1984

National Certification As A Performance Measure: Pediatric Nurse Practitioners, 1977-1982

Barbara Hall Dunn
bhdunn114@comcast.net

School of Social Work
Virginia Commonwealth University

This is to certify that the dissertation prepared by Barbara Hall Dunn entitled National Certification as a Performance Measure: Pediatric Nurse Practitioners, 1977-1982 has been approved by her committee as satisfactory completion of the dissertation requirement for the degree of Doctor of Philosophy.

[redacted]

Lynn D. Nelson, Ph.D., Committee Chairman

[redacted]

Marilyn A. Biggerstaff, D.S.W., Committee Member

[redacted]

JoAnne K. Henry, Ed.D., Committee Member

[redacted]

Neil W. Henry, Ph.D., Committee Member

[redacted]

Barbara A. Munjas, Ph.D., Committee Member

[redacted]

L. Diane Bernard, Ph.D., Faculty Coordinator,
Doctoral Program, Social Policy and Social Work

[redacted]

David P. Fauri, Ph.D., Associate Dean and Director,
Doctoral Program, Social Policy and Social Work

[redacted]

Grace E. Harris, Ph.D., Dean, School of Social Work

April 10, 1984

April 10, 1984

National Certification as a Performance Measure: Pediatric Nurse Practitioners, 1977-1982

A dissertation submitted in partial fulfillment of the requirements for the degree Doctor of Philosophy at Virginia Commonwealth University

By
Barbara Hall Dunn
BSN, Virginia Commonwealth University, 1970
MSN, University of Virginia, 1974

Committee Chairman: Dr. Lynn D. Nelson
Associate Professor of Sociology & Anthropology
Virginia Commonwealth University
Richmond, Virginia
April, 1984

Acknowledgments

No one completes a doctoral program or a dissertation without help from many people.

I am particularly indebted to my dissertation committee--Drs. Marilyn Biggerstaff, JoAnne Henry, Neil Henry, Barbara Munjas, and Lynn Nelson --for making the process as painless as possible. I am also indebted to Lynn Nelson, my committee chairman, for his skills as a facilitator and master teacher. To Neil Henry, my statistical advisor, my gratitude for many, many hours of patient discussion and explanation, with the dissertation and other doctoral course work.

This research was conducted in cooperation with the National Board of Pediatric Nurse Practitioners and Associates (NBPNPA), Rockville, Maryland, and its testing agency, the National Board of Medical Examiners (NBME), Philadelphia. I am obligated to the NBPNPA Executive Board for granting me access to the data base used in this study. To Nancy Dickenson-Hazard, NBPNPA Executive Director, and Nancy Cohen, Executive Secretary, I extend my appreciation for support and assistance with data collection. I also recognize the assistance of Dr. Anita Bell Campbell, Senior Psychometrician, who was my liaison with the NBME.

For their cooperation in providing information and unpublished research, I would like to thank the American Association of Nurse Anesthetists, the American College of Nurse Midwives, and the NAACOG Certification Corporation. In addition, unpublished data regarding the Longitudinal Study of Nurse Practitioners was obtained from the State University of New York at Buffalo. For those data I extend my appreciation to Dr. Harry Sultz and his co-investigators.

At Virginia Commonwealth University's Academic Computing Center, I received assistance with data processing from Vernon Williams, Senior Programmer Analyst, and Toby Trott, Programmer Analyst. Their tutorials and debugging were invaluable.

Dr. Diane Bernard, faculty coordinator of the doctoral program, has been a staunch advocate for all students in the program, for which I am especially grateful. I would also like to acknowledge the contributions of my fellow students (present and former) to my doctoral education.

Finally, thanks to my sister, Deborah Dunn, for help in typing the various drafts of this document, and to my family and friends, whose support I depend on always. This dissertation is dedicated to the memory of my grandmother, who never doubted.

Table of Contents

List of Tables

Table Page

List of Abbreviations

AAP--American Academy of Pediatrics

AANA--American Association of Nurse Anesthetists

ACNM--American College of Nurse Midwives

AFPNA/PP--Association of Faculties of Pediatric Nurse Associate/ Practitioner Programs

ANA--American Nurses Association

DHEW--Department of Health, Education and Welfare

DHHS--Department of Health and Human Services

GMENAC--Graduate Medical Education National Advisory Committee

IOM--Institute of Medicine (National Academy of Sciences)

MCAT--Medical College Admissions Test

NAACOG--Nurses Association of the American College of Obstetricians and Gynecologists

NAPNAP--National Association of Pediatric Nurse Associates and Practitioners

NBME--National Board of Medical Examiners

NCHCA--National Commission for Health Certifying Agencies

NLN--National League for Nursing

NQE--National Qualifying Examination (National Board of Pediatric Nurse Practitioners and Associates)

SAS--Statistical Analysis System

Abstract

NATIONAL CERTIFICATION AS A PERFORMANCE MEASURE: PEDIATRIC NURSE PRACTITIONERS, 1977-1982

Barbara Hall Dunn, Ph.D.

Virginia Commonwealth University, 1984

Committee Chairman: Dr. Lynn D. Nelson

This research involved an analysis of data for 3,387 candidates who took the National Qualifying Examination for pediatric nurse practitioners/associates between 1977 and 1982. Those data were available from the National Board of Pediatric Nurse Practitioners and Associates, which administers the examination, and its testing agency, the National Board of Medical Examiners. Included in the data were sociodemographic characteristics of examinees, characteristics of their nurse practitioner educational programs, and their composite examination scores.

The purpose of the research was to determine the ability of these sociodemographic and educational program variables to predict examination performance. Sociodemographic variables included: examinee age; highest level of education; months of experience as a registered nurse and as a nurse practitioner; formal or informal preparation as a nurse practitioner; current function (nurse practitioner skills or not); employment setting; year of examination; and, status as a first-time examinee or repeater. The educational program variables available for study were: current program status (operational or not); educational level (certificate or masters); institutional setting/sponsors;

accreditation status; administrative control; discipline of program director(s); year established; class size; and length in hours and weeks.

The data were analyzed by descriptive and multivariate techniques. There were statistically significant differences in the sociodemographic and program profiles of examinees from year to year. Between 1977 and 1982 the average age and, consequently, the length of experience of examinees decreased. Their highest level of education increased, and there were a larger proportion of masters programs and masters program graduates. In terms of functions and settings, the number of examinees who were not functioning as nurse practitioners increased, as did the number who were unemployed. Over this 6 year period, educational programs have moved into the mainstream of nursing education: they are typically located in schools of nursing that are accredited by the National League for Nursing, with administrative control vested in nursing and with a nurse director or nurse and physician co-directors. Additionally, programs have increased in both length in hours and in weeks.

Regression analysis was used to evaluate the relationships between examination scores and the various predictor variables. An exploratory analysis using stepwise regression procedures eliminated those variables with little predictive significance. Further analyses with the five sociodemographic and five program variables remaining in the regression equations indicated that the largest contributions to differences in examination scores were made by the following variables: examinees' status as first-time takers or repeaters, their highest education, their age, and the educational level and accreditation status

of their nurse practitioner program.

At the individual level of analysis, sociodemographic variables were better predictors than program variables and explained between 8-26% (R=.28 to .51) of the variance in examination scores. On the other hand, the program variables explained 28% (R=.53) of the variance in average performance from program to program. That is, at the aggregate level of analysis (by program) there is obviously less individual variation around the program means and, therefore, greater predictive ability.

Based on the results of this research the investigator made recommendations regarding educational and regulatory policy and suggestions for further research. In particular, further research on certification in nursing was encouraged.

Chapter I. Problem Statement

Economics is the force driving health policy in 1984, and for the forseeable future. With expenditures for health care accounting for 10.4% of the gross national product (Davis, 1983a), cost containment has become the national priority for health care.

Educational institutions that prepare health professionals are faced with budget reductions and an unstable federal funding future. As a result, they are reassessing their programs and priorities. Nurse practitioner programs are under intense scrutiny for several economic and political reasons, including the expense of conducting them. It was estimated, for example, that in 1979 the average annual federal cost per student in nurse practitioner programs was $12,900, compared to $14,200 for medical students and $5,262 for undergraduate nursing students (LeRoy & Solkowitz, 1981, p. 17). (No data are available regarding the cost of masters programs in nursing.)

Although two recent studies recommended continued federal support for nurse practitioner programs (Institute of Medicine (IOM), 1983, p. 16; Report of the Graduate Medical Education National Advisory Committee (GMENAC), 1981a), the future of federal funding is uncertain. Total federal appropriations for nurse practitioner education have decreased in recent years and, with potential evaporation of that source, more than 50% of these programs are concerned with survival. In addition,

state budget reductions raise the possibility of lack of funding for those programs in state-supported institutions. The expense of nurse practitioner education, fewer extramural dollars, and a general concern about the supply and regulation of health professionals pose significant questions for policymakers at the institutional, state, and federal levels.

These questions include: Should the preparation of nurse practitioners be continued? If so, how many are needed and at what level (certificate/masters) should they be prepared? If certificate education is continued, should admission be restricted to registered nurses with certain educational/experiential backgrounds? How can program costs be reduced without sacrificing quality? Can the length of programs be shortened by reducing content--classroom, clinical, or preceptorship? Can the number of students per class and student-faculty ratios be increased? Can the number of faculty, particularly expensive physician faculty be decreased? Do schools of nursing need nurse faculty who are prepared as practitioners? How should the practice of nurse practitioners be regulated?

Assuming that decision makers want to continue to prepare nurse practitioners, they are left with the task of determining how to do this in the future. One way to determine "how" is to look at available data on the postgraduate performance of nurse practitioners and see whether it is related to particular student or program characteristics. Variation in achievement (scores) on national specialty certification examinations is one such performance measure. In fact, it is the only standard performance measure available for nurse practitioners.

Although it is recognized that relationships between performance

on a cognitive examination and in actual clinical practice are equivocal, certification by examination has become an institutionalized indicator of competence to practice in a particular specialty area. Additionally, national specialty certification is being used more frequently to assure various publics--employers, consumers, state and federal regulatory agencies, third-party payors--that certain predetermined skills and knowledge have been mastered by individual health professionals.

Education and certification are related. Educational institutions are responsible for providing a learning experience in accordance with stated objectives, and for assuring that those objectives have been achieved. The granting of a degree or certificate is recognition of that achievement. On the other hand, certification mechanisms are responsible for assuring that those individuals who have been awarded a degree or certificate possess the knowledge and skill required to function at subsequent levels of professional responsibility (Report of the Committee on Goals, 1973, p. 25).

While certification examinations "should not be designed or used to confirm that educational objectives have been achieved," they should "provide validation that the individual who has met the institution's requirements is competent to assume new responsibilities for patient care" (Report of the Committee on Goals, 1973, p. 26). Relationships between intramural and extramural evaluation are depicted in Table 1.

If a school of nursing claims that its students are educated so as to be eligible to apply for state licensure, then its curriculum must reflect licensure requirements. Likewise, if a nurse practitioner program claims that its graduates will be eligible to apply for state

Table 1
Evaluation in the Continuum of Professional Education

Evaluation	Beneficiary	Focus	Purpose	Application
Intramural-educational institution				
Learning	Student	Learning	Guidance	In-Course
Academic achievement	Institution & student	Mastery	Promotion & placement	End of course & end of year
Extramural-external agency				
Licensure	Public	Competence	Author-ization	Entry to professional practice
Certification				Entry to specialty practice

Note. Adapted from Evaluation in the Continuum of Medical Education, Report of the Committee on Goals and Priorities of the National Board of Medical Examiners, Philadelphia, 1973, p. 22.

and national specialty certification, its curriculum must reflect those requirements (Hinsvark & Dorsch, 1979, p. 374).

