Square | Estimate of <br> Variance <br> Component | Frequency <br> Within | Eo² (x) |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Instructors <br> (I) | 4.39 | 2 | 2.20 | -0.087 | 3 | ${ }^{*} 0.0000$ |
| Weeks <br> (W) | 1.03 | 2 | 0.52 | -0.014 | 3 | $* 0.0000$ |
| Judges <br> (J) | 9.23 | 2 | 4.61 | 0.124 | 3 | 0.0413 |
| Students <br> S:I | 47.85 | 9 | 5.32 | 0.591 | 12 | 0.0492 |
| IW | 4.79 | 4 | 1.20 | 0.015 | 9 | 0.0016 |
| IJ | 11.93 | 4 | 2.98 | 0.236 | 9 | 0.0262 |
| WJ | 0.16 | 4 | 0.04 | -0.011 | 9 | $* 0.0000$ |
| SW:I | 18.40 | 18 | 1.02 | 0.341 | 36 | 0.0095 |
| SJ:I | 2.74 | 18 | 0.15 | 0.051 | 36 | 0.0014 |
| IWJ | 1.64 | 8 | 0.21 | 0.009 | 27 | 0.0003 |
| SWJ:I | 6.14 | 36 | 0.17 | 0.171 | 108 | 0.0016 |

[^2]$E \sigma^{2}(x)$ means expected variance of each source.
S:I means students nested within instructor.

TABLE 6

## VARIANCE COMPONENTS CONTRIBUTING TO UNIVERSE SCORE VARIANCE AND EXPECTED OBSERVED SCORE VARIANCE OF JUDGES AND WEEKS, AND GENERALIZABILITY

 COEFFICIENTS OF SEMESTER II| Source <br> of <br> Variance | Universe Score <br> Variance | Expected Observed Score <br> Variance | Generalizability <br> Coefficient |
| :---: | :---: | :---: | :---: |

Judges 0.0413
0.0413

Students:I
0.0492
0.0492

U IJ
0.0262

WJ
0.0000

D SJ:I
0.0014

IWJ
0.0003

G SWJ:I
0.0016

E Total
0.0905
0.1200

S

| Weeks | 0.0000 | 0.0000 |
| :--- | :--- | :--- |
| Students:I | 0.0492 | 0.0492 |
| IW |  | 0.0016 |
| WI |  | 0.0000 |
| SW:I |  | 0.0095 |
| IWJ |  | 0.0003 |
| SWJ:I |  | 0.0016 |
|  |  | 0.0622 |

$$
\frac{0.0905}{0.1200}=0.75
$$

Weeks
0.0000
0.0000

E
IW
0.0016
0.0000

E SW:I
0.0003

K SWJ:I
0.0492
0.0622

$$
\frac{0.0492}{0.0622}=0.79
$$

| Key: | Week $=$ W | Student $=$ S |
| :--- | :--- | :--- |
|  | Judge $=$ J | Student nested within Instructor $=S: 1$ |
| Instructor $=1$ |  |  |

Note: Above terminology is elaborated in Appendix J.

TABLE 7

## ANALYSIS OF VARIANCE OF CLINICAL SCORES OF SEMESTER III STUDENTS: EFFECTS OF TIMES ON JUDGES

| Source | Sum <br> of <br> Variance <br> Squares | Degrees <br> of <br> Freedom | Mean <br> Square | Estimate of <br> Variance <br> Component | Frequency <br> Within | $E^{2}(x)$ |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Instructors <br> (I) | 0.67 | 1 | 0.67 | -0.127 | 2 | $*_{0} 0.0000$ |
| Weeks <br> (W) | 0.20 | 2 | 0.10 | -0.128 | 3 | ${ }^{*} 0.0000$ |
| Judges <br> (J) | 0.52 | 2 | 0.26 | -0.014 | 3 | $* 0.0000$ |
| Students |  |  |  |  |  |  |
| S:I | 75.25 | 10 | 7.53 | 0.836 | 12 | 0.0697 |
| IW | 2.84 | 2 | 1.42 | -0.183 | 6 | $* 0.0000$ |
| IJ | 5.41 | 2 | 2.71 | 0.109 | 6 | 0.0181 |
| WJ | 2.55 | 4 | 0.64 | 0.030 | 9 | 0.0033 |
| SW:I | 94.13 | 20 | 4.71 | 1.569 | 36 | 0.0436 |
| SJ:I | 14.91 | 20 | 0.75 | 0.249 | 36 | 0.0069 |
| IWJ | 1.36 | 4 | 0.34 | 0.010 | 18 | 0.0005 |
| SWJ:I | 11.26 | 40 | 0.28 | 0.282 | 108 | 0.0026 |

*Expected variances of negative estimates of variance components are set to zero.
$E \sigma^{2}(x)$ means expected variance of each source.
S:I means students nested within instructor.
coefficients for the judges and weeks were .69 and .58 , respectively, indicating that the clinical scores show more stability over judges than across time in Semester III. A possible reason for the latter could be explained by the predicted growth in students over time.

Semester V. Table 8 reports a summary of the ANOVA which supports the generalizability analysis for Semester $V$. Estimates of the variance components for instructors, instructors-by-weeks-by-judges (IWJ), instructors-by-iudges (IJ), and weeks-by-judges (WJ) were negative, and therefore given a value of zero.

Generalizability coefficients for judges and weeks were . 90 and . 62 respectively, indicating that the clinical scores show more stability over judges than across time in Semester V. Again, a possible reason for the latter could be explained by the predicted growth in the performance of these students over time.

## Correlation Coefficients Versus Generalizability Coefficients

Table 9 provides the generalizability $(G)$ for judges and the correlation coefficients of Semesters II, III, and V. Both the G coefficient and the correlation coefficient for Semester II were identical, .75. The G coefficient of Semester III, however, was higher than the correlation coefficient. This discrepancy of .11 is perhaps due to the fact that the $G$ coefficient is a more sensitive measure of reliability than the conventional correlational coefficient. Finally, the $G$ coefficient of Semester $V$ was .90 , larger than the correlation coefficient which was .75 . The discrepancy of .15 between the two coefficients of Semester $V$ was not only due to

## TABLE 8

## ANALYSIS OF VARIANCE OF CLINICAL SCORES OF SEMESTER V STUDENTS: EFFECTS OF TIMES ON JUDGES

| Source <br> of <br> Variance | Sum <br> of <br> Squares | Degrees <br> of <br> Freedom | Mean <br> Square | Estimate of <br> Variance <br> Component | Frequency <br> Within | $E^{2}(x)$ |
| :--- | :---: | :---: | :---: | :---: | :---: | :---: |
| Instructors <br> (I) | 8.96 | 1 | 8.96 | -0.032 | 2 | $* 0.0000$ |
| Weeks <br> (W) | 30.79 | 2 | 15.39 | 0.465 | 3 | 0.1548 |
| Judges <br> (J) | 6.04 | 2 | 3.02 | 0.131 | 3 | 0.0437 |
| Students |  |  |  |  |  |  |
| S:I | 39.35 | 4 | 9.84 | 1.093 | 6 | 0.1822 |
| IW | 20.90 | 2 | 10.45 | 0.380 | 6 | 0.0633 |
| IJ | 1.15 | 2 | 0.57 | -0.009 | 6 | $* 0.0000$ |
| WJ | 0.35 | 4 | 0.09 | -0.101 | 9 | $* 0.0000$ |
| SW:I | 56.26 | 8 | 7.03 | 2.344 | 18 | 0.1302 |
| SJ:I | 5.26 | 8 | 0.66 | 0.219 | 18 | 0.0122 |
| IWJ | 0.91 | 4 | 0.23 | -0.156 | 18 | $* 0.0000$ |
| SWJ:I | 11.13 | 16 | 0.70 | 0.696 | 54 | 0.0129 |
| $=$ |  |  |  |  |  |  |

*Expected variances of negative estimates of variance components are set to zero.
$E \sigma^{2}(x)$ means expected variance of each source.
S:I means students nested within instructor.

## TABLE 9

## A COMPARISON OF GENERALIZABILITY COEFFICIENTS AND MEAN CORRELATION COEFFICIENTS ACROSS JUDGES IN SEMESTERS II, III, AND Va

| Semester | Generalizability | Mean <br> Correlation <br> Coefficients |
| :---: | :---: | :---: |
| 111 | .75 | .75 |
| V | .69 | .58 |

*This mean correlation coefficient was computed on the scores of the same sample size of six students that was used to calculate the $G$ coefficient in Semester $V$.
${ }^{a}$ Semester IV data were not suitable for generalizability analysis because the same three judges did not score the anecdotal records of all corresponding students.
the fact that $G$ coefficient is a more sensitive measure of reliability, but also possibly that only the data generated by a small sample of six students was utilized to compute the generalizability coefficient.