A review of the literature reveals little research in nursing related to certification, and a limited amount of research on this topic in medicine.

Purpose

This research involved an analysis of data on pediatric nurse practitioners (PNPs) who had taken the National Qualifying Examination for pediatric nurse practitioners/associates (the terms practitioner and associate are synonymous). Those data were available from the National Board of Pediatric Nurse Practitioners and Associates, which administers the examination, and its testing agency, the National Board of Medical Examiners. Included in the data were sociodemographic characteristics of examinees, characteristics of their nurse practitioner educational

programs, and their examination scores. The investigator sought to determine whether those sociodemographic and educational program characteristics were related to examination performance.

This information will assist nurse educators and administrators in making decisions about admissions policy and design and organization of nurse practitioner programs. It may also assist other policymakers in decisions regarding the types of programs to fund in the future. This research was also conducted to develop new knowledge about certification as a performance measure in nursing, and to add to the general body of knowledge on nurse practitioner performance as it relates to relevant predictor variables.

Objectives

The objectives of this research were to: (a) develop a sociodemographic profile of pediatric nurse practitioners who sought certification through the National Board's mechanism; (b) create a data file on characteristics of educational programs preparing pediatric nurse practitioners; (c) design a model to explore relationships among sociodemographic characteristics of examinees, educational program characteristics, and examination performance; and (d) determine the model's ability to predict performance on the National Qualifying Examination for pediatric nurse practitioners.

Variables

The major variables under consideration in this research were:

1. Sociodemographic characteristics of examinees--sex, age, highest level of education, type of nurse practitioner preparation, months of experience as a registered nurse, months of experience as a nurse practitioner, and current employment setting and job function.

2. Nurse practitioner program characteristics--level of program, current status of program (active/inactive), location and setting of program, administrative control, discipline of program director(s), accreditation status, number of students per class, length in hours and weeks, and year program established.

3. Examination performance--standard composite scores.

Scope

The scope of this research was limited to analysis of data on pediatric nurse practitioners, their educational programs, and their performance on a national specialty certification examination. Chapter II describes the context of the problem in terms of nurse practitioners, relevant policy-related issues, and credentialing.

Chapter II. Problem Context

To put this problem into perspective, information regarding nurse practitioners, policy-related issues, and credentialing is reviewed.

Nurse Practitioners

In 1965 the first nurse practitioner program in the country, for pediatric nurse practitioners, was developed at the University of Colorado. That short-term continuing education curriculum was designed to prepare registered nurses to assume greater responsibility for care of patients in primary care settings, in areas that had traditionally been the province of physicians. Initially the purpose of preparing nurse practitioners was to increase the access to and availability of care for rural and other medically underserved populations. Within a relatively short period of time, the nurse practitioner curriculum was integrated into the mainstream of nursing education, such that today almost 60% of these programs are offered at the masters level (Sultz, Henry, Kinyon, Buck, & Bullough, 1983a).

Based on projections of the Division of Nursing, Department of Health and Human Services (DHHS), it is estimated that there are between 22,000-24,000 graduates of formal nurse practitioner programs in 1984 (Report of the GMENAC, 1981b, pp. 17-21; Sultz, Henry, Kinyon et al., 1983a). (There are no estimates of the number of informally trained nurse practitioners.) Of that number, 75% are either pediatric, family,

or adult nurse practitioners, equally distributed. The remaining 25% are a combination of obstetric-gynecologic, geriatric, emergency, and other subspecialty nurse practitioners (Report of the GMENAC, 1981b, pp. 18-21). Pediatric nurse practitioners are the prototype group in terms of design of educational programs, role development in practice, and establishment of national specialty certification mechanisms.

Related organizations. There are four major membership organizations relating to pediatric nurse practitioners. They are the American Nurses Association (ANA), the American Academy of Pediatrics (AAP), the National Association of Pediatric Nurse Associates and Practitioners (NAPNAP), and the Association of Faculties of Pediatric Nurse Associate/Practitioner Programs (AFPNA/PP).

In 1971 the ANA and the AAP jointly developed and issued the first guidelines for educational programs preparing pediatric nurse practitioners, including behavioral objectives, curriculum content, and academic and organizational structure (Guidelines, 1971). Two years after the formation of the AFPNA/PP in 1973, more specific behavioral objectives and curriculum content were developed. In its most recent publication (1982), that association revised behavioral and curriculum objectives and issued a position statement supporting programs in accredited graduate schools of nursing, with the practitioner optimally prepared at the masters level.

Educational programs. The most comprehensive data about nurse practitioners and their educational programs are found in the Longitudinal Study of Nurse Practitioners, phases I, II, and III (Sultz, Bullough, Kinyon, Buck, & Sherwin, 1983; Sultz, Henry, Bullough, Buck, & Kinyon, 1983; Sultz, Henry, Kinyon et al., 1983a, 1983b; Sultz & Kinyon,

1976; Sultz, Zielezny, Gentry, & Kinyon, 1978, 1980). This study was funded by the Division of Nursing, DHHS, and data were collected in 1973, 1977, and 1980.

The 1980 data indicated that there were 83 certificate programs for nurse practitioners and 116 masters programs. Of those, 43 programs were for pediatric nurse practitioners (21.6%); 22 certificate and 21 masters level (Sultz, Henry, Kinyon et al., 1983a). All masters programs were sponsored by colleges or universities, and two pediatric certificate programs were sponsored by hospitals; 13 other programs were sponsored by hospitals or voluntary nonprofit agencies (Sultz, Bullough et al., 1983).

The organization of programs has been influenced by the source of funding, particulary for those programs receiving grant support from the federal government (Kahn, 1979). In 1976 the Division of Nursing, Department of Health, Education, and Welfare (DHEW; now DHHS), issued guidelines for those programs applying for grant support. Requirements included: affiliation with a collegiate school of nursing, medicine, or public health; program length of at least one academic year (nine months); minimum student enrollment of eight per class; and, curriculum to include classroom and clinical instruction with an optional preceptorship (DHEW, 1976).

The Sultz data indicated the following sources of funding for 199 programs surveyed in 1980: (a) federal (70%)--Division of Nursing (50%), capitation funds (7%), other: National Institute of Mental Health, Public Health Service, Veterans Administration, National Health Service Corps, military (13%); (b) non-federal (37%)--state (21%), other: university funds and foundations, including Robert Wood Johnson, Kellogg,

March of Dimes, Noyes (16%) (Sultz, Bullough et al., 1983).

Most programs, in 1980, were found in National League for Nursing (NLN) accredited schools of nursing; 22 programs (5 pediatric) were accredited by the American Nurses Association (ANA, 1982, pp. 12-13). These programs were most likely to be located in the northeastern or western regions of the country; 66% of the masters programs and 50% of the certificate programs that were active in 1980 were initiated during or after 1974 (Sultz, Bullough et al., 1983; Sultz, Henry, Kinyon et al., 1983a).

Program directors were usually nurses; 93% of certificate programs reported that a nurse was the primary program director, and 46% of masters programs reported that there was also a physician co-director (Sultz, Henry, Kinyon et al., 1983b). Certificate level programs averaged 2.9 full-time faculty (2.2 pediatric), and masters programs averaged 3.2 (2.5 pediatric). When faculty were divided by discipline, there were 2.3 nurse practitioners in certificate programs (1.6 pediatric) and 2.5 in masters programs (1.9 pediatric). Physician faculty averaged 0.3 and 0.2 for certificate and masters programs respectively (Sultz, Bullough et al., 1983). While certificate programs averaged 5 students for every nurse practitioner faculty member (5:1) and 41 students for every physician faculty member (41:1) (25:1 pediatric), masters programs had one nurse practitioner faculty member for every four students (4:1) and one physician faculty member for every 58 students (58:1) (37:1 pediatric) (Sultz, Henry, Kinyon et al., 1983a).

Most programs admitted one class per year, with an average class size of 8-9 for certificate and 12-13 for masters programs. Programs ranged in length from 4-22 months for certificate and from 9-24 months

for masters (9-22 pediatric). The average length of programs was 11.7 months, certificate (11.3 pediatric) and 16.1 months, masters (14.9 pediatric) (Sultz, Henry, Kinyon et al., 1983a).

Athough the masters programs were longer, the total number of hours spent in classroom, clinical, and preceptorship was greater for certificate programs. Certificate programs averaged 430 hours of classroom instruction (325 pediatric), 353 hours of supervised clinical practice (292 pediatric), and 524 hours of preceptorship (399 pediatric), for a total of 1,307 hours (1,016 pediatric). (Total hours for each component were derived from hours per week times the number of weeks as reported by program directors.) Eighty-nine percent (89%) of these programs (91% pediatric; 20/22) reported a preceptorship requirement (Sultz, Bullough et al., 1983; Sultz, Henry, Kinyon et al., 1983b).

The masters programs averaged 326 hours of classroom instruction (397 pediatric), 390 hours of clinical practice (347 pediatric), and 236 hours of preceptorship (265 pediatric), for a total of 1,052 hours (1,009 pediatric). While 60.3% of all masters programs reported a preceptorship requirement, 66.7% of the pediatric programs have this requirement (14/21) (Sultz, Bullough et al., 1983).

Educational trends. Between 1973 and 1980 the total number of certificate programs decreased (particularly pediatric programs) and the number of masters programs increased. All masters and most certificate programs were located in college or university settings, and were directed by a nurse or had nurse and physician co-directors. Although the largest number of programs were located in the south in 1973, in 1980 more programs were found in the northeast and west (Sultz, Bullough et al., 1983; Sultz, Henry, Kinyon et al., 1983a, 1983b).

The most common designation for these programs was "nurse practitioner" program, rather than nurse associate, nurse clinician, clinical nurse specialist, or any other title. While there has been no appreciable increase in the length of masters programs, the certificate programs have increased from an average of 8.5 months in 1973 (7.9 pediatric) to 11.7 months in 1980 (11.3 pediatric). Paradoxically, students in certificate programs spent more total hours in their programs.

There has also been an increase in the number of masters programs that require preceptorships; a decrease in the number of classes offered per year, and a related increase in the number of students per class with larger student-faculty ratios. Finally, there seemed to be less reliance on physician preceptors in the programs and greater numbers of nurse faculty who maintained their own clinical practice (Sultz, Henry, Kinyon et al., 1983a, 1983b).

Student characteristics. Of more than 1,500 nurse practitioner students surveyed in 1980, most were married (52%; 56.4% pediatric), white (93%; 91% pediatric), women (94.8%; 98.6% pediatric), between the ages of 25-34 (65.8%; 60.1% pediatric). The average age for all nurse practitioners was 32.9 years; 34.1 years for pediatric students. While all students in masters programs presumably entered the program with at least a bachelors degree, certificate students' prior nursing education was varied. Certificate students were divided as follows: 33.4% with diplomas (40.8% pediatric); 24% with associate degrees (7.2% pediatric); 35.6% with bachelors degrees (42.8% pediatric); 6.9% with masters degrees (9.2% pediatric); and 0.1% with doctoral degrees (Sultz, Bullough et al., 1983; Sultz, Henry, Bullough et al., 1983).