In summary, the generalizability coefficients and the correlation coefficients across the three judges were comparable in Semesters II, III, and V. Phase two analyses--multiple regressions--were thus initiated.

## Regression Analyses

Preliminary Analyses: Raw Clinical Scores Versus Standardized Clinical Scores
Semester II. Table 10 shows the means, standard deviation, correlation coefficients, and coefficient of determination $\left(R^{2}\right)$ for both the raw and the standardized clinical scores of Semester II, Week 3. Comparing the statistics of the raw clinical scores to the standardized clinical scores, we notice that their standard deviations, correlation coefficients, and the $R^{2}$ are very close. For example, the $R^{2}$ of the raw and standardized clinical scores were .419 and .432, respectively. Furthermore, the order of entry of the independent variables in the regression equations was the same--Psychology first, then Biology, and finally Nursing.

In conclusion, there were no significant differences between the regression results of the raw and standardized clinical scores of Semester II, Week 3.

Semester III. The mean, standard deviation, correlation matrix, and $R^{2}$ for both the raw and the standardized clinical scores of Semester III, Week 3 are shown in Table 11. Like the Semester II regression analyses, the results of

SUMMARY OF STEPWISE MULTIPLE REGRESSION ANALYSIS OF WEEK 3 OF SEMESTER II: A COMPARISON OF RAW CLINICAL SCORES TO STANDARDIZED CLINICAL SCORES

| Raw vs Standardized | Variable | Statistics |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\bar{x}$ | SD |  |  |  |  | $\mathrm{R}^{2}$ |
|  |  |  |  | $\begin{aligned} & \text { Clinical } \\ & \text { Scores } \end{aligned}$ | Nursing | Psychology | $\overline{\text { Biology }}$ |  |
| Raw | Clinical score | 4.708 | 0.899 |  | . 384 | . 620 | -. 015 |  |
|  |  |  |  |  |  |  |  |  |
|  | Nursing | 2.417 | 0.515 |  |  | . 534 | -. 255 | . 419 |
| Clinical |  |  |  |  |  |  |  |  |
|  | Psychology | 2.833 | 0.937 |  |  |  | -. 280 | . 385 |
| Scores |  |  |  |  |  |  |  |  |
|  | Biology | 2.083 | 0.289 |  |  |  |  | . 412 |
| Standardized | Clinical score | -0.002 | 0.852 |  | . 394 | . 636 | -. 039 |  |
|  |  |  |  |  |  |  |  |  |
|  | Nursing | 2.417 | 0.515 |  |  | . 534 | -. 255 | . 432 |
| Clinical |  |  |  |  |  |  |  |  |
|  | Psychology | 2.833 | 0.937 |  |  |  | -. 280 | . 404 |
| Scores |  |  |  |  |  |  |  |  |
|  | Biology | 2.083 | 0.289 |  |  |  |  | . 425 |

$N=12$ (the same sample of 12 students used in generalizability analysis)

SUMMARY OF STEPWISE MULTIPLE REGRESSION ANALYSIS OF WEEK 3 OF SEMESTER III: A COMPARISON OF RAW CLINICAL SCORES TO STANDARDIZED CLINICAL SCORES

| RawvsStandardized | Variable | Statistics |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\bar{X}$ | SD | r |  |  |  | $\mathrm{R}^{2}$ |
|  |  |  |  | Clinical Scores | Nursing | Psychology | Biology |  |
|  | Clinical score | 5.541 | 0.879 |  | . 761 | . 761 | . 465 |  |
| Raw |  |  |  |  |  |  |  |  |
|  | Nursing | 2.833 | 0.577 |  |  | 1.000 | . 369 | . 579 |
| Clinical . |  |  |  |  |  |  |  |  |
|  | Psychology | 2.833 | 0.577 |  |  |  | . 369 |  |
| Scores |  |  |  |  |  |  |  |  |
|  | Biology | 3.000 | 0.853 |  |  |  |  | . 618 |
|  | Clinical score | -0.021 | 0.857 |  | . 742 | . 742 | . 448 |  |
| Standardized |  |  |  |  |  |  |  |  |
|  | Nursing | 2.833 | 0.577 |  |  | 1.000 | . 369 | . 551 |
| Clinical |  |  |  |  |  |  |  |  |
|  | Psychology | 2.833 | 0.577 |  |  |  | . 369 |  |
| Scores |  |  |  |  |  |  |  |  |
|  | Biology | 3.000 | 0.853 |  |  |  |  | . 586 |

$$
N=12
$$

Semester III revealed that the standard deviations, correlation coefficients, and the $R^{2}$ 's of both the raw and the standardized clinical scores were approximately the same. For example, the $R^{2}$ 's of the raw and the standardized analyses were .618 and .586 , respectively. Furthermore, the order of the entry of the independent variables in the regression equations was the same--Nursing and Biology. It should be noted that because Nursing and Psychology were perfectly correlated ( $r=1.00$ ), the computer algorism selected Nursing.

In conclusion, there were no important differences between the results of the regression analyses of the raw and the standardized clinical scores of Week 3 of Semesters II and III. Hence, it was decided to use the mean raw clinical scores to do the regression analyses--the clinical scores given by the three nursing judges--without standardizing them.

## Final Analyses: Raw Clinical Scores

Semester II. Table 12 shows the means, standard deviations, correlation matrices, coefficients of determination $\left(R^{2}\right)$ and the partial $r$ 's for the regression analysis of the data of Semester II. The results of Week I show that only Psychology entered the multiple regression equation, producing an $R^{2}$ of .259 , which was significant at the .10 level or better. An inspection of the partial $r^{1}$ 's shows that Biology would have entered next in the regression equation. However, its contribution to the prediction of the clinical scores was not significant at the . 10 level. Consequently, it did not enter the regression equation.

TABLE 12

SUMMARY OF STEPWISE MULTIPLE REGRESSION ANALYSES: SEMESTER II

| Week | Variable | Statistics |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | $\bar{x}$ | SD | $r$ r |  |  | $\mathrm{R}^{2}$ | Partial r |
|  |  |  |  | Clinical Nursing | Psychology | Biology |  |  |
| 1 | Clinical score | 4.93 | 0.81 | . 104 | . 509 | . 085 |  |  |
|  | Nursing | 2.21 | 0.80 |  | . 499 | -. 077 |  | -. 201 |
|  | Psychology | 2.79 | 0.89 |  |  | -. 253 | . 259 |  |
|  | Biology | 2.07 | 0.27 |  |  |  |  | . 257 |
| 2 | Clinical score | 4.67 | 0.91 | . 580 | . 411 | -. 053 |  |  |
|  | Nursing | 2.21 | 0.80 |  | . 499 | -. 077 | . 337 |  |
|  | Psychology | 2.79 | 0.89 |  |  | -. 253 | . 356 |  |
|  | Biology | 2.07 | 0.27 |  |  |  |  | . 034 |
| 3 | Clinical score | 4.56 | 1.02 | . 643 | . 636 | . 031 |  |  |
|  | Nursing | 2.21 | 0.80 |  | . 499 | -. 077 | . 413 |  |
|  | Psychology | 2.79 | 0.89 |  |  | -. 253 | . 545 |  |
|  | Biology | 2.07 | 0.27 |  |  |  | . 577 |  |
| Combined | Clinical score | 4.72 | 0.75 | . 560 | . 632 | . 023 |  |  |
|  | Nursing | 2.21 | 0.80 |  | . 499 | -. 077 | . 480 |  |
|  | Psychology | 2.79 | 0.89 |  |  | -. 253 | . 400 |  |
|  | Biology | 2.07 | 0.27 |  |  |  | . 509 |  |

$$
N=14
$$

The results of Week 2 show that both Nursing and Psychology entered in the regression equation, producing an $R^{2}$ of .356 , which was significant at $\alpha=.10$. The partial r of Biology, when both Nursing and Psychology had entered in the regression equation, was .034 , confirming that it would not have added much to the regression coefficient.

The results of Week 3, on the other hand, show that all three independent variables entered the regression equation, producing an $R^{2}=.577$, which was significant at $\alpha=.10$. The order of the entry was: Nursing, Psychology, Biology.

Finally, the regression results of Weeks 1, 2, and 3 combined show that all three variables entered the regression equation, producing an $R^{2}$ of .509 , which was significant at the .10 level. In contrast to Week 3, the order of entry of the independent variables was: Psychology, Nursing, Biology.

In summary, 51 percent of the variance of the combined clinical scores in Semester II can be attributed to variations in grades in Psychology, Nursing, and Biology.

Semester III. Table 13 shows the means, standard deviations, correlation matrices, coefficients of determination $\left(R^{2}\right)$, and the partial $r^{\prime}$ s for the regression analyses of the data of Semester III. An examination of the correlation matrices shows that Nursing and Psychology had a correlation coefficient of 1.00 . Consequently, only one of these two variables would have had the opportunity to enter in the regression equation $(\alpha=.10)$. However, either of these variables-Psychology or Nursing--when used alone in the regression analysis, would have yielded substantially the same results.

SUMMARY OF STEPWISE MULTIPLE REGRESSION ANALYSES: SEMESTER III

| Week | Variable | $N$ | $\bar{\chi}$ | SD | Correlation Matrix |  |  |  | $R^{2}$ | Partial r |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Clinical Score | Nursing | Psychology | Biology |  |  |
| 1 | Clinical score | 11 | 6.06 | 0.89 |  | . 45 | . 45 | . 46 |  |  |
|  | Nursing | 11 | 2.82 | 0.60 |  |  | 1.00 | . 43 |  |  |
|  | Psychology | 11 | 2.82 | 0.60 |  |  |  | . 43 |  |  |
|  | Biology | 11 | 3.09 | 0.83 |  |  |  |  |  |  |
| 2 | Clinical score | 12 | 5.64 | 0.82 |  | . 79 | . 79 | . 65 |  |  |
|  | Nursing | 12 | 2.83 | 0.58 |  |  | 1.00 | . 37 | . 621 |  |
|  | Psychology | 12 | 2.83 | 0.58 |  |  |  | . 37 |  | . 000 |
|  | Biology | 12 | 3.00 | 0.85 |  |  |  |  | . 768 |  |
| 3 | Clinical score | 12 | 5.54 | 0.88 |  | . 76 | . 76 | . 47 |  |  |
|  | Nursing | 12 | 2.83 | 0.58 |  |  | 1.00 | . 37 | . 579 |  |
|  | Psychology | 12 | 2.83 | 0.58 |  |  |  | . 37 |  | . 000 |
|  | Biology | 12 | 3.00 | 0.85 |  |  |  |  | . 618 |  |
| Combined | Clinical score | 12 | 5.74 | 0.71 |  | . 80 | . 80 | . 62 |  |  |
|  | Nursing | 12 | 2.83 | 0.58 |  |  | 1.00 | . 37 | . 637 |  |
|  | Psychology | 12 | 2.83 | 0.58 |  |  |  | . 37 |  | . 000 |
|  | Biology | 12 | 3.00 | 0.85 |  |  |  |  | . 760 |  |

Data for one student was missing for one week; in this case clinical scores for Weeks 1, 2, and 3 combined were computed by averaging her scores on the remaining two weeks.

In Week 1, none of the variables entered the regression equation. Unlike Week 1, in Week 2 both Nursing and Biology entered the regression equation, producing an $R^{2}$ of .768 which was significant at the .10 level or better. The partial r of Psychology was . 000, an expected value, given the fact that Psychology and Nursing had a perfect $r$ of 1.00 .

In Week 3, again Nursing and Biology entered in the regression equation, producing an $R^{2}$ of .620 which was significant at $\alpha=.10$.

Finally, combining Weeks 1,2 , and 3 yielded $R^{2} .76$, significant at $\alpha=.10$.

In summary, 76 percent of the variance of the clinical scores of Semester III students could be attributed to their first semester Nursing theory and Biology grades.

Semester IV. Table 14 shows the means, standard deviations and the correlation matrices for Semester IV. An inspection of the correlation matrices shows that the correlation coefficients between the dependent variable--clinical scores--and the independent variables--Nursing, Psychology, and Biology--were smaller than the correlation coefficients between the independent variables. The above, together with the problem of small sample size, created a situation where none of the variables entered the regression equation at the set $\alpha$ level of .10 .

Semester V. Table 15 shows the means, standard deviations and the correlation matrices for Semester V. Like Semester IV, the correlation coefficients between the dependent and the independent variables were in general smaller than the r's between the independent variables. The latter, together with the problem

TABLE 14
SUMMARY OF STEPWISE MULTIPLE REGRESSION ANALYSES: SEMESTER IV

| Week | Variable | N | $\bar{X}$ | SD | Correlation Matrix |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Clinical Score | Nursing | Psychology | Biology |
| 1 | Clinical score | 15 | 5.31 | 1.42 |  | . 331 | . 282 | . 015 |
|  | Nursing | 15 | 2.93 | 0.70 |  |  | . 620 | . 514 |
|  | Psychology | 15 | 3.00 | 2.65 |  |  |  | . 452 |
|  | Biology | 15 | 2.67 | 0.72 |  |  |  |  |
| 2 | Clinical score | 14 | 5.24 | 0.90 |  | . 397 | . 359 | . 250 |
|  | Nursing | 14 | 3.00 | 0.68 |  |  | . 552 | . 486 |
|  | Psychology | 14 | 3.07 | 0.62 |  |  |  | . 396 |
|  | Biology | 14 | 2.79 | 0.70 |  |  |  |  |
| 3 | Clinical score | 14 | 5.35 | 1.20 |  | . 289 | . 213 | . 020 |
|  | Nursing | 14 | 2.85 | 0.69 |  |  | . 684 | . 591 |
|  | Psychology | 14 | 3.00 | 0.71 |  |  |  | . 488 |
|  | Biology | 14 | 2.77 | 0.73 |  |  |  |  |
| Combined | Clinical score | 16 | 5.33 | 0.93 |  | . 389 | . 298 | . 037 |
|  | Nursing | 16 | 2.94 | 0.68 |  |  | . 620 | . 513 |
|  | Psychology | 16 | 3.00 | 0.63 |  |  |  | . 449 |
|  | Biology | 16 | 2.69 | 0.70 |  |  |  |  |

Data for five students was missing for one week; in these cases clinical scores for Weeks 1, 2, and 3 combined were computed by averaging their scores on the remaining two weeks.

## TABLE 15

SUMMARY OF STEPWISE MULTIPLE REGRESSION ANALYSES: SEMESTER V

| Week | Variable | N | $\bar{\chi}$ | SD | Correlation Matrix |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  |  |  |  | Clinical <br> Score | Nursing | Psychology | Biology |
| 1 | Clinical score | 17 | 5.30 | 1.16 |  | -. 311 | . 303 | . 035 |
|  | Nursing | 17 | 3.12 | 0.86 |  |  | -. 115 | . 499 |
|  | Psychology | 17 | 3.12 | 0.78 |  |  |  | -. 333 |
|  | Biology | 17 | 2.82 | 0.64 |  |  |  |  |
| 2 | Clinical score | 16 | 4.73 | 0.91 |  | . 056 | . 109 | . 094 |
|  | Nursing | 16 | 3.19 | 0.83 |  |  | -. 136 | . 436 |
|  | Psychology | 16 | 3.13 | 0.81 |  |  |  | -. 367 |
|  | Biology | 16 | 2.88 | 0.62 |  |  |  |  |
| 3 | Clinical score | 17 | 5.29 | 0.98 |  | . 095 | . 349 | . 002 |
|  | Nursing | 17 | 3.12 | 0.86 |  |  | -. 115 | . 499 |
|  | Psychology | 17 | 3.12 | 0.78 |  |  |  | -. 333 |
|  | Biology | 17 | 2.82 | 0.64 |  |  |  |  |
| Combined | Clinical score | 17 | 5.14 | 0.83 |  | -. 137 | . 311 | . 002 |
|  | Nursing | 17 | 3.12 | 0.86 |  |  | -. 115 | . 499 |
|  | Psychology | 17 | 3.12 | 0.78 |  |  |  | -. 333 |
|  | Biology | 17 | 2.82 | 0.64 |  |  |  |  |

Data for one student was missing for one week; in this case clinical scores for Weeks 1, 2, and 3 combined were computed by averaging her scores on the remaining two weeks.
of small sample size, created a situation where the contribution of the variables was not significant at the .10 level, and therefore did not enter in the regression equation.

In summary, 51 and 76 percent of the variance in the clinical performance scores of Semesters II and III nursing students could be accounted for by their first semester Nursing theory, Biology, and Psychology grades. However, in both Semesters IV and $V$, the coefficients of determination were small and not significant. Further discussions of these findings are presented in Chapter V.

## Chapter V

SUMMARY, FINDINGS AND CONCLUSIONS, LIMITATIONS, AND RECOMMENDATIONS

The purpose of this study was to increase knowledge of factors which correlate with the clinical performance of nursing students, thereby assisting nurse educators to better predict students' performance, and accordingly give counsel.

A search of the literature revealed that very few studies have been done in the area of prediction of clinical performance of nursing students. Of the few studies done, however, the reported relationships between performance in theory courses and clinical practice courses were weak. Taylor et al. (1965) suggested that perhaps the clinical grades represent measures of behaviors and abilities additional to those measured by academic courses. Furthermore, none of the studies addressed the question of reliability of the clinical grades that were being predicted. The latter has disturbed many nurse educators, who have expressed dissatisfaction with the typical process of clinical evaluation.