Students usually had between 1 and 5 years of experience in nursing prior to entering the program, and that experience was most likely to have been in an inpatient hospital setting. The average number of years experience was 8 (9.5 pediatric); 9.5 years for certificate students (11.9 pediatric) and 6.8 years for masters students (same for pediatric). When asked whether they were members of the American Nurses Association, the 1980 cohort typically answered "no" (36.3% were members). As might be expected, more masters (47.6%) than certificate (22.5%) students were members of ANA (Sultz, Bullough et al., 1983; Sultz, Henry, Bullough et al., 1983).

Financial aid to students, like funding for programs, most often came from federal sources. Fifty-seven percent (57%) of the programs indicated that they received federal student aid (36% from the Division of Nursing, DHHS; 21% from other sources). Non-federal sources of student aid were received by 27% of the programs (8% state; 1% city/local; 17% other) (Sultz, Bullough et al., 1983).

Student trends. In comparing the 1973 and 1980 data, the following trends were noted. Students were younger (average 32.9 years versus 35.2 years), and consequently had less experience when they entered the program (averaged 8 years versus 9.4 years). While that experience was most likely to be in an inpatient hospital setting, as it was in 1973, the tendency was more pronounced (47.9% versus 36.2%) (Sultz, Bullough et al., 1983; Sultz, Henry, Bullough et al., 1983).

Fewer students were married (52% versus 55.3%), and the number of nonwhite students has decreased (7% versus 10.1%). While there tend to be more black students in certificate programs than in masters programs, that discrepany has lessened between 1973 and 1980 (5% more blacks in

certificate programs than masters in 1973, 2.3% more blacks in certificate programs in 1980). There were more than twice as many men in nurse practitioner programs in 1980 than in 1973 (5.2% versus 2%), although the numbers remained small (Sultz, Bullough et al., 1983; Sultz, Henry, Bullough et al., 1983).

The trend toward baccalaureate education in nursing was evident (67.8% versus 53.3%), with fewer students entering certificate programs with diplomas in nursing (33.4% versus 46.7%) and more students entering masters programs in general. Finally, students were less likely to be members of ANA in 1980 (36.3%) than they were in 1973 (45.4%) (Sultz, Bullough et al., 1983; Sultz, Henry, Bullough et al., 1983).

Policy-Related Issues

The national policy issues most germane to this area of research are primarily economic and political. That is, they are questions related to health manpower, regulation of health occupations and the industry as a whole, and the spiraling cost of health care. Although economic forces are driving health policy, there are conflicting interpretations of what the problems are and conflicting opinions about alternatives to solve the problems.

Health manpower and regulation. In the 1960s, there was a perceived physician shortage and concern about the availability of and access to care for rural and other medically underserved populations. The real problem was maldistribution of physicians by specailty and geographic location (Lee, LeRoy, Stalcup, & Beck, 1976, p. 18), with physicians prefering specialty practice and metropolitan locations. Although most problems for which patients seek care are minor, by 1970 80% of physicians were specialists and only 20% were in general primary

care practice (Roemer, 1977, p. 60).

When it became apparent that the demand for services could not be met by physicians (Fuchs, 1974, pp. 68-69), the training and utilization of nonphysician primary care providers was proposed. Nurses were recognized as an underutilized resource who could assist in meeting primary care needs; therefore, in 1971 the DHEW Committee to Study Extended Roles for Nurses recommended increased use of nurses as providers (DHEW, 1971, p. 8).

In the 13 years since that recommendation was made, nurse practitioners have proven their ability to meet the primary care needs of various populations, providing care comparable to that of primary care physicians (Dunn & Chard, 1980). They are less expensive to educate than physicians, and can be prepared in a much shorter length of time. System costs are lower when nurse practitioners are used, because most are employees whose salaries are about one-third of the average primary care physician's salary. Evidence of reduced costs to the consumer, however, are less apparent (LeRoy & Solkowitz, 1981, pp. 21,24).

Economic, political, and legal barriers. Despite the advantages of nonphysician providers, economic and legal barriers to their practice have prevented their full utilization. These barriers include state regulations that usually require physician supervision, lack of consistency among states in regulatory control of practice, and limited third-party reimbursement for the services of nurse practitioners (Hutton & Rorabaugh, 1983; Lee et al., 1976, p. 119).

The economic and political climate of the 1980s is radically different than it was in the 1960s and 1970s. There is a general feeling that health care costs are out of control, and the emphasis at

the federal level is on cost containment and shifting responsibility for programs from the public to private sector (Davis, 1983a). In the single most significant health financing legislation since the enactment of Medicare in 1965, federal policymakers have altered the mechanism for reimbursement of hospitals under Medicare (Social Security Amendments of 1983; P.L. 98-21). This is the forerunner of other proposals that are designed to increase competition in the health care industry and offer incentives for cost containment (Davis, 1983a).

While increased competition and efforts to contain costs might appear to support the continued use of nonphysician providers, there are complicating factors that make the actual outcomes uncertain. In its attempts to lower federal expenditures for health, the federal government has reduced allocations to educational institutions for the training of health professionals. It has also reduced allocations for federally funded programs, such as family planning and maternal-child health, and shifted the responsibility for those programs to the states. Thus, there may be fewer educational programs preparing nurse practitioners, and fewer public settings employing nurse practitioner program graduates.

Exacerbating these problems are projections that predict a surplus of primary care physicians by 1990 (Report of the GMENAC, 1981a). In reaction to these projections, there is evidence that physician groups are stepping up their efforts to reduce competition with nonphysician groups. Their methods include: attempts to amend state statutes or regulations authorizing the practice of nurse practitioners or to make those regulations more restrictive; individual lawsuits against nurse practitioners for "practicing medicine without a license;" denial of

hospital privileges to nurse practitioners and nurse midwives; and, efforts to intimidate physicians who employ, supervise, or are supportive of these providers (Dunn & Brown, 1982b; Pollard & Schultheiss, 1983).

Many observers believe that nurse practitioners and other non-physician providers must be reimbursed by third-party payors, as a matter of economic survival, and 13 states currently provide for direct reimbursement of certain nurses (Mezey, 1983). However, the federal government has just begun to address the problem of fee-for-service reimbursement of physicians, proposing changes in the rate of reimbursement and encouraging prepaid health plans (such as health maintenance organizations) as an alternative to the private practice, fee-for-service model. Federal policymakers are not likely to extend reimbursement to other types of providers until they have dealt with physician reimbursement. In addition, extending reimbursement usually means the introduction of new services rather than substitution of one provider for another. Therefore, more longitudinal research is needed to determine the costs and benefits of reimbursing other provider groups (Davis, 1983a, 1983b).

Nursing and nursing education. Finally, there are factors that relate specifically to nursing and nursing education. In periods of economic constraint, fewer students enroll in graduate nursing programs on a full-time basis (IOM, 1983, p. 150). Schools of nursing that offer nurse practitioner minors in their masters programs have found that this is one way of attracting students. Students making decisions about continuing education are increasingly attracted to masters programs, because there is relatively little difference in the length of certifi-

cate and masters nurse practitioner programs and they earn academic credit and a degree for their efforts (some certificate programs offer continuing education credit rather than academic credit).

With organized nursing supporting the baccalaureate degree in nursing as the qualification for entry into professional nursing, it is logical that the masters degree would become the qualification necessary for specialty practice. In the past 5 years, the National League for Nursing (1979), American Nurses Association (1980), and the Association of Faculties of Pediatric Nurse Associate/Practitioner Programs (1982) have issued position statements in support of masters level preparation for nurse practitioners. Additionally, at least 5 state boards of nursing have stipulated masters level preparation as an eligibility requirement in obtaining state certification (Hutton & Rorabaugh, 1983).

The basic question in terms of economic and legal considerations is: Is this movement toward graduate education aimed at producing a better qualified nurse practitioner who will practice more competently than the certificate graduate, or is this a self-serving, arbitrary standard? If regulatory agencies (whether private or public) restrict state or national certification to those nurse practitioners who are masters prepared, they may leave themselves open to charges of "restraint of trade"--a violation of federal antitrust laws.

The Federal Trade Commission (FTC), which has jurisdiction in the area of anticompetitive business practices in the professions, has been keeping a watchful eye on certification mechanisms (including those for nurse practitioners) in recent years (Pollard & Leibenluft, 1981). Although there are no known lawsuits against state regulatory boards or

agencies or associations that certify nurse practitioners, the probability that this will occur becomes more likely as these groups introduce more restrictive requirements. This is an interstate problem, therefore falling within FTC purview, because restrictive and conflicting standards among states prevent practitioners from moving easily from one state to another to practice.

Credentialing

Credentialing is generally understood to involve three processes--licensure, accreditation, and certification. Licensure represents a set of legal requirements primarily concerned with public safety, while accreditation is concerned with evaluation of programs and institutions. As Passarelli notes, certification reverses the priorities by concentrating on the individual practitioner (1979, p. 79). These three credentialing functions are related. That is, eligibility for licensure and specialty certification usually includes program completion in an accredited educational institution. While licensure, which is the most restrictive form of occupational regulation, controls practice, certification is usually "title control" (Shimberg, 1982, pp. 15-17).

Health occupational credentialing has been a major topic of policy debate since the early 1970s. At that time there was a perceived physician shortage and massive federal funding of nonphysician provider programs was initiated. The federal government quickly realized that there were problems regulating quality with the influx of new categories of health providers, and that the state licensure system was inadequate to deal with the problem. In 1976, therefore, the Department of Health, Education, and Welfare recommended a program for national, non-federal certification of these new providers (Subcommittee on Health Manpower

Credentialing, 1976).

Nurse credentialing. In nursing, educational programs are accredited on the national level by the National League for Nursing (NLN). Since 1975, certain short-term continuing education programs, including several nurse practitioner programs, have been accredited by the American Nurses Association. Other groups involved in the accreditation or approval of programs include state boards of nursing and national certification boards. State boards "approve" schools of nuring in their determination of eligibility to sit for licensure exams. Likewise, national certification boards "approve" the educational programs of those persons applying for specialty certification.

All states require that registered nurses be licensed to practice, which includes passing a written examination administered by the regulatory agency in each state that is responsible for nursing or health occupations. Some jurisdictions also provide for state certification of certain nursing specialties; and, at least 14 states require national certification for nurse practitioners (Hutton & Rorabaugh, 1983). (Those nursing specialties most frequently regulated are nurse midwives, nurse anesthetists, and nurse practitioners.) For the most part, however, certification is understood to be a "process by which a nongovernmental agency or association grants recognition to an individual who has met predetermined qualifications specified by that agency or association" (Subcommittee on Health Manpower Credentialing, 1976, p. 1; emphasis added).

The first certification program in nursing was initiated in 1946 by the American Association of Nurse Anesthetists. With that exception, certification in nursing is really a phenomenon of the last decade. The

need for certification developed as nursing evolved from a generalist to specialist profession. It also evolved in response to federal initiatives making national specialty certification a condition for reimbursement under federal programs (DHEW, 1978; Subcommittee on Health Manpower Credentialing, 1976). In 1982 there were 32 national specialty certification mechanisms in nursing, offered in 28 clinical or functional areas by 14 different organizations (Dunn & Brown, 1982a).

Certification of nurse practitioners. There are currently three organizations that certify nurse practitioners. The ANA offers certification mechanisms for pediatric, school, adult, family, and gerontologic nurse practitioners, and the NAACOG (Nurses Association of the American College of Obstetricians and Gynecologists) Certification Corporation certifies obstetric-gynecologic nurse practitioners. Pediatric nurse practitioners are also certified by the National Board of Pediatric Nurse Practitioners and Associates (hereafter called the National Board). The duplication in certification mechanisms for pediatric nurse practitioners exists because of philosophical and political differences about the manner in which these practitioners should be certified. Of an estimated 5,600 pediatric nurse practitioners in the country, about 54% are nationally certified; 88% by the National Board and 12% by the ANA (IOM, 1983, p. 258; National Board, internal document, 1982).