The problem of the present study was to determine the relative degree to which (a) academic grades and (b) complexity of the nursing situation predict scores in clinical performance for first and second year nursing students.

The student sample consisted of 14 Semester II students; 12 Semester III students; 16 Semester IV students; and 17 Semester $V$ students. Of the 16 instructors who took part in the study, two acted only as nursing judges, and 14 acted as both clinical instructors and judges.

Rines' (1963) anecdotal record form was adapted to meet the needs of the study. The scoring instrument, designed by the investigator, contained criteria with which the clinical instructors and judges were to use a 9 point Likert-type scale to rate the degree of student achievement of stated clinical course objectives. The scale was made more explicit by employing the conventional college grading system to explain the numerical scores on the scale. For example, any score from 8 to 9 was equivalent to superior performance, or a letter grade of "A" standing.

The instrument for measuring complexity of the nursing situation was also designed for this study. A three point scale reflecting different levels of complexity was used. Each level of complexity was described by four major criteria which were to be used by the clinical instructors to make their decisions (see Appendix G).

The videotape-evaluation of a student's clinical performance was designed by the investigator. Its purpose was to train instructors in the writing of anecdotal records about students' clinical performance. The tape, consisting of five vignettes, ran for a total of nine minutes and depicted the performance of a nursing student caring for two patients.

Separate workshops were conducted by the investigator to train the instructors of Semesters II, III, IV, and V, in writing and scoring anecdotal records, and determining the level of complexity of the nursing situation. The workshops were conducted in two parts. First, the advantages and disadvantages of the anecdotal record technique of evaluation were discussed, followed by a discussion of each of the five vignettes. This was followed by practice in recording and scoring the anecdotal record of the student's performance observed on the videotape, and determining the level of complexity of this student's clinical assignment.

Data for the study were collected over a three-week period. These data included: (a) clinical performance scores obtained from anecdotal records of students' clinical performance prepared by their instructors and, subsequently, scored by the clinical instructor and two nursing judges using the student's clinical objectives and the scoring instrument; (b) rating of the level of complexity of the students' clinical assignments by their clinical instructors, using the complexity of the nursing situation instrument; and (c) first semester Nursing theory, Psychology, and Biology grades obtained from the students' permanent records in the Registrar's office.

Data analyses were carried out in two phases. Phase one involved the utilization of two statistical methods to analyze the reliability of the clinical performance scores: (1) conventional correlation analyses; and (2) generalizability analyses. The dependent variable was the average of the clinical performance scores of the students given by the clinical instructors (who wrote and scored
the anecdotal records) and two nursing judges (who scored the written anecdotal records). The independent variables were the grades in first semester Nursing, Biology, and Psychology; and the complexity of the nursing situation. In phase two, multiple regression analyses were conducted to determine the proportion of variance of the clinical performance scores in Semesters II, III, IV, and V accounted for by the academic grades obtained in Semester I Nursing, Psychology, and Biology.

Findings and Conclusions

Both the generalizability and correlational analyses showed that the inter-rater reliabilities among the three judges in all of the four semesters were moderate to moderately strong (from .58 to .84 ). The generalizability (G) coefficients were usually higher than the correlation coefficients. This discrepancy can be explained by the fact that the $G$ coefficient is a more sensitive measure of reliability when compared with the conventional correlation coefficient.

Given that the clinical scores were found to be sufficiently reliable, the data analysis proceeded to phase two. Inspection of the ratings of the levels of complexity of the nursing situation in each semester revealed minimal to no variation from level two complexity. Consequently, the independent variable complexity of the nursing situation could not be used in the regression analysis.

The regression analysis of Semester II data revealed that 51 percent of the variance of the clinical scores could be explained by the three remaining independent variables--academic grades in first semester Nursing, Psychology, and Biology. The squared multiple correlation coefficient was significant at . 10 level of significance.

For Semester III, 76 percent of the variance of the clinical score was accounted for by the Nursing or Psychology grades, and the Biology grades (p. < .10). In this analysis, Nursing and Psychology grades were perfectly correlated. Consequently, either of these variables, when used alone, explained the same amount of variance in the clinical scores.

In contrast, the regression analyses for Semesters IV and $V$ revealed that only 20 and 16 percent of the variance of the clinical scores respectively, was accounted for by the three independent variables. These results were not significant. In summary, it appears that the academic grades in Nursing, Psychology, and Biology courses obtained in first semester of the nursing program are better predictors of clinical performance in earlier semesters, Semesters II and III, than in later semesters, Semesters IV and V.

Based on the above findings, the following conclusions were made:

1. The anecdotal record method of clinical performance evaluation yielded acceptably reliable clinical scores where all the clinical instructors (those instructors who wrote the anecdotal records) and all the judges (those who evaluated the written anecdotal records) were trained using a planned instructional program.
2. The grades in Psychology, Nursing, and Biology courses during the first semester of the nursing program accounted for a significant proportion of the
variance in the clinical performance scores of nursing students in Semester II. Likewise, Nursing or Psychology, and Biology grades obtained in the first semester of the nursing program accounted for a significant proportion of the variance in the clinical performance scores of Semester III nursing students.

In Semesters IV and V, however, these grades failed to account for a significant proportion of the variance in clinical performance scores of these students. Therefore, hypothesis 1--grades in Nursing, Biology, and Psychology courses during the first semester of the nursing program account for a significant proportion of the variance in clinical performance scores of students--is accepted in the case of Semester II and III students, and rejected in the case of Semester IV and V students.
3. The complexity of the nursing situation could not be used in the regression analysis because it showed no variability across the three weeks and the four semesters. Therefore, hypothesis 2--complexity of the nursing situation accounts for an additional significant proportion of the variance in clinical performance scores of students--could not be tested.

There are several interrelated factors which may explain the differences noted between Semesters II and III and Semesters IV and V clinical performance and hence the multiple correlations with the independent variables selected. As the nursing students progress from one semester to another, their responsibilities increase. They are expected to (1) be more independent; (2) have acquired a repertoire of skills and competencies to make more independent decisions that require higher levels of cognitive abilities; (3) provide nursing care to individuals
with complex health problems; (4) provide nursing care for a larger number of individuals; (5) assume a greater role as a member of the nursing team; and (6) require less direct instructor supervision.

All the above factors may have influenced the nature of the students' clinical assignments, the amount of required and/or desired instructor supervision, and the accessibility of the clinical instructor to the student. Based on this information, it can be safely assumed that as the students progress from Semester II to Semester $V$, their clinical assignments will increase in number and complexity, they will require less direct instructor supervision, and will have less access to their clinical instructors. It can be conjectured that the aforementioned may affect the nature of the content of the anecdotal records written by the clinical instructors, and hence the clinical scores of the students. Consequently, it can be expected that these factors may influence the performance scores of Semester IV and $V$ students more than Semester II or III students. Therefore, to the extent that the nature of the students' assignments change as a result of their natural progression in the nursing program, their clinical performance may be affected, and thus may not correlate as highly with previously identified variables.

Furthermore, there were temporal factors which may have contributed to the differences noted in the multiple correlation coefficients of Semesters IV and $V$. These senior students had additional clinical experiences, and higher level Nursing, Psychology, and Biology courses which may have affected their attitudes, motivation, competencies, perceptions of the role of the nurse, perceptions of the reality of clinical nursing, hence their clinical performance and clinical scores.

Finally, a possible factor which may have influenced the performance of Semester II and III students was the increase in the amount and type of clinical experiences of these students in their first semester of nursing. These junior students had more clinical experiences in their Semester I clinical nursing course compared to the senior students of Semesters IV and V. It is important to note, however, that this factor has similarities to the temporal factors described above, which were conjectured to have contributed to the differences noted in the multiple correlation coefficients between Semesters II and III and Semesters IV and V.

There were several factors which may have influenced the ratings of the independent variable, the complexity of the nursing situation, yielding complexity levels which were essentially the same--level two. The criteria for level two complexity were: 1. activities of the student--treatments prescribed by the doctor and/or the nursing team are moderate in number; the student has had numerous opportunities to practise the skills and activities of the previous semester, and has minimal/moderate opportunities to practise the skills and activities learned in the present semester; 2. condition of the patient--the patient's condition is changing and requires moderate modification of the student's plan of care; 3. patient experiences--available patient experiences meet the learning needs of students in a given semester; suitable student/patient ratio, with patients who provide desired experiences; and 4. revisions of student's assignment based on the instructor's assessment of the student's abilities/per-formance--the assignment requires: integration of previously acquired and some
newly learned knowledge; coordination of nursing activities; flexibility in the student's organization and plan of care; problem-solving activities (assessment, planning, implementation, and evaluation); and the student can cope with assignments for his semester level (for further explanation see Appendix G). This variable had three levels of complexity which may have encouraged the clinical instructors to perceive and choose level two complexity as the desired and ideal complexity of the nursing situation. It is also possible that in reality complexity of the nursing situation does not vary.

## Limitations of the Study

1. Conclusions can only be generalized to the particular institution that was sampled. However, there are no known reasons for believing that college nursing instructors are not typical of a much larger population of nursing instructors in other institutions offering a diploma in nursing.
2. The fact that the complexity scale had only three levels of complexity, a psychological phenomenon of rating towards the center may have occurred yielding complexity scores which were essentially the same. Also, this scale was not pilot tested.
3. The dependent variable--the clinical performance--was only assessed for the first half of the semester, and therefore may not reflect the students' performance for the whole semester.

## Recommendations

This study was exploratory in nature and served to identify the degree to which academic grades in Nursing, Psychology, and Biology courses in first semester of the nursing program correlate with the clinical performance scores of nursing students. The recommendations based on the results of the study will be presented in two parts, recommendations for future research and recommendations for nurse educators.

The recommendations for future research are as follows:

1. The complexity of the nursing situation be studied further to establish whether the tool used to assess complexity was indeed insensitive in differentiating levels of complexity, or whether the nursing situation itself lacked complexity. Furthermore, the tool to measure the complexity of the nursing situation be refined using four levels rather than three, so that differences in the complexity of the nursing situation are more finely differentiated.
2. A study be conducted with a second group of Semester IV and V students, substituting the theory grades obtained in their first semester Nursing, Psychology, and Biology courses with their third and fourth semester Nursing, Psychology, and Biology courses respectively to control the possible influence of the temporal factor of additional learning experiences.
3. A study be conducted to determine whether the multiple correlation coefficients obtained utilizing the clinical scores over a three-week period during the first half of the semester are comparable to the correlation coefficients obtained utilizing the clinical grades at the end of the semester.

The recommendations for nurse educators are as follows:

1. Results of this study be used with caution when providing guidance to nursing students enrolled in two-year nursing programs.
2. Results of this study be used with caution when predicting clinical performance scores of students enrolled in two-year nursing programs from their first semester grades in Nursing, Psychology, and Biology.

## REFERENCES

Ahmann, S. J., \& Glock, M. D. Evaluating Pupil Growth (2nd ed.). Boston: Allyn and Bacon, Inc., 1971.

Blood, D. F., \& Budd, W. C. Educational Measurement and Evaluation. New York: Harper and Row, Publishers, Inc., 1972.

Champney, B., Chatfield, K. M., \& Sims, A. Methods of Assessment. A Guide for Teachers of Nurses, edited by Elizabeth Raybould. Oxford: Blackwell Scientific Publication, 1975, 155-173.

Cronbach, L. J., et al. The Dependability of Behavioral Measurements: Theory of Generalizability for Scores and Profiles. New York: John Wiley and Sons, Inc., 1972.

Fivars, G., \& Gosnell, D. Nursing Evaluation: The Problem and the Process. New York: The Macmillan Co., 1966.

Fuerst, E. V., \& Wolff, L. Teaching Fundamentals of Nursing (4th ed.). Philadelphia: J. B. Lippincott Co., 1969.

Gonnella, J. S., Goran, M. J., Williamson, J. W., \& Cotsonas, N., Jr. Evaluation of Patient Care: An Approach. Journal of American Medical Association, 1970, 214, 2040-2043.

Gordon, P. Evaluation: A Tool in Nursing Service. The American Journal of Nursing, 1960, 3, 364-367.

Gore, B. W. The Multivariate Prediction of College Grade Point Average and State Board Examination Scores in Selected Mississippi Associate Degree Nursing Programs (Doctoral Dissertation). Mississippi: University of Southern Mississippi, 1973.

Goza, J. T. An Investigation of the Academic Potential, Academic Achievement, and Personality of Participants in an Associate Degree Nursing Program (Doctoral Dissertation). Texas: East Texas State University, 1970.

Green, P. E., \& Tull, D. S. Research for Marketing Decisions (2nd ed.). New Jersey: Prentice-Hall, Inc., 1970.

Gronlund, N. E. Measurement and Evaluation in Teaching (3rd ed.). New York: Macmillan Publishing Co., Inc., 1976.

Hart, W. M. The Evaluation of the Performance of Baccalaureate Students in Clinical Nursing (Doctoral Dissertation). Indiana University, May 1974.

Hayter, J. An Approach to Laboratory Evaluation. The Journal of Nursing Education, 1973, 12, 17-22.

Heslin, P. Evaluating Clinical Performance. Nursing Outlook, 163, 11, 344-345.

Infante, M. S. The Clinical Laboratory in Nursing Education. New York: John Wiley and Sons, Inc., 1975.

Kerlinger, F. N., \& Pedhazur, E. J. Multiple Regression in Behavioral Research. New York: Holt, Rinehart and Winston, Inc., 1973.

Korman, M., \& Stubblefield, L. R. Medical School Evaluation and Internship Performance. Journal of Medical Education, 1971, 46, 670-673.

Kruger, M. Evaluation of a Student's Clinical Performance, a Videotape. Vancouver: Vancouver Community College, Langara, 1978.

Lancaster, A. Nurse Teachers: the Report of an Opinion Survey. Belfast: W. \& G. Baird Ltd., 1972.

Leaper, D. J., Gill, P. W., Staniland, J. R., Horrocks, J. C., \& De Dombal, F. T. Clinical Diagnostic Progress: An Analysis. British Medical Journal, 1973, 3, 569-574.

Litwack, L., Sakata, R., \& Wykle, M. Counselling, Evaluation and Student Development in Nursing Education. Philadelphia: W. B. Saunders, 1972.

MacKay, R. C. Evaluation of Faculty and Students. . . . A Means Towards Fuller Communication and Greater Productivity. The Journal of Nursing Education, 1974, 13, 3-7.

McManus, R. L. The Effect of Experience on Nursing Achievement. New York: Teachers College, Columbia University Bureau of Publications, 1949.

National League for Nursing. Evaluation - An Objective Approach. New York: The League, 1972.

Powers, B. Prediction of Success of Community College Nursing Students (Doctoral Dissertation). Arizona State University, December 1974.

Publication Manual of the American Psychological Association. Maryland: Garamond/Pridemark Press, Inc., 1975.

Rines, A. R. Evaluating Student Progress in Learning the Practice of Nursing. New York: Bureau of Publications, Teachers College, Columbia University, 1963.

Ryden, M. B. The Predictive Value of a Clinical Examination of Interpersonal Relationship Skills. Journal of Nursing Education, 1977, 16, 5, 27-31.

Schweer, J. E., \& Gebbie, K. M. Creative Teaching in Clinical Nursing (3rd ed.). Saint Louis: The C. V. Mosby Company, 1976.

Sister Miller, P. A Teaching Practicum. Nursing Outlook, 1976, 24, 5, 752-753.

Tate, B. L. Evaluating the Nurse's Clinical Performance. Nursing Outlook, 1962, 10, 35-37.

Taylor, C. W., Nahm, H., Quinn, M., Harms, M., Mulaik, J., \& Mulaik, S. A. Selection and Recruitment of Nurses and Nursing Students: A Review of Research Studies and Practices. Utah: University of Utah Press, 1963.

Taylor, C. W., Nahm, H., Quinn, M., Harms, M., Mulaik, J., \& Mulaik, S. A. Report of Measurement and Prediction of Nursing Performance, Part I. Salt Lake City: University of Utah Press, 1965.

Thorndike, R. L., \& Hagen, E. Measurement and Evaluation in Psychology and Education (2nd ed.). New York: John Wiley and Sons, Inc., 1961.

Wiener, S. L., Koran, L., Mitchell, P., Schattner, G., Fierstein, J., \& Hotchkiss, E. Clinical Skills: Quantitative Measurement. New York State Journal of Medicine, 1976, 76, 610-612.

Wood, V. A. A Problem That Won't Go Away. International Nursing Review, 1972, 19, 336-343.

APPENDICES

APPENDIX A

APPENDIX B

## MEMO

to Mrs. M. Kruger, Nursing Department
from A.L. Dartnell, Acting Principal
date January 9, 1978
subject Research

I have very carefully read your proposal and also have discussed the matter with Mrs. Knor.

This memo will give you permission to proceed as outlined in your proposal. If there is any way in which I can help would you please let me know.

ALD/SB
Enc:
c.c. Mrs. E. Knor

APPENDIX C

## EXPLANATIONS TO INSTRUCTORS

## Dear Colleague:

I discussed my proposal with you at faculty meetings in both April and October, 1977 and would now like to request your participation.

As you will remember, the study is to determine what proportion of the variance in the clinical performance score of the nursing student is contributed by the grades in the first semester nursing, biology and psychology courses, and the complexity of the nursing situation.

The information gathered will help nurse educators be cognizant of the effect of such factors on the clinical performance grade of the nursing students and accordingly evaluate the students' performance.