Summary of Problem Context

In summary, a variety of economic and political factors will influence the future supply, training, regulation, and employment of nurse practitioners and other nonphysician providers. It is not clear what the future will be, but it is clear that we are witnessing dramatic

changes in health policy at the federal level--changes that will affect policy at the state, local, and institutional levels.

Chapter III reviews the literature related to research on measurement of performance, in-school classroom and clinical performance in the health occupations, postgraduate performance in nursing and medicine, and certification mechanisms for nursing and medicine.

Chapter III. Related Research

Prediction of academic performance and measurement of academic achievement are extensively researched areas of educational and psychological evaluation. A large volume of literature exists describing relationships, or the lack thereof, between various predictor variables and academic outcome measures. The literature reviewed includes general educational and psychological research, research on in-school academic and clinical performance in the health occupations, and research on postgraduate clinical performance and competency measurement in nursing and medicine.

Measurement of Performance

General findings. Tests are usually designed to measure an individual's aptitude (future oriented) or achievement (past/present oriented) in a particular content domain. The fundamental objective of achievement testing is generalization (Bejar, 1983, p. 18). Interpretation of test scores is part of the process of evaluating individuals; for achievement tests, it is also part of the process of evaluating educational curricula (Hopkins & Stanley, 1981, pp. 8,289,384; Mehrens & Lehmann, 1978, p. 528; Thorndike & Hagen, 1977, p. 191).

Thorndike and Hagen describe the social good that testing tries to achieve as: (a) protection from incompetence; (b) efficient use of resources (regarding selection and training procedures); (c) efficient

educational procedures (achievement measures as one indicator of the outcomes of education); (d) characterization of each person as an individual (rather than by group membership); and (e) contributing additional information to our knowledge about individuals (1977, pp. 620-622).

A taxonomy of cognitive behaviors that can be measured objectively has been described by Kane. Those behaviors are: knowledge, comprehension, application, analysis, synthesis, and evaluation (1980a, pp. 41-43). In general, however, strong predictors of academic achievement have not been found. This is due, in part, to the fact that the construct called "general ability" is often left out of prediction equations (Lavin, 1965, p. 19). Additionally, studies indicate that it is more difficult to predict performance for graduate school than for college. That is, with a more highly selected group, there is less variation in ability and, therefore, lower correlations (Lavin, 1965, pp. 50-51). This restriction in range on the criterion measure attenuates validity coefficients (Cullen, Dohner, Peckham, & Sampson, 1980, p. 263).

Noncognitive personality and sociodemographic variables that have been related to academic achievement include: positive relationships with independence, impulse control, introversion, positive self-image, higher socioeconomic status, female versus male, and urban/suburban versus rural students. Negative relationships have been found with increased age and anxiety (Lavin, 1965, pp. 43-44, 79-82, 132-133).

Extraneous factors. A variety of extraneous factors influence performance on cognitive tests. Among them are test sophistication, practice, coaching, anxiety and motivation, response styles or sets, and certain administrative factors (Hopkins & Stanley, 1981).

Test sophistication, or "test-wiseness", is defined as "an examinee's ability to use the characteristics and formats of the test and/or the test-taking situation to increase his score" (p. 141). In regard to practice, studies generally show improvement of test scores (10-20 points) on retest. This has been shown for the Scholastic Aptitude Test (SAT), the Graduate Record Examination (GRE), and the Medical College Admissions Test (MCAT). The effects are greater for people with limited educational background or experience with test-taking, for speeded tests, and for repeat rather than parallel forms of a test. There is usually no effect after the second retest, and there is little effect at all if the interval between the first and second tests is more than three months (pp. 143-144).

Coaching usually produces a small gain in scores, and is dependent on length and type of coaching. Although an inverse relationship between test scores and anxiety has been found, there is no evidence that this is a causal relationship. Response styles or sets are the test-taking habits that cause people of equal ability to score differently. These styles are known as: the speed versus accuracy set; the acquiescence set; the positional-preference set; the option length set; and the set to gamble (pp. 144-148).

In relation to the speed versus accuracy set, it has been shown that older people tend to work more slowly, which has led to gross overestimation of the effect of increasing age on performance (particularly on speeded tests). The acquiescence set describes the tendency of examinees to mark more true than false answers on true-false formats of a test, and the tendency of test item-writers to write more correct answers that are true (pp. 146-147).

Research regarding the positional-preference set indicates that test-takers, when they do not know the correct answer, do not randomly choose an option on multiple choice tests. Instead, they go through a particular decision-making process in selecting their answer. Examinees also tend to choose the longest option on multiple choice tests, which is called the option length set (p. 147).

The final style, known as the set to gamble, describes individual differences in the tendency to guess at answers. This tendency is consistent within and between tests, and most examinees can guess better than chance. Since most standardized tests do not correct for chance, gamblers have an advantage over non-gamblers in improving their test scores (p. 148).

Administrative factors such as the method of administration, the examiner, preannouncement, answer sheet format, scoring, disturbance during testing, and answer changing all affect test performance. Of these, the most interesting and relevant (to this study) research is related to answer changing. There is a widely disseminated myth that test-takers should stay with their first impression and not change answers. Research, however, indicates that test-takers are more likely to change wrong answers to right answers than the reverse (pp. 154-156).

Health occupations. In the health occupations, measurement of performance is complicated by the fact that there are three relevant domains of behavior to be examined--cognitive, affective, and psychomotor. Only the cognitive domain can be adequately evaluated by written tests. To provide information on the affective and psychomotor domains, clinical performance evaluations are routinely conducted. Unfortunately, these clinical performance evaluations are usually based on the

subjective ratings of faculty or supervisory personnel. It is unfortunate because comparisons are often made between performance on these dissimilar measures, which makes the research in this area difficult to interpret.

One of the challenges for credentialing mechanisms (licensure and certification) is to design appropriate test instruments that effectively measure the knowledge, skills, and professional attributes deemed essential for competent practice. Those areas are defined by selected educators and practitioners in a particular specialty ("expert consensus"), and are presented in a written examination that reflects a systematic method of instruction. Thus, credentialing examinations serve largely to validate levels of academic achievement (Report of the National Commission for Health Certifying Agencies (NCHCA), 1981, pp. 13-14).

Although there are no reliable studies that establish the extent to which credentialing and quality of care are interdependent, it is reasonable to expect that there should be a relationship between the competence of a health professional and the service that they provide (Passarelli, 1979, pp. 77,82). Before an attempt to measure competence can be made, however, it must be defined in terms of a particular specialty (Lloyd, 1980, p. 294).

For example, the American Board of Pediatrics has defined three dimensions of competence for pediatricians--subject matter, abilities, and tasks. The subject matter dimension is the clinical content of pediatric practice, that is, the most commonly encountered problems. There are five categories in the abilities dimension: attitudes, factual knowledge, interpersonal skills, technical skills, and clinical

judgment. In the tasks dimension, the categories are gathering, organizing, and recording data; assessing data; and managing problems and maintaining health (Burg, Brownlee, Wright, Levine, Daeschner, Vaughan, & Anderson, 1976).

As might be expected, these dimensions overlap with the blueprint used by the National Board for test item development for pediatric nurse practitioners (discussed in Chapter V). The subject matter dimension can be directly measured by cognitive examinations, as can factual knowledge and clinical judgment in the abilities dimension, and data gathering, organizing, and assessing, and management of problems in the tasks dimension. Therefore, cognitive examinations are one component in the assessment of competence (Salman, 1981, p. 9). These examinations have not, however, been "extended to provide a valid and reliable indicator of an individual's ability to apply knowledge to the care of patients" (Egan, 1982, p. 2933).

In-School Performance

Medicine. There appear to be no high magnitude, consistently significant relationships between preselection characteristics of medical students and their academic achievement in medical school. Undergraduate grade point average (GPA) is said to be the most accurate single predictor of medical school performance, with scores on the Medical College Admissions Test (MCAT) a less accurate predictor (DeVaul, Jervey, O'Keefe, & Short, 1982, p. 195). When these measures are used to predict clinical performance in medical school, the relationship is close to zero (Deighton, Smith, & Gallagher, 1979, p. 133; Gough, 1978; Murden, Galloway, Reed, & Colwill, 1977, pp. 181-186).

Freidman, Cheatham, Porter, and Bakewell (1979) looked at the

relationship between certain preselection characteristics of medical students (N=398) and their academic achievement in a medical school. In this study, academic achievement was designated satisfactory, unsatisfactory, or honors; the preselection characteristics were two interaction variables (undergraduate GPA x selectivity of the undergraduate school; undergraduate science GPA x total undergraduate science hours). Using discriminant analysis, they were able to predict 82% of those students in the satisfactory achievement group and 61% of those in the unsatisfactory group; however, they were unable to predict those in the honors group as a function of the preselection variables (pp. 145-147).

Using MCAT scores and undergraduate GPA as predictors, Gough (1978) found positive relationships with grades in the first 2 years of medical school, but no relationships with grades in the fourth year. Tucker and McGaghie (1982) designed a more elaborate study that related preselection variables to performance in the first 2 years of medical school. The dependent variables in this study were scores on end-of-year examinations for 655 medical students at one school. Predictor variables included: age, marital status, number of children, minority group membership, sex, undergraduate science GPA, adjusted undergraduate GPA (adjusted for the quality of the undergraduate school), number of hours of graduate level science classes, and MCAT scores.

Regression of these variables on examination scores explained 48% of the variance in scores on the first-year examinations and 38% of variance in scores on second-year examinations. Undergraduate science GPA and marital status (favoring those married) made the most significant contributions, while age, sex, minority group membership, and

number of children were not significant predictors when other factors were controlled. Tucker and McGaghie were particularly interested in the relationship between age and examination performance. In the regression equations described above, they found that the independent contribution of age to explanation of variance in examination scores was 0.16% in the first year and 0.28% in the second year.

Others who have examined the relationship between age and academic performance have determined that it is related to attrition and rank in medical school. Johnson and Hutchins (1966) found that, for academic reasons alone, the attrition rate for students between ages 23 and 33 increased from 6% to 11%, and was even higher for students over age 33 at the time of entry in medical school. Likewise, Conger and Fitz (1963) found that students over 24 at the time of entry had higher attrition rates and lower class ranks than younger students.

Daegenais and Rosinski (1975) looked at the relationship between the social class of medical students and certain cognitive (undergraduate GPA, MCAT scores, Miller Analogies scores, Parts I and II of the National Board of Medical Examiners (NBME) exams) and affective (personality and attitude scales) variables at one school (N=497). Their analysis of variance demonstrated that undergraduate GPA was the only variable significant across social class levels, and was inversely related (p. 202).

O'Donnell (1982) was interested in the personality type of medical students (N=99) and their performance on Part I of the NBME examination. He used the Myers-Briggs Personality Type Indicator to assess personality, and categorized students according to one of four personality types (sensing-thinking; sensing-feeling; intuitive-thinking; intuitive-

feeling). Other studies have indicated that sensing types consistently score lower on aptitude tests such as the American College Testing Program (ACT), SAT, and MCAT. Conversely, intuitive types score higher, and are more easily admitted to medical school. While intuitive types are estimated to make up 25-35% of the general population, they account for 52% of the medical student population (medical school is apparently particularly attractive to the intuitive-feeling type student). O'Donnell found, however, that 42% of the intuitive-feeling type students in his sample failed NBME Part I on their first attempt. Based on the fact that intuitive students are most easily admitted to medical school but have the greatest difficulty in passing the NBME examination, O'Donnell concluded that perhaps the NBME content or the curriculum content of the medical school needed reevaluation.

In relation to clinical performance in medical school, some researchers indicate that there are no consistent or strong relationships with preselection characteristics (Carline, Cullin, & Scott, 1982). There are conflicting findings reported about the relationship of age and clinical performance. While Benor and Hobfoll (1981) found that students in the 20-24 age group performed better in clinical, Carline, Cullin, and Scott found that older students scored higher on faculty ratings of clinical competence (1982, pp. 205-207).

Other research shows no relationship between clinical performance and MCAT scores or undergraduate GPA (Gough, 1978), and no significant relationship with preadmission interview ratings (Hobfoll & Benor, 1981). Murden and others, however, demonstrated that interview ratings, which are purported to assess characteristics such as maturity, are more important than GPA or MCAT scores in predicting clinical performance

(1977, pp. 181-186). Benor and Hobfoll (1981) conclude that, above some minimum threshold, academic achievement is of limited value in predicting clinical success.

Dawson-Saunders and Doolen (1980) have also studied the relationship between preselection characteristics and clinical performance. In their sample of 143 students at one medical school, they used nine preselection variables (four MCAT scores; science and nonscience undergraduate GPA; traditional-premedical/science undergraduate background versus nontraditional; age; sex) and faculty ratings of clinical performance along four dimensions (clinical sophistication; cognitive knowledge; personal maturity; communications skills).

In regression equations for the four dimensions of clinical competence, Dawson-Saunders and Doolen found that only nonscience GPA was significant in all four. They were able to explain 39% of the variance in ratings of clinical sophistication (nonscience GPA, science GPA, science MCAT, and quantitative MCAT significant) and 48% of the variance in cognitive knowledge (nonscience GPA, science GPA, general information MCAT, science MCAT, quantitative MCAT, and age--higher ratings for younger students--significant). Thirty-four percent (34%) of the variance in ratings of maturity was explained (nonscience GPA significant), and 37% of the variance in communications skill was explained (nonscience GPA significant). Using canonical redundancy analysis, they demonstrated that 16% of the variance in clinical performance was explained by the preselection variables. These relationships were generally positive but low to moderate in magnitude. They conclude that this might be expected, because of the 3 year interval between prediction measures (at entry to medical school) and clinical performance measures

(third year of medical school). That is, that you might expect that the correlation between preselection variables and clinical performance would decrease over time (1980, pp. 246-248).

Beneson, Stimmel, and Aufses (1981) looked at concordance between the surgical clerkship performance of medical students (as rated by faculty) and their surgical subtest scores on Part II of the NBME examination. Through discriminant analysis, they were able to accurately predict those students given honors ratings in their clerkship only 39.2% of the time (that is, those students scoring in the top 18% on the NBME examination who were also rated in the top 18%--honors--on clinical performance).

Finally, Raffeto and Zabarenko (1979) designed a paracognitive evaluation form to rate clinical skills and four related areas for medical students. They found that summary ratings by faculty on this evaluation instrument were not significantly related to scores on Part I or Part II of the NBME examination, and had a significant but low (r= .10) relationship with MCAT scores. Since the instrument was designed to measure noncognitive clinical performance, they concluded that the lack of correlation was evidence that the measures did in fact assess different domains of behavior.

In summary, research on in-school academic and clinical performance of medical students indicates that it is difficult to find significant relationships with predictor variables, both cognitive and noncognitive, that are consistent and of high magnitude. This is particularly evident when measures of cognitive performance are used to predict clincial performance in medical school. This is due, in part, to the importance of noncognitive attributes in clinical practice and in the assessment of

clinical competence by faculty ratings. The statistical techniques used most frequently by researchers in medicine were correlational analysis, regression analysis, discriminant analysis, and occasionally, canonical redundancy analysis.

Nursing. For undergraduate nursing students, several researchers have demonstrated relationships between preselection characteristics and academic achievement in the nursing program. SAT scores have been correlated with first year nursing GPA for 1,510 associate degree, diploma, and baccalaureate students (Mundy & Hoyt, 1965); with nursing GPA and fourth quarter clinical course grades for 79 diploma students (Plapp, Psathas, & Caputo, 1965); and with nursing GPA for 112 associate degree (Backman & Steindler, 1971) and 219 baccalaureate degree (Tillinghast & Norris, 1968) students.

Other researchers have found relationships between nursing GPA and high school GPA or high school rank. For 100 diploma students, Michael, Haney, and Jones (1966) claim that high school GPA was the best predictor of academic success in nursing school; Litherland (1966) confirmed this claim in his study of 3,358 diploma and baccalaureate students. Further, Tillinghast and Norris (1968) found correlation coefficients in the range of r=.30 to .60, indicating significant relationships of moderate magnitude between nursing GPA and high school GPA for 219 baccalaureate students.

Research in the general educational literature suggests that high school rank is the best single predictor of college performance (Lavin, 1965, p. 52). In nursing, Backman and Steindler (1971) found significant but low correlations (r=.28) between high school rank and nursing GPA for 112 associate degree nursing students.

The prediction efficiency of academic achievement for undergraduate nursing students has been increased by consideration of other variables. Zagar, Arbit, and Wengel (1982) were interested in predicting attrition and cumulative GPA for 570 students in a diploma nursing program. Their predictor variables were composite scores on the ACT test, three scales of the Minnesota Multiphasic Personality Inventory (MMPI), and 10 subscales of the Edwards Personal Preference Schedule (EPPS). In zero-order correlations, ACT scores were the only significant predictor of GPA (r=.35). When the 14 predictor variables were regressed on GPA, 38% of the variance was explained, with ACT scores significant. To predict attrition (graduation versus nongraduation), discriminant analysis was performed, with results parallel to those of the multiple regression.

Aldag and Rose (1983) were also interested in attrition, nursing GPA, and ACT scores. In their study of 787 associate degree nursing students, they examined the relationship between students' age on admission to the nursing program and these criterion variables. They were particularly concerned with students' age because they felt that traditional preselection criteria, including ACT scores, underestimated the performance of older students. No significant differences were found between student age at entry and cumulative nursing GPA, nor between GPA and ACT scores. There was a significant negative relationship between age and ACT scores, which confirmed their belief that there is an age bias in this measure. Age was also related to attrition rates, with a higher percentage of those in the 30-39 age group graduating as compared to students under 30 and over 40; there were significant differences between these three age groups.

Conversely, Reed, Feldhusen, and Van Mondfrans (1973) found that

student age (in months) at entry in an associate degree nursing program and previous education were positively related to first semester nursing GPA (N=665).

Most of the research on academic achievement in masters programs in nursing has focused on prediction of the masters GPA. Ainslie, Colby, Hoffman, Meserve, O'Conner, and Quimet (1976) found low correlations between baccalaureate GPA and masters GPA for 193 students in one program. Previous research cited by Sime, Corcoran, and Libera (1983) found correlation coefficients in the r=.05 to .37 range for baccalaureate and masters GPAs; their research confirmed this relationship (r=.32) for 138 students at one school.

In a multiple regression equation regressing baccalaureate GPA and Graduate Record Examination verbal and quantitative scores (GRE-V; GRE-Q) on masters GPA, Thomas (1974) explained 15% of the variance in distribution of grade point averages at the masters level. Other studies that have examined the relationship between GRE scores and masters GPA have demonstrated zero-order correlations of r=.23 to .41 for GRE quantitative scores (Ainslie et al., 1976; Stein, 1978), and conflicting findings regarding GRE verbal scores. While Stein (1978) found no significant relationships between GRE verbal scores and masters GPA, Ainslie and others (1976) describe significant correlations of moderate magnitude (r=.37).

Sime, Corcoran, and Libera (1983) were interested in the predictive validity of certain measures for success in one masters degree program (N=138). Their prediction variables included undergraduate GPA, an aptitude test, a measure of creativity, and a measure of flexibility in thinking. Criterion variables were cumulative masters GPA and faculty

ratings of students' overall competence and five other personal attributes. For regression of the predictor variables (excluding the measure of flexibility in thinking) on masters GPA, 23% of the variance was explained (R=.48), and the verbal subscale of the aptitude measure made the most significant contribution (R=.41).

When these variables were regressed on faculty ratings of students' overall competence, 12% of the variance was explained (R=.34); again, the verbal subscale of the measure of aptitude made the most significant contribution (R=.24). They concluded that the use of multiple predictors to determine success in masters programs did not appear warranted, and that they were unable to identify valid predictors of noncognitive attributes as measured by faculty ratings (Sime et al., 1983).

Finally, Tripp and Duffey (1981) looked at the relationship between three preselection characteristics (undergraduate GPA, GRE-V, GRE-Q) and the status of students in one masters program. Their criterion variables were graduation from the program (N=102), and two categories of nongraduation (65 applicants not admitted and 103 students who dropped out before graduation). Through discriminant analysis with three predictor variables and three groups, they determined that there was an underlying one-dimensional space that explained the data (extracting 98% of the variance in the discriminant space). Baccalaureate GPA made the greatest contribution to that discriminant function, extracting 78% of the variance. They caution, however, that the smaller contributions of GRE-V and GRE-Q (33% and 26% of variance, respectively) must be interpreted in view of the redundancy between these measures (r=.47; 22% of the variance in one GRE score explained by variance in the other).

Tripp and Duffey were able to correctly identify 72% of those

students who were not accepted for admission, but were not able to identify either those who ultimately graduated or dropped out of the program. They concluded that, for their program, use of traditional predictors such as GRE scores and undergraduate GPA was not helpful in determining those students who (once admitted) would succeed in the masters program.

Since 1965, a number of studies specifically related to nurse practitioners have been conducted. Most of that research involves description of demographic characteristics, discussion of attitudes toward and acceptance of nurse practitioners, their impact on the quality and type of care provided, and assessment of certain psychosocial factors (Dunn & Chard, 1980).

Previous research by this investigator examined relationships between preselection variables and academic achievement in one university setting. The sample included 196 graduates of three certificate and masters nurse practitioner programs (pediatric, family, and obstetric-gynecologic) between 1974 and 1980 (Dunn, 1981). The predictor variables were undergraduate GPA, type of nursing education, years of experience as a registered nurse, years out of school at entry in the program, age at program entry, and scores on state board licensure examinations (SBE scores). The criterion variable was the average final grade in nurse practitioner courses (numerical average of all courses).

When the predictor variables were regressed on average final grade, this six variable model explained only 18% of the variance ($\underline{R}$=.42). Age (negatively related) and undergraduate GPA (positively related) made the largest contributions ($\underline{R}$=.30 and $\underline{R}$=.39 respectively), and were statistically significant at the $\underline{p}$=.01 level (Dunn, 1981).

In relation to in-school clinical performance in nursing, very little research was identified. Generally, low clinical performance evaluations are related to lack of academic success in nursing school, regardless of other factors (Hutcheson, Garland, & Prather, 1973). Kissinger and Munjas (1982) reported on the relationship of certain predictor variables and success in using the nursing process on written clinical simulations, for 201 baccalaureate nursing students in six programs. The study considered a number of predictor variables--personality variables, intellectual skills, demographic data, cognitive style; however, only four variables emerged as significant predictors. They were: vocabulary knowledge (measured by the Extended Range Vocabulary Test); verbal ability (measured by SAT verbal scores); convergent thinking ability (measured by the vocabulary test and the Inference Test); and, field independent perceptual style (measured by the Witkin's Group Embedded Figures Test). The authors conclude that, since successful use of the nursing process is essential for clinical practice and completion of nursing programs, these measures might be used in admissions screening.