The anecdotal records will be typed and coded to ensure confidentiality. I would like to make it very clear that this is not an evaluation of your ability to write anecdotal notes, make clinical assignments, or grade students' performance.

A copy of the completed study will be provided to the nursing department and made available to all participants.

Your participation will entail:
(1) writing anecdotal notes about the performance of 4-8 participating students on the designated forms during the weeks of January 23, 30 and February 6 and 13, 1978.
(2) giving a score to each anecdotal record according to the clinical objectives of the student and the scoring system provided to you by the investigator.
(3) determining the level of complexity of the nursing situation according to the criteria (see Appendix G).
(4) volunteering about 30 to 40 minutes for each participating student to complete activities 1-3.
(5) acting as a nursing judge, to evaluate the performance of 10-14 students on the basis of anecdotal records written by other instructors from your semester.

APPENDIX D

## CONSENT FORM

Dear Colleague:
Please indicate your willingness to participate in this study by signing your name in the appropriate space provided below, and as soon as possible return the consent form to Mary Kruger, office number Al67c.

You may keep the attached information if you wish.

I, $\qquad$ , agree (OR I, $\qquad$ , do not wish) to participate in the study by Mary Kruger titled: What proportion of the variance in the clinical performance score of the nursing student is contributed by the grades in the first semester nursing, biology, and psychology courses and by the complexity of the nursing situation. I understand that the study will involve writing anecdotes about the performance of participating students in my clinical area for the weeks of January 23, 30, and February 6 and 13, 1978; determining the complexity of the nursing situation according to described criteria; and scoring the anecdotal notes of participating students from the semester employing the course objectives and the scoring system.

1 have been assured that the findings will not influence my position in the school and that my name will not appear in any written report.

I am aware that I am free to withdraw from the study at any time.
$\qquad$

APPENDIX E

## APPENDIX F

## CONSENT FORM

## Dear Student:

Please indicate your willingness to participate in this study by signing your name in the appropriate space below, and return the consent form to Mary Kruger, office number Al67c as soon as possible.

1, $\qquad$ , agree (OR I, $\qquad$ ,
do not wish) to participate in the study by Mary Kruger titled: What proportion of the variance in the clinical performance score of the nursing student is contributed by the grades in the first semester nursing, biology, and psychology courses, and by the complexity of the nursing situation. I understand that the study will involve obtaining my first semester nursing, biology, and psychology grades from the registrar's office and scoring of the anecdotal records about my performance, anonymously, for the weeks of January 23 to February 20, 1978.

I have been assured that the findings will not influence my standing in the school and that my name will not appear in any written report.

I am aware that I am free to withdraw from the study at any time.

APPENDIX G

## COMPLEXITY OF THE NURSING SITUATION

$$
\text { level } 1 \quad \text { level } 2 \quad \text { level } 3
$$

On a complexity scale ranging from level 1 to level 3 sketched above, circle on the right the level that best represents the student's assignment according to the criteria described.

$$
\text { Circle only ONE level level } 1 \text { level } 2 \text { level } 3
$$

Description of levels 1,2 and 3

| LEVEL 1 | LEVEL 2 | LEVEL 3 |
| :---: | :---: | :---: |
| activities of the student | activities of the student | activities of the student |
| Treatments prescribed by the Dr. and/ or the nursing team are minimal in number; the student has had numerous previous opportunities to practise these skills and activities, and has minimal or no opportunity to practise the activities and skills learned in the present semester. | Treatments prescribed by the Dr. and/ or the nursing team are moderate in number; the student has had numerous opportunities to practise the skills and activities of the previous semester ${ }_{r}$. and has minimal/moderate opportunities to practise the skills and activities learned in the present semester. | Treatments prescribed by the Dr. and/ or the nursing team are numerous; the student has had minimal/moderate opportunities to practise the skills and activities learned in the previous semester, and has minimal/moderate opportunities to practise the skills and activities learned in the present semester. |

2. 

| LEVEL 1 |
| :---: |
| The patient's condition of the patient is stable and | requires minimal or no modification of the student's plan of care.

## patient experiences

(a) not enough patients on the ward providing the desired experiences.
(b) too many students on the ward; therefore, limited suitable patient experiences for the students of the given semester.
(c) initial appropriate patient assignment; however, the clinical instructor changes the assignment because
(1) patient transferred to another ward
(2) patient's condition has worsened
(3) patient is discharged
(4) patient has expired.
LEVEL 2
condition of the patient

| The patient's condition is changing |
| :---: |
| and requires moderate modification | and requires moderate modification of the student's plan of care.

patient experiences
available patient experiences meet the learning needs of students in a given semester; suitable student/ patient ratio, with patients who provide desired experiences.

LEVEL 3
condition of the patient
The patient's condition is changing rapidly and requires maximum modification of the student's plan of care.

> patient experiences
the student's initial assignment is according to the learning needs of the students in a given semester; however, the patient's(s') condition changes and the clinical instructor decides that the student should continue to care for the patient(s) in collaboration with other member(s) of the nursing team.
4.
revisions of student's assignment based on the instructor's assessment of the student's abilities/performance

The assignment requires:
(a) integration of previously acquired knowledge;
(b) coordination of nursing activities;
(c) problem-solving activities:
assessment, planning, implementation, and evaluation.

Student can not cope with the type of assignment for his semester level. Therefore,
(a) the clinical instructor changes the assignment
(b) the clinical instructor chooses assignments that she perceives that the student is able to manage.

LEVEL 3
revisions of student's assignment based on the instructor's assessment of the student's abilities/performance

The assignment requires:
(a) integration of previously and newly learned knowledge;
(b) coordination of nursing activities;
(c) maximum flexibility in the student's organization and plan of care;
(d) constant problem-solving activities: assessment, planning, implementation, and evaluation.

Student can cope with assignments for his semester level.

To determine the complexity level of the student's assignment, you must. follow the sequential procedure below.
(1) Read the variables described under levels 1, 2, and 3.
(2) Determine how many of the variables in each level apply to the student's assignment.
(3) Decide on the level that best represents the student's assignment. Below are three possible situations and the appropriate decision in each.
(a) if all the variables from one level apply to the student's assignment, then that is the appropriate level.
(b) if two of the variables from one level and one variable from each of the other levels apply to the student's assignment, then the decision depends on the pattern. There are three situations under this category:
(1) if two variables from level 1, and one variable from each of the levels 2 and 3 apply, then the decision is towards the central tendency, thus level 2.
(2) if one variable from each of the levels 1 and 3 , and two variables from level 2 apply, then the decision is level 2.
(3) if one variable from each of the levels 1 and 2 , and two variables from level 3 apply to the student's assignment, then the decision is level 3.
(c) if two of the variables from one level and two variables from another level apply to the student's assignment, then the decision is the higher level. For example, if there are two variables from level 2, and two variables from level 3, then the decision is level 3.
(4) Circle the level decided in step 3.

APPENDIX H

## ANECDOTAL RECORD

Semester

Date $\qquad$

Description of the Situation

Instructor Number
Student Number $\qquad$

APPENDIX I

## SCORING SYSTEM

On a scale of 1 to 9,1 representing unsatisfactory to 9 representing superior performance, circle on the right the number that best describes the student's achievement of the 6 major objectives.


SEMESTER II

| Objectives Circle ONE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 1. Utilize the problem solving approach to provide systematic care. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 2. Communicate with others in an effective and systematic manner. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 3. Value individuals as unique and having worth. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 4. Demonstrate organizational skills. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 5. Assume his role as a member of the health team. | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 6. Demonstrate responsibility and accountability for his own behavior | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

Note: The above are the terminal objectives under which you have specific objectives for your clinical course. Using the criteria under the scoring system described in Appendix I ranging from "unsatisfactory" to "superior" along with the corresponding letter grades used at the college will enable you to select the appropriate figure ranging from 1 to 9.

## SCORING SYSTEM

On a scale of 1 to 9,1 representing unsatisfactory to 9 representing superior performance, circle on the right the number that best describes the student's achievement of the 6 major objectives.

| F | P | C |  | B |  |  | A |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Unsatisfactory | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Marginal |  | Average | Above <br> Average | Superior |  |  |  |

## SEMESTER III

$$
\begin{array}{lllllllllll}
\text { Objectives } & \text { Circle ONE } & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{array}
$$

1. Utilize the nursing process to provide systematic nursing care. $\quad \begin{array}{llllllllll} & 2 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\end{array}$
2. Communicate with others in an effective and systematic manner. $\begin{array}{llllllllll}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\end{array}$
3. Value individuals as unique and having worth. $\quad \begin{array}{lllllllll} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8\end{array}$
4. Demonstrate organizational skills,
5. Demonstrate accountability for the care of his assigned patients. $\begin{array}{llllllllll} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\end{array}$
6. Demonstrate responsibility for his behavior.
$\begin{array}{lllllllll}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\end{array}$
Note: The above are the terminal objectives under which you have specific objectives for your clinical course. Using the criteria under the scoring system described in Appendix I ranging from "unsatisfactory" to "superior" along with the corresponding letter grades used at the college will enable you to select the appropriate figure ranging from 1 to 9 .