In summary, research on in-school academic and clinical performance of nursing students demonstrates some relationships between cognitive and noncognitive attributes and measures of performance. Like the research in medicine, however, these relationships are neither of high magnitude or consistently significant. Researchers in this area tend to use correlational analysis, regression analysis, and discriminant analysis as their primary statistical techniques.

Postgraduate Performance

Medicine. Most of the research on postgraduate performance in

medicine has focused on clinical performance of physicians in residency programs. Veloski and Gonnella's (1980) study of the relationship between residents' performance on Part III of the NBME examination and type of residency program is an exception. They looked at 1,028 graduates of one medical school a year after graduation. At that time graduates had completed 1 year of residency in eight specialty programs and had taken Part III of the NBME examination. Although they adjusted scores on Part III based on prior differences in scoring on Part II of the NBME examination, they found significant differences in Part III scores related to type of residency.

Those graduates who entered more specialized residency programs (psychiatry and pathology) scored significantly lower than those who entered general practice programs (family, internal medicine, or flexible) or general specialty programs (surgery, obstetrics-gynecology, or pediatrics). They concluded that, because the NBME examination assesses knowledge essential for all practicing physicians, early specialization after graduation should be discouraged (Veloski & Gonnella, 1980, pp. 142-147).

The research on clinical performance of residents usually uses subjective ratings of supervisory personnel as the criterion measure. Paiva (1979) looked at the reliability of self and supervisor ratings for residents and concluded that there was substantial agreement on measures of 18 attributes (pp. 118-123). Keck, Arnold, Willoughby, and Calkins found that clinical performance of residents was related to supervisor ratings, but was not related to academic performance in medical school (1979). An earlier study by Margolis and Cook (1974) demonstrated that the clinical performance of pediatric residents was

not related to intern-matching rank or scores on a pediatric certification pretest.

When Rose, Corman, and Roberts examined the relationship between clinical performance of residents as measured by chart audit and knowledge as measured by oral examination, they found only 54% agreement and no significant correlation (1979, pp. 113-117). Likewise, end-of-year clinical performance of pediatric interns (judged by chief residents) was not related to their preadmission rank, sex, or marital status (Werner, Adler, Robinson, & Korsch, 1979).

Donnelly (1979) was interested in the clinical performance of interns, as measured by supervisory ratings, and their stage of ego development, as measured by Loevinger's eight stages of ego development. She found that interns received higher performance ratings when they were judged to be at or above stage five of ego development (conscientious; individualistic; autonomous; integrated). Those below this state (impulsive; self-protective; conformist; self-aware) received lower performance ratings (pp. 99,101).

In summary, research on the postgraduate performance of physicians has found few relationships between ratings of clinical performance and cognitive or noncognitive predictor variables. One study did indicate a relationship between scores on a cognitive examination and the type of postgraduate training of physicians; and, another found relationships between clinical performance and stage of ego development of interns. The studies reviewed in this section were limited to correlational analysis for their statistical technique.

Nursing. The majority of research on postgraduate achievement in nursing has focused on performance on state board licensure examinations

(SBE), probably because it is the only standard performance measure available for all registered nurses (RNs). Until recently the SBEs were divided into 5 subscales: medical nursing, surgical nursing, psychiatric nursing, pediatric nursing, and obstetric nursing; therefore, 5 scores were reported. (The SBE now reports 1 score.)

Wolfle and Bryant (1979) reviewed the literature on relationships that have been demonstrated between SBE scores and other cognitive variables. Ten (10) studies found low positive correlations with SAT quantitative scores (range: r=.07 to .26; mean r=.18); and, 15 studies found moderate or high magnitude correlations with SAT verbal scores (range: r=.34 to .77; mean r=.42). Another 10 studies examined relationships with cumulative nursing GPA and found moderate magnitude positive correlations (range: r=.30 to .54; mean r=.45). The largest number of studies (32) demonstrate low to high magnitude positive correlations with National League for Nursing (NLN) test scores (range: r=.21 to .77; mean r=.49). (The NLN tests are standardized national in-school achievement examinations that parallel the content areas of the SBE.)

Two reports found significant relationships between SBE scores and age. Aldag and Rose (1983) looked at 787 associate degree graduates and found age significantly related to each subscale. Those examinees in the 30-39 year age group demonstrated higher average scores on each subscale and higher pass rates (97.3%) as compared to those under 30 (89.3% pass) or over 39 (94.9% pass).

Reed and Feldhusen (1972) studied 155 graduates of five associate degree nursing programs. They regressed 18 pre-nursing school predictors on each subscale of the SBE, producing multiple Rs in the range

of .49 to .69. Only examinee age (in months) and SAT verbal scores were significant for each subscale of the examination. However, when they added nursing GPA to the regression equation, age dropped out as a significant predictor. By adding GPAs and their squares for each of four semesters in the program, they were able to improve the prediction efficiency of the equation (R= .65 to .73; R^2=.42 to .53). In these equations, SAT verbal scores and the square of second semester nursing GPA were the only variables significant for each subscale of the SBE.

Tillinghast and Norris (1968) had earlier reported relationships between SBE scores and SAT scores, in the range of r=.20 to .78. In addition, Litherland (1966) found relationships with high school GPA when he studied 3,358 diploma and baccalaureate graduates; and Mueller and Lyman (1973) found relationships with NLN achievement tests for 110 diploma graduates.

Two studies deserve particular attention. Bell and Sanchez (1980) claim that the literature suggests that verbal ability (SAT verbal) and NLN test scores are the best predictors of SBE scores. Their research involved three samples from three baccalaureate nursing programs. One sample's SBE scores were used as the criterion measure (N=128); the scores of the other two samples (N=312; N=101) were used to predict those of the criterion group. Regression analysis revealed that scores for the five subscales were relatively stable across time and universities (R=.64 to .83). A subsequent factor analysis showed that one factor explained 65-67% of the variance in SBE scores, NLN scores, and SAT verbal scores for two samples. They concluded that this was evidence that verbal ability is the construct underlying these cognitive measures.

Wolfle and Bryant's research (1978) is remarkable because it is the only study identifed that used path analysis, in this case attempting to demonstrate causal relationships between SBE scores and other cognitive measures. While between 33% and 51% of the variance in SBE subscale scores was explained when nursing GPA and NLN scores were regressed on SBE scores, some of those effects were indirect or spurious. Path analysis determined that the association between GPA and SBE scores was 40% direct effect, 40% indirect effect, and 20% spurious effect (these percentages are averages for the five subscales). On the association between SBE scores and NLN scores, the subscales showed an average 70% direct effect and 30% spurious effect of prior ability. The NLN tests were not particularly good predictors of scores on the pediatric and psychiatric subscales, where 40% and 47%, respectively, of the effects were spurious.

Researchers who have examined relationships between the clinical performance of nurses and other factors have usually used self or supervisor ratings as their criterion measure. As previously noted, these comparisons are problematic because they are probably measures of a different domain of behavior; therefore, lack of correlation should not be surprising. Both Bohan (1966) and Saffer and Saffer (1972) looked at relationships between self and employer ratings of staff nurses and their performance in nursing school; they found no significant correlations with grades. Brandt and Metheny (1968) found that self and employer ratings for staff nurses were related to grades in clinical courses, but not to SBE scores.

McCloskey (1983) was particularly interested in the relationship between type of nursing education and job effectiveness. Staff nurses

(N=299) from 12 Chicago hospitals were chosen as the sample: 53 licensed practical nurses, 49 RNs with BSNs, 63 RNs with ADNs, and 134 RNs with diplomas. The criterion measure was head nurse ratings of job effectiveness. The predictor variables were 36 indicators of formal education, continuing education, and job skills.

The job skills ratings were based on Schwirian's 6-D Scale of nursing behaviors, and were completed by head nurses as well as the staff nurses themselves. When the 36 predictors were regressed on job effectiveness, 57% of the variance was explained by two head nurse skills ratings. When those ratings were removed from the equation, the remaining 9 variables explained only 19% of the variance in job effectiveness (years of nursing education and clinical experience in the educational program each accounted for 1% of the variance).

When self-ratings of job skills were removed from the equation, 13% of the variance in job effectiveness was explained; again, years of nursing education and in-school clinical experience each accounted for 1% of the variance. To determine whether nursing education had an effect on the job skills ratings, the other variables from the original equations were regressed on skills. Of the five variables that entered the equation, nursing degree accounted for 4% of the variance and clinical experience in the educational program accounted for another 5% of the variance.

McCloskey's regression analyses demonstrated that years of nursing education has a significant but small effect on head nurse ratings of job effectiveness. Variables that made no significant contribution to the explanation of variance included nursing GPA, SBE scores, and job experience.

In summary, a review of research conducted on the postgraduate performance of registered nurses revealed some relationships between SBE scores and other cognitive variables. Few significant relationships were found between the clinical performance of nurses and cognitive or noncognitive predictor variables. The research reviewed used correlational analysis, regression analysis, factor analysis, and, in one case, path analysis as the statistical techniques.

Nurse practitioners. For nurse practitioners, there are few studies that relate postgraduate performance to specific predictor variables. One study conducted by Farrand, Holzemer, and Schleutermann (1982) examined the performance of 46 masters and certificate prepared nurse practitioners, and 31 RNs (MSN, BSN, ADN, and diploma) on three patient management problems. There were no significant differences in scores within the nurse practitioner group (between masters and certificate), although there were differences between the nurse practitioner group and the RN group. There was an interesting result within the RN group: BSN prepared RNs scored higher than the MSN, ADN, or diplomas nurses. When the researchers controlled for length of experience, the scores for the RN group improved but remained lower than nurse practitioners' scores.

Another study looked at the relationship between nurse practitioners' performance on a 240 item test of medical knowledge and the quality of their clinical practice (Hastings, Sasmor, & Murray, 1975). Their Performance on the test and in practice was compared with that of physicians in internal medicine practice. The quality of care measures were a retrospective chart audit and random physician recheck of patients and charts (intercorrelation, r=.72). Although the nurse

practitioners scored lower on the test of didactic knowledge, the clinical measures indicated that the care they provided was equal to or better than that of the physician group (pp. 272-275).

Hoekelman and others (1975) reported that the job performance of graduates of one university nurse practitioner program was not related to academic achievement in the program or to type of nursing education. Carfang (1979), however, found that self-ratings of job effectiveness for 29 pediatric nurse practitioners was related to years of experience as an RN; perceived effectiveness was not related to years of experience as a nurse practitioner or to age.

Data from the 1980 survey for the Longitudinal Study of Nurse Practitioners indicated that the employment status of nurse practitioners was related to their level of nurse practitioner preparation, prior nursing education, the length of their nurse practitioner program, and the size of classes in their program.

The 1980 Sultz data for 1,579 nurse practitioners showed that graduates of certificate nurse practitioner programs were more likely to be employed in a nurse practitioner position (direct care, teaching, or consulting) than were graduates of masters programs (75.6% versus 60.4%; pediatric: 76.7% versus 57.3%). Although these percentages were not substantially different than they were for the 1973 survey, there was a tendency for more certificate graduates to be employed in non-nurse practitioner nursing positions (1973: 9.8%; 1980: 13.1%) and for fewer masters graduates to be employed in these positions (1973: 34.8%; 1980: 24.9%).

In regard to the prior nursing education of certificate graduates, those with diplomas were most likely to be employed in nurse practition-

er positions (80.3%) and those with masters degrees were least likely to be employed in these positions (68.4%). This was also true for pediatric certificate graduates, and was not different than the 1973 findings.

Graduates of all programs were more likely to be employed as a nurse practitioner if the length of their program was between 6 and 17 months (78%; 77% pediatric). Those who attended very short programs (3-5 months) or long programs (17+ months) had rates of 37.5% and 60.2%, respectively. While there were no apparent differences in employment status for the graduates as a whole, or for certificate graduates, in relation to class size, there were differences for those who attended masters programs. Masters graduates who attended programs with larger classes (8+) were more likely to be employed as nurse practitioners than those who attended programs with smaller classes (less than 8 students/class) (61.9% versus 55.7%). This was also true for the pediatric masters graduates (57.8% versus 40%) (Sultz, Bullough et al., 1983).

Finally, previous research by this investigator looked at the relationship between the self-rated clinical performance of 196 graduates of one university nurse practitioner program and several predictor variables. Seven predictor variables were used: age in years, final average grade in nurse practitioner courses, undergraduate GPA, years out of school at time of entry in the nurse practitioner program, prior nursing education, years of experience as an RN, and SBE scores. When these predictors were regressed on performance ratings, 20% of the variance was explained ($\underline{R}$=.45). Only final grade in the program and undergraduate GPA were significant predictors (Dunn, 1981).

An additional finding in that research (Dunn, 1981) was that the employment status of graduates was not related to level of education,

marital status, number of dependents, or geographic location. Seventy-nine percent (79%) of the graduates were employed in nurse practitioner positions at the time of the survey.

In summary, most of the literature on postgraduate performance of nurse practitioners is descriptive and adds little information about relevant predictor variables.

Certification

Medicine. Other literature reviewed concerned specialty certification in medicine. In relation to certification in general, a longitudinal study of medical school graduates of 1960, conducted by the Association of American Medical Colleges, found that board certified physicians had significantly higher incomes than their noncertified colleagues (Erdmann, Jones, & Tonesk, 1979).

Early studies by Trussell (1962) and Morehead, Trussell, and Ehrlich (1964) examined the relationship between the certification status of physicians and the quality of care that they provided. These studies found no significant relationships (cited by Downing & Maatsch, 1979, p. 124). Later studies demonstrated no differences between general practitioners, noncertified physicians, and certified physicians on patient management problems (McGuire & Williamson, 1968), and low magnitude correlations between scores on multiple choice questions and a measure of the physicians' diagnostic accuracy and patient management (Gonnella, 1973). Yet another study (Payne & Lyons, 1972) found that board certification, type of specialty, and years of practice did not correlate well with process audit ratings of clinical performance (cited by Downing & Maatsch, 1979, p. 125).

In an attempt to establish the concurrent validity of certification

examinations for Canadian family practice physicians, Pawluk, Roberts, Tech, and Neufeld (1976) compared various examination formats with an independent indicator of quality of care. They found that the composite examination score was not a useful predictor of quality of care (r=.12, not significant), nor was the multiple choice format score (the relationship was close to zero when interaction between the various subscales was controlled). They did, however, find a significant relationship between management of selected indicator conditions and patient management problems on the examinations, when years of practice was controlled (r=.60) (pp. 301-302). The researchers could not establish concurrent validity based on these results.

On the other hand, Downing and Maatsch (1979) contend that multiple choice questions can be written that are powerful discriminators of varying levels of clinical competence. They tested three groups with known differences in levels of competence: medical students (N=22), emergency medicine residents (N=36), and graduates of emergency medicine residency programs with at least one year of practice in an emergency room (N=22). This research was conducted for the American College of Emergency Physicians for subsequent use by the American Board of Emergency Medicine in development of the first certification examination in emergency medicine. The researchers developed two multiple choice subscales; one subscale contained items of medium difficulty and the second subscale contained items of high clinical relevance (high magnitude correlations with an independent rating of simulated clinical performance) (pp. 124,126).

As expected, both multiple choice subscales discriminated examinees according to their known levels of competence; however, all examinees

scored higher on the high clinical relevance subscale. Stepwise discriminant analysis produced two discriminant functions: the high clinical relevance function and the medium difficulty subscale function. Using these two functions, a classification analysis correctly identified 81.3% of the examinees. Taken separately, the high clinical relevance scale classified 76.3% correctly and the medium difficulty scale classified 71.2% correctly (p. 127).

Further examination revealed that the high clinical relevance scale was statistically higher in internal-consistency reliability than the other scale. Downing and Maatsch concluded that test items chosen for their high clinical relevance are more reliable, more efficient in discriminating levels of physician clinical competence, and less difficult for clinicians than items of medium difficulty that are typically used for certification or licensure (1979, pp. 127-128).

In later research on Part II of the emergency medicine certification examination, Maatsch (1981) was interested in the discriminant and concurrent validity and the reliability of examination ratings for multiple choice questions, patient management problems, and simulated clinical encounters (examiner administered oral evaluations). Again, the examination subscales were administered to fourth year medical students (N=22), residents in emergency medicine (N=36), and graduates of emergency residency programs with at least one year in practice (N=22). A fourth group, physicians with at least five years of experience in emergency medicine who were not graduates of residency programs, was added for this study (N=14).

The composite examination score accurately discriminated between graduates, residents, and students, but the results for physicians in

the fourth group were not statistically different than the resident or graduate group. In regard to concurrent validity, Maatsch found that patient management problems did not correlate as well with examiner ratings for the simulated clinical encounters as did the multiple choice questions. Although other studies had reported correlations between written and oral examinations between r=.27 to .54, the correlation between objective scores and simulated clinical encounters was r=.83 for this study. Maatsch cautions, however, that this high correlation is probably due to inclusion of students (with known less competence), thereby increasing the range of ability of those tested.

Finally, two other studies examined raltionships between certification examination performance and other noncognitive variables. Mellsop (1981) reports results of 10 years of experience with the Australian certification examination in psychiatry (N=531). This examination is taken after 3 years of postgraduate training in psychiatry, and has an average pass rate of 68%. Mellsop found significant relationships between examination performance and sex, age, prior postgraduate experience, and length of postgraduate psychiatric training. Women were less likely to pass the examination on their first attempt than men (58% versus 71%), and those under age 30 were more likely to pass than those over age 30 (85% versus 65%).

In addition, those examinees who had 1 or 2 years of general medicine residency before entering psychiatric training were more likely to pass than those with 3 or more years in a general medicine residency before psychiatric training (72.5% versus 61%). He also found an inverse relationship between length of postgraduate psychiatric training and successful performance on the examination (87%-3 years; 71%-4 years;

over 4 years). His conclusion was that the paradoxical finding regarding length of specialty training may be an indirect relationship related to the age of examinees (Mellsop, 1981).

Martin, Gullickson, and Gerken (1980) were particularly interested in the relationship between performance on the certification examination for physical medicine and rehabilitation and the size of the residency programs from which examinees graduated. Programs were classified in one of three categories: less than five residents, five to 10 residents, and more than 10 residents. When these three classifications were compared with scores on Part I and Part II of the examination, no significant relationships and low magnitude correlations (r=.10 and r=.18, respectively) were found.

To summarize, the literature related to certification in medicine suggests that certain noncognitive variables such as age, sex, prior experience, and length of specialty training may be related to performance. In relation to examination validity, while some researchers have established discriminant validity for known groups, no research was identified in which the concurrent validity of certification examinations (with actual clinical practice) or predictive validity was established. Statistical techniques used in these studies included correlational and discriminant analysis.

While a limited amount of research on certification in medicine was found, it should not be concluded that little research has been done. Most of the statistical analyses done by testing agencies is for the internal use of specialty certification boards and, therefore, is not published in the medical literature. .

Nursing. No published studies have been identified that use certi-

fication in nursing as a performance measure, or that relate credentialing to the outcomes of care (Lang, 1979, p. 336). In a survey of certified nurses conducted by Edari and staff for the Study of Credentialing in Nursing, there was no indication that certification had either a positive or negative influence on the perceived quality of patient care. That survey's distribution of certified nurses according to level of education, when compared to national statistics, showed gross underrepresentation of diploma nurses and gross over-representation of nurses with masters degrees in nursing (1979, pp. 331-332).

Because no published research was found, the investigator queried three agencies that certify nurses in expanded roles: the American Association of Nurse Anesthetists (AANA), the American College of Nurse Midwives (ACNM), and the NAACOG Certification Corporation. As a result of those queries, descriptive information and the results of research submitted for publication were obtained.

Certification for nurse anesthetists is conducted by the AANA Council on Certification; this is a criterion-referenced examination. In 1979, Fleming published description information in the association's journal about the results of six examinations between 1975 and 1978 (the exam is administered twice a year: June and December). Over those six examinations, the reliability, average difficulty, and average discrimination of test items has remained relatively stable (KR20= .92 to .93; mean p=.69 to .74; mean r=.30 to .34). The average scores have increased during this period, which Fleming notes coincides with revision of accreditation standards for educational programs. This revision resulted in fewer numbers of accredited programs, but no significant decrease in numbers of graduates.

Before 1978, examinees who failed the examination were limited to three attempts to pass; that restriction on repeaters was eliminated after December 1977. When repeaters' scores were averaged with those of first-time takers, they lowered the overall average and increased the standard deviation. First-time examinees had average scores that were 28 points higher than those of repeaters, and there was less likelihood that repeaters would pass the examination as the number of retakes increased (Fleming, 1979).

Examinees with baccalaureate degrees in nursing had the highest pass rate and highest average scores. There was no appreciable difference in performance based on the type of previous clinical experience. In regard to the educational programs from which examinees graduated, there was no real difference in scoring between graduates of programs greater or less than 24 months in length (the revised accreditation standards required a minimum program length of 24 months). There were differences, however, in terms of the type of educational program: graduates of MSN, BSN, and military nurse anesthesia programs performed better than graduates of programs that had been discontinued (Fleming, 1979).

Further descriptive information for examinations between 1979 and 1982 (seven exams) was obtained from the AANA. Those data show that repeaters continue to score lower than first-time takers (average 30 points), and that their scores decrease with successive repeats. Examinees with BSN preparation continue to have higher average scores than those with diplomas, associate degrees, or masters degrees in nursing. In terms of the type of nurse anesthesia program, graduates of military, MSN, and BSN programs had higher average scores than graduates

of discontinued programs (listed in order of decreasing scores) (AANA, 1983).

Beginning with the December 1980 examination, recertification candidates were included in the data. Those data for four examinations indicate that the 21 recertification candidates scored an average 27 points lower than candidates who were taking the examination for the first time (AANA, 1983).

The second certifying agency, the NAACOG Certification Corporation, has administered an examination for obstetric-gynecologic nurse practitioners since 1980. Data for the 1980 and 1981 examinations were obtained from the NAACOG Certification Corporation's newsletters. This is a norm-referenced examination; information on the examinations' reliability, average difficulty, and average discrimination on the test items is discussed in Chapter V. Informally trained practitioners were allowed to sit for the examination between 1980 and 1982. In the first 2 years in which the examination was administered, informally prepared candidates had failure rates that were significantly higher than those for graduates of formal nurse practitioner programs (1980: 23% versus 6%; 1981: 10% versus 4%) (NAACOG Certification Corporation, 1980, 1982).