## SCORING SYSTEM

On a scale of 1 to 9,1 representing unsatisfactory to 9 representing superior performance, circle on the right the number that best describes the student's achievement of the 6 major objectives.


## SEMESTER IV

$$
\begin{array}{lllllllllll}
\text { Objectives } & \text { Circle ONE } & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9
\end{array}
$$

1. Utilize the problem solving approach to provide systematic nursing $\begin{array}{lllllllllllllll}1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9\end{array}$ care.
2. Communicate with others in an effective and systematic manner.
3. Value individuals as unique and having worth.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

Note: The above are the terminal objectives under which you have specific objectives for your clinical course.
Using the criteria under the scoring system described in Appendix ! ranging from "unsatisfactory" to "superior" along with the corresponding letter grades used at the college will enable you to select the appropriate figure ranging from 1 to 9 .

## SCORING SYSTEM

On a scale of 1 to 9,1 representing unsatisfactory to 9 representing superior performance, circle on the right the number that best describes the student's achievement of the 6 major objectives.


SEMESTER V

| Objectives $\quad$ Circle ONE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |

1. Utilize the problem solving approach to provide care.
2. Communicate with others in an effective and systematic manner.
3. Value individuals as unique and having worth.
4. Demonstrate organizational skills.
5. Assume his role as a member of the health team.
6. Demonstrate responsibility for his own behavior.

| Circle ONE | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- | :--- |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|  | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |

Note: The above are the terminal objectives under which you have specific objectives for your clinical course.
Using the criteria under the scoring system described in Appendix I. ranging from "unsatisfactory" to "superior" along with the corresponding letter grades used at the college will enable you to select the appropriate figure ranging from 1 to 9.

| Categories | Letter Grade | Numerical Grade | Criteria |
| :---: | :---: | :---: | :---: |
| Unsatisfactory | F | 12 | The student fails to meet the course objectives; <br> (1) requires consistent detailed explanation of principles, concepts, and procedures; <br> (2) requires consistent teacher or staff supervision; <br> (3) fails to apply the problem-solving process; <br> (4) is unable to demonstrate initiative in meeting objectives; <br> (5) is inconsistent in applying previously learned knowledge; <br> (6) is frequently inconsistent and unsafe in his performance of communication and psychomotor skills. |
| Marginal | P | 3 | The student is inconsistent in meeting the course objectives: <br> (1) requires frequent detailed explanation of principles, concepts, and procedures; <br> (2) requires frequent teacher or staff supervision; <br> (3) occasionally applies the problem-solving process; <br> (4) occasionally demonstrates appropriate initiative in meeting objectives; <br> (5) occasionally applies previously learned knowledge; <br> (6) is usually inconsistent and unsafe in his performance of communication and psychomotor skills. |

## Definitions:

Usually - customarily, habitually

Frequently - happening at short intervals, often occurring

Occasionally - happening irregularly, coming now and then

Rarely - seldom, not often
Appropriate - specially suitable, proper, fit

Consistent - marked by regularity, or steady continuity throughout; showing no significant change, unevenness, or contradiction

Initiative - ability to originate actions, to initiate desirable steps

Marginal - close to the lower limit of qualification or acceptability

| Categories | Letter Grade | Numerical Grade | Criteria |
| :---: | :---: | :---: | :---: |
| Average | C | 45 | The student meets all of the overall course objectives: <br> (1) requires occasional explanation of principles, concepts, and procedures; <br> (2) requires occasional teacher or staff supervision; <br> (3) usually applies the problem-solving process; <br> (4) usually demonstrates appropriate initiative in meeting objectives; <br> (5) usually applies previously learned knowledge; <br> (6) is usually consistent and safe in his performance of communication and psychomotor skills. |
| Above Average | B | 67 | The student meets all of the overall and most of the specific objectives: <br> (1) requires occasional explanation of principles, concepts, and procedures; <br> (2) requires appropriate teacher or staff supervision; <br> (3) frequently applies the problem-solving process; <br> (4) frequently demonstrates appropriate initiative in meeting objectives; <br> (5) frequently applies previously learned knowledge; <br> (6) is consistent and safe in his performance of communication and psychomotor skills. |

## Definitions:

Usually - customarily, habitually

Frequently - happening at short intervals, often occurring

Occasionally - happening irregularly, coming now and then

Rarely - seldom, not often
Appropriate - specially suitable, proper, fit

Consistent - marked by regularity, or steady continuity throughout; showing no significant change, unevenness, or contradiction

Initiative - ability to originate actions, to initiate desirable steps

Marginal - close to the lower limit of qualification or acceptability

| Categories | $\begin{array}{l}\text { Letter } \\ \text { Grade }\end{array}$ | $\begin{array}{c}\text { Numerical } \\ \text { Grade }\end{array}$ | Criteria |
| :--- | :--- | :--- | :--- |
| Superior | A | 89 | $\begin{array}{l}\text { The student meets all of the overall and } \\ \text { specific objectives: } \\ \text { (1) rarely requires explanation of principles, } \\ \text { concepts, and procedures; }\end{array}$ |
| (2) requires appropriate teacher or staff |  |  |  |
| supervision; |  |  |  |
| consistently applies the problem-solving |  |  |  |
| process; |  |  |  |\(\left.] \begin{array}{l}(4) consistently demonstrates appropriate <br>


initiative in meeting objectives;\end{array}\right\}\)| (5)consistently applies previously learned <br> knowledge; |
| :--- |
| (6) is consistent and safe in his performance |
| of communication and psychomotor skills. |

## Definitions:

Usually - customarily, habitually
Frequently - happening at short intervals, often occurring

Occasionally - happening irregularly, coming now and then

Rarely - seldom, not often
Appropriate - specially suitable, proper, fit

Consistent - marked by regularity, or steady continuity throughout; showing no significant change, unevenness, or contradiction

Initiative - ability to originate actions, to initiate desirable steps

Marginal - close to the lower limit of qualification or acceptability

APPENDIX J

## GENERALIZABILITY ANALYSIS

The theory of generalizability proposed by L. J. Cronbach et al. (1972) employs the analysis of variance to determine coefficients of reliability. More specifically, this theory is motivated by the following questions: (1) how can a person's score best be estimated? and (2) how large is the error that arises from incomplete observations?

The term "generalizability" refers to the generalization from observed (sample) scores, to universe scores (persons' mean scores over all acceptable observation). Cronbach et al. state that the universal score "emphasizes that the investigator is making an inference from a sample of observed data, and also that there is more than one universe to which he might generalize" (p. 18).

In a generalizability ( $G$ ) study, the researcher collects data from which estimates can be made of the components of variance for measurements made for a certain procedure. Furthermore, the researcher obtains two or more scores for the individual by observing him under different conditions, and then examines the consistency of these scores.

It is necessary here to define some of the essential terminology of the theory of generalizability, to facilitate the description of its application in the present study.

Any observations made are described in terms of conditions, defined as "the task or stimulus presented, the day and hour, the setting in which the observation is made, the observer, and possible additional features of the operations performed" (Cronbach et al., p. 17). The term facet, however, is used to refer to conditions of a certain kind.

The facets, alone or in combination, define the universe in question. Therefore, a universe of observations could be characterized with respect to one, two, or more facets.

The theory of generalizability makes the following assumptions:
(1) conditions of observations are not necessarily parallel;
(2) conditions (particularly test items) are randomly supplied from a universe, or according to a stratified design; and
(3) two or more facets may be analyzed simultaneously.

Generalizability analysis provides estimate components of those facets represented in the experimental design. These estimates of variance components are then used to determine the coefficient of generalizability. Finally, the coefficient of generalizability $\left(\mathrm{E}^{2}\right)$ is determined by the "ratio of the universescore variance to the expected observed-score variance--the variance likely to be obtained under a certain experimental plan" (Cronbach et al., p. 17).

In the notation $E \rho^{2}, E$ stands for expected value, and $\rho^{2}$ is the squared correlation of observed score with universe score. The value of the $G$ coefficient ranges from zero to one, and it expresses how well the observations are likely to locate the individual relative to other members of the population.

In this study, the BMD08V computer program was used to analyze the data. This program generated estimates of variance components (variances of hypothesized components of an observed score), which were utilized in calculating generalizability coefficients.

This study was mainly concerned with the accuracy of generalizing over:
(1) ratings of three nursing judges, and
(2) ratings gathered during three different weeks.

The study employed a nested design with three facets. These were:
(1) Students, nested within instructors ( $\mathrm{S}: 1$ ),
(2) Judges (J), and
(3) Weeks (W)

The design of the study was $(S: 1) \times J \times W$. The diagram below shows schematically the forms of overlapping variances.


An examination of the design of the study identified 11 sources of variance (i.e., components). These were: instructors (I); weeks (W); students, nested within instructors (S:I); judges (J); instructor-by-week interaction (IxW); instructor-by-judge interaction ( $\mathrm{l} \times \mathrm{J}$ ); week-by-judge interaction (W×J); student-by-week interaction ( SxW ); student-by-judge interaction ( $\mathrm{S} \times \mathrm{J}$ ); instructor-by-week-by-judge interaction (íxWxJ); and, finally, student-by-week-by-judge interaction (S: $1 \times W \times J$ ).

Finally, the generalizability coefficients were calculated with the following formula:

$$
E \rho^{2}=\frac{\sigma^{2}(\mathrm{p})+\left(1 / n_{i}^{\prime}\right) \sigma^{2}\left(\mathrm{Pi} / 1^{*}\right)}{\sigma^{2}(\mathrm{p})+\left(1 / \mathrm{n}_{\mathrm{i}}^{\prime}\right) \sigma^{2}\left(\mathrm{Pi} / 1^{*}\right)+\sigma^{2} \delta}
$$

where, $\sigma^{2} p=$ variance component for persons. (In the study, either judges or weeks)
$n_{i}^{\prime}=$ number of conditions for facet $\mathbf{i}$ (items)
$\sigma^{2}\left(\mathrm{Pi} / \mathrm{I}^{*}\right)=$ variance component for the persons-by-items interactions for a fixed set of items. (In this study, students nested within instructor)
${ }^{\circ}{ }^{\circ} 2_{\delta}=$ the error, which includes pi, e components of variance

The numerator and the denominator represent the universe-score variance and the expected observed-score variance, respectively.

APPENDIX K

## STANDARDIZATION PROCEDURE

The following steps explain the procedure used to standardize the raw clinical scores given by the clinical instructor and two other nursing judges.

For steps 1 to 8 refer to Figure 1.
(1) The means of each judge was computed.
(2) The variance of each judge was computed.
(3) The standard deviation of each judge was computed.
(4) Next, from the score of each student in Column $\mathrm{J}_{1}$ was subtracted the mean of $J_{1}$, and subsequently, that number was divided by the standard deviation (SD) of $\mathrm{J}_{1}$. For example, the standardized score of student 303 given by Judge 1 was calculated by this formula. $\frac{3.5-5.75}{\operatorname{SD}\left(\mathrm{~J}_{\mathfrak{j}}\right)}$
(5) Using the above method, the other two scores given to these students were adjusted for Judges 2 and 3.
For example, $\frac{4.0-5.5}{\mathrm{SD}\left(\mathrm{J}_{2}\right)}$ and $\frac{5.0-5.38}{\mathrm{SD}\left(\mathrm{J}_{3}\right)}$ were the standardized scores of student 303 for Judges 2 and 3 .
(6) The pooled mean standardized score of each student was then calculated by adding together the three standardized scores given to her by Judges 1, 2, and 3, and dividing this sum by three. For example, the standardized score of student 303 across the three judges was calculated by this formula:

$$
\frac{\left(\frac{3.5-5.75}{\mathrm{SD}\left(\mathrm{~J}_{1}\right)}\right)+\left(\frac{4.0-5.5}{\mathrm{SD}\left(\mathrm{~J}_{2}\right)}\right)+\left(\frac{5.0-5.38}{\mathrm{SD}\left(\mathrm{~J}_{3}\right)}\right)}{3}
$$

(7) The above computations were repeated to standardize the clinical scores of Week 3 of both Semesters II and III students.

## SCORES GIVEN BY THE JUDGES FOR THE STUDENTS IN SEMESTER III

|  | JUDGES (J) SCORE |  |  |
| :---: | :---: | :---: | :---: |
| STUDENT NO. | $J_{1}$ | $J_{2}$ | $J_{3}$ |
| 303 | 3.5 | 4.0 | 5.0 |
| 304 | 4.5 | 4.5 | 5.0 |
| 305 | 6.5 | 6.0 | 6.0 |
| 306 | 7.0 | 6.0 | 5.0 |
| 307 | 7.0 | 6.0 | 6.0 |
| 308 | 6.0 | 6.0 | 6.0 |
| 309 | 4.0 | 6.0 | 5.5 |
| 310 | 5.0 | 5.0 | 4.0 |
| 311 | 6.0 | 6.0 | 7.0 |
| 312 | 6.5 | 6.0 | 6.0 |
| 313 | 5.0 | 5.5 | 4.0 |
| 314 | 5.75 | 5.5 | 5.38 |
| $\bar{X}$. |  |  | 5.0 |

Figure 1

APPENDIX L

## STEPWISE MULTIPLE REGRESSION ANALYSIS

The BMD:02R computer program was employed to do the stepwise multiple regression analysis. This program printed out the following statistics: the mean and standard deviation of each variable; the covariance and correlation matrices; the analysis of variance table as each independent variable entered the regression equation; the multiple correlation coefficient ( $R$ ) and its standard error of estimate for each step; the intercept constant, the partial regression weights (beta weights), their standard error and level of significance; and finally, the coefficients of determination $\left(R^{2}\right)$, the semipartial correlations and their level of significance.

The coefficient of multiple correlation ( $R$ ) is defined by Kerlinger and Pedhazur (1973) as the product moment correlation of the predicted variable ( $Y^{\prime}$ )--which is a linear combination of the independent variables $(X$ 's)--and the observed $Y$. The values of $R$ range from zero to one, and when squared ( $\mathrm{R}^{2}$ ), and multiplied by 100 yield the percent of variance of the dependent variable ( $Y$ ) which is explained by the linear combination of the independent variables.

The omnibus F statistic is calculated as each independent variable enters the regression equation. This statistic is calculated by utilizing the summary data from the ANOVA table. The F statistic was calculated with the following formula:

$$
F=\frac{S S_{\text {regression }} / \mathrm{df}_{1}}{S S_{\text {residual }} / \mathrm{df}_{2}}
$$

where $\mathrm{df}_{1}=$ degrees of freedom associated with sums of squares of regression which is equal to $K$, the number of independent variables in the regression equation.
$d f_{2}=N-K-1$, where $N$ is equal to the sample size

This F statistic was used to test the significance of the multiple correlation coefficient ( $R$ ), as each new variable entered the regression equation. In addition, at each step the standard error of estimate ( $\mathrm{SE}_{\mathrm{est}}$ ) of R was calculated, which is the standard deviation of the residuals in Y at a particular step in the analysis.

The formula for the $S E_{\text {est }}$ was as follows:

$$
\sqrt{\frac{S S_{\text {residual }}}{N-K-1}}
$$

This statistic is an index of the dispersion of the predicted $Y$ measure ( $Y^{\prime}$ ) about the regression line, and is used to evaluate the estimate of the dependent variable $(Y)$ based on the independent variables. For example, if the $S E_{\text {est }}$ of $R$ is larger when compared to the standard deviation of the dependent variable $(Y)$, then the estimate of $Y$ on the basis of the independent variables is poor. The beta weights ( $b$ 's) are the sample partial regression coefficients which indicate the relative importance of the independent variables ( $X^{\prime}$ s) in making the predictions to the dependent variable (Y). The standard error of regression weights is the statistic which indicates the variability of the errors and provides a measure with which to compare the statistics whose significance are being tested. The formula used to calculate the latter is:

$$
S E_{b_{1}}=\sqrt{\frac{\left(\mathrm{SE}_{\mathrm{est}}\right)^{2}}{S S_{\mathrm{X}_{1}\left(1-R^{2} 1.23\right)}}}
$$

where $\begin{aligned} . \mathrm{SE}_{\mathrm{b}_{1}}= & \begin{array}{l}\text { Standard error of beta weight of independent variable } \\ \text { number } 1\end{array} \\ \mathrm{SS}_{\mathrm{X}_{1}}= & \text { Sums of squares of independent variable number } 1\end{aligned}$

The partial correlation is used when statistical control of variance is desired. At each step of the multiple regression analysis, the partial r's are calculated which show the correlation between two variables with the third variable held constant.

Finally, the semipartial correlation is calculated. Unlike the partial $r$ 's, the semipartial r's represent the correlation between two variables with the influence of another variable(s), removed from only one of the variables being correlated. More specifically, the square of this statistic tells the unique contribution to the variance of the dependent variable that each independent variable adds after the variance contribution of the preceding variable(s) is accounted for; however, this applies only to the particular order in which the independent variables enter the regression equation.


[^0]:    *Vancouver Community College, Langara 1977-1978 Calendar.

[^1]:    I If a student was absent for a week, her score for the three weeks combined was determined by averaging her scores of the other two weeks.

[^2]:    *Expected variances of negative estimates of variance components are set to zero.