The report on the first administration of the NAACOG Certification Corporation's examination indicated that there were significant differences in average scores according to level of education, length of experience in nursing, and area of clinical practice. Those candidates with BSN educational background scored significantly higher than those with diploma or associate degree education. Scores for candidates with more than 20 years of experience in nursing were significantly lower than those with less than 20 years of experience (this is an indirect

measure of the effect of age: those over 40 years of age versus younger candidates). In regard to area of clinical practice, there were significant differences in performance between candidates who practiced in obstetric-gynecologic settings compared to candidates whose practice was limited to gynecology or to obstetrics (highest scores for combined practice, lowest for obstetric practice) (NAACOG Certification Corporation, 1980).

A report of research submitted for publication was obtained from a third agency, the American College of Nurse Midwives. That research was conducted by the test consultant for the ACNM Division of Examiners (Fullerton) and the chairperson of that division (Thompson). They were particularly interested in assessment of the continuing competence of nurse midwives, and the use of entry-level certification examinations as the criterion measure (Fullerton & Thompson, 1983).

The ACNM examination is norm-referenced, with a modified essay format. Sixty-two (62) volunteers participated in the study, which was conducted during the fall of 1981 and winter of 1982. The sample was divided into two major groups: first-time certification candidates and recertification candidates. The recertification group was stratified by date of initial certification (retroactive certification before 1971; certification between 1971-1975; certification between 1976-1979) and primary employment focus (clinical practitioners; nurse midwifery faculty; nurse midwifery administrators/other).

Analysis of variance for the recertification group revealed no significant differences in average scores (raw scores) that were attributable to age, number of years since graduation from the nurse midwifery program, type of nurse midwifery program (certificate, mas-

ters), or level of education (diploma, BSN, MSN). There were significant differences in examination performance, however, between the certification and recertification groups.

The failure rate for first-time takers (certification group) was lower than that for recertification candidates; and, the average score for certification candidates was 358.4 versus 330.5 for recertification candidates (faculty mean=349.3; clinician mean=325.3; administrator/other mean=317.5). Analysis of variance demonstrated significant differences between the four subgroups on five of the six subscales of the examination. (The six subscales are: antepartum, intrapartum, postpartum, newborn, family planning/gynecology, and professional issues. There were no differences between groups on the postpartum subscale.)

Post-hoc analysis of the five subscales on which significant differences between groups were found revealed that, for three of the subscales (antepartum, intrapartum, family planning/gynecology), faculty recertification candidates and certification candidates scored alike (higher average scores) and clinicians and administrators/others in the recertification group scored alike. On the remaining subscales, the first-time takers scored higher than the three recertification groups on the newborn subscale; and, the clinician recertification group scored lower than the other three groups on the professional issues subscale.

Fullerton and Thompson (1983) conclude that the recertification candidates were a norm group different than the first-time takers. The test scores of faculty recertification candidates tended to increase the performance of the recertification group toward the norm; and, clinical preceptors, whether viewed as faculty or clinicians, tended to raise the group's level of performance. Because of the results of this research,

they suggested that the competency of nurse midwives over time cannot be assured by reevaluation with entry-level certification examinations.

As mentioned in the summary on certification in medicine, there is probably other information regarding certification in nursing that the investigator was not able to obtain from public sources. It is also true, however, that a more limited amount of research on certification has been conducted, due to the fact that it is a relatively recent phenomenon in nursing.

Based on available data, there is no indication that certification examinations in nursing have established concurrent, predictive, or construct validity. It could be claimed, however, that a certain degree of differential validity can be inferred. Differential validity refers to the "degree to which different demographic groups perform equally well on a test" (Report of the NCHCA, 1981, p. 19). In this regard, differential validity could be inferred according to characteristics such as level of education, type of education, employment setting, length of experience, and age of certification examinees in nursing.

Finally, if it is assumed that cognitive examinations, such as those for credentialing in nursing, are actually measures of academic achievement, discriminant validity can be inferred in some cases. Campbell and Fiske's concept of discriminant validity, an extension of construct validity, implies that "traits that are truly distinct from one another should lead to different results even if they are measured by the same method" (Carmines & Zeller, 1979, p. 54). It could be expected, therefore, that certification candidates with different levels of competence would differ in terms of performance.

Summary of Research Results

A review of the literature reveals some relationships between certain cognitive and noncognitive variables and measures of academic achievement, although the evidence is sometimes conflicting and correlations are generally of low magnitude. In regard to clinical performance, either in-school or postgraduate, there is little evidence that strong predictors have been identified. Part of the problem is undoubtedly due to the fact that measures used to assess academic achievement and those used to assess clinical performance are not measuring the same construct or domain of behavior. The decreased efficiency of certain predictor variables over time is probably another factor.

The limited amount of research on certification in nursing and medicine suggests that relationships between performance on examinations and in actual clinical practice are, at best, equivocal. While it could be argued that some examinations are valid discriminators of certain examinee characteristics, other types of validity cannot be inferred.

Researchers in this area most often used statistical techniques such as correlational analysis, regression analysis, and discriminant analysis. A significant limitation of much of the research in this area is the lack of control for intervening or explanatory variables.

Chapter IV describes the conceptual model for this research, and discusses modeling, frameworks for modeling, the research model, and assumptions of the study.

Chapter IV. Conceptual Model

One of the objectives of this research was to develop a model to explore relationships between sociodemographic characteristics of certification examinees, characteristics of their educational programs, and their examination performance. That conceptual model was then tested to determine its ability to predict examination performance. This chapter describes the purpose and process of modeling, frameworks for modeling, the research model, and assumptions underlying this theoretical approach.

Modeling

Modeling was used in this research primarily as a heuristic device. That approach to theorizing is helpful in developing hypotheses and in exploring relationships among variables (Asher, 1976, p. 8). It was not the investigator's intent to propose a functionally specific model, that is, a model expressing exact causal relationships between variables.

As Achen notes, "any attempt at specifying exact causal functions must necessarily result in oversimplified explanations" (1983, p. 15). In addition, because of the incomplete data base and the lack of specificity in the model, causal analysis techniques would be difficult to use or interpret. As with other research in this area, this study was essentially descriptive and attempted to discover and test theory related to academic achievement.

Frameworks for Modeling

The conceptual model for this research was constructed from other literature describing educational and social processes. Conceptual frameworks based on educational models have been discussed by Carroll, Cooley and Lohnes, DeCecco and Crawford, Reed and Riley, and Stufflebeam.

Educational models. DeCecco and Crawford's basic teaching model includes four components that they believe are essential to the teaching process. Those components are: instructional objectives, entering behaviors, instructional procedures, and performance assessment; there are feedback loops between and among these components (cited by Frisbie, 1979, p. 2). In relation to the learning process, Carroll (1963) proposed a conceptual model of factors affecting student success. In this model, students' degrees of learning are contingent on their time spent in learning activities and the time needed to master the information. Time spent is related to opportunities provided for learning and students' motivation to learn, while time needed is related to students' aptitudes, ability to understand instruction, and the quality of that instruction (cited by Cooley & Lohnes, 1976, pp. 187-188).

Stufflebeam and others have described a model for educational evaluation that includes four elements: context, input, process, and product (cited by Steele, 1978). Reed and Riley, in their comprehensive evaluation model for nursing education, have expanded on Stufflebeam's model. Their three-dimensional model explains the who, what, and why of educational evaluation. Stufflebeam's four elements are the "why" dimension of evaluation. The "what" (is evaluated) dimension also contains four elements: students, faculty and administration, curriculum, and resources. Finally, the "who" (does the evaluation) dimension

of this model incorporates students, faculty, administration, and others. In this systems model, students are considered "throughputs" of the system (Reed & Riley, 1979, pp. 442-443).

Cooley and Lohnes (1976) have specified a taxonomy of student output measures in terms of the type of outcome and the type of data; their taxonomy was adapted from work by Astin, Panos, and Creager (1967). The two types of data are psychological and behavioral, and the two types of outcomes are cognitive and affective. In this taxonomy, knowledge and academic achievement are cognitive outcomes based on psychological data, whereas level of academic achievement and vocational achievement are cognitive outcomes based on behavioral data (p. 146). (See Table 2.)

Table 2
Cooley & Lohnes' Taxonomy of Student Output Measures

Type of Data	Type of Outcome	
Psychological	Cognitive	Affective
	Knowledge	Self concept
	General intelligence	Interests
	Critical thinking	Values
	Basic skills	Beliefs
	Special aptitudes	Drive for achievement
	Academic achievement	Satisfaction with school
Behavioral	Cognitive	Affective
	Level educational attainment	Career choice
	Vocational achievement: level of responsibility; income; awards of special recognition	Avocations Mental health Citizenship Interpersonal relations

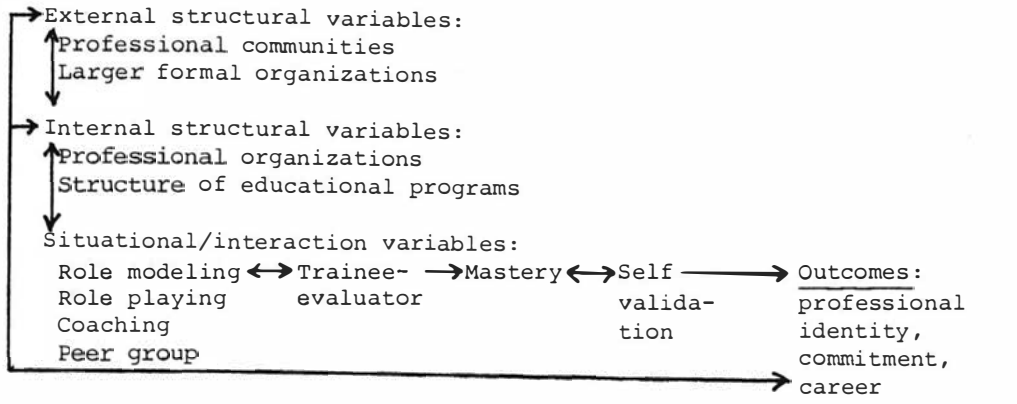
Note. Adapted from Evaluation Research in Education by W.W. Cooley & P.R. Lohnes, NY: Irvington Publishers, Inc., 1976, p. 146.

Cooley and Lohnes also developed an educational model that specifies functional relationships between various sets of variables and measures of educational outcomes. In this model, educational outcomes are dependent on family, instruction, peer groups, and initial abilities and motives sets. The family set is dependent on community culture and initial abilities and motives; the instruction set is dependent on school resources and policies, peer groups, and initial abilities and motives; and, the peer groups set is dependent on communtiy culture, school resources and policies, and initial abilities and motives. Finally, the school resources and policies set is dependent on the community culture set (1976, p. 153).

In contrast to these educational models other researchers have looked at evaluative outcomes in terms of social process models.

Social process models. Based on their research on physicians, Bucher and Stelling (1977) developed a model for the professional socialization process (see Table 3). They concluded that the outcomes of socialization were in large part determined by the structure of professional education programs, and there was nothing to support the idea that the socialization process established effective mechanisms for individual internal control or colleague control. When judging their own performance and competence, physicians tended to emphasize self-evaluation and self-validation, and focused on the process rather than the outcomes of their work. In effect, they felt that they were accountable only to themselves. Because of this, Bucher and Stelling discussed the need for greater visibility, accountability, and external review of professional performance (pp. 257, 281, 283, 284).

Table 3
Bucher & Stelling's Model
for the Professional Socialization Process



Note. Adapted from Becoming Professional by R. Bucher & J.G. Stelling, Beverly Hills, CA: Sage Publications, Inc., 1977, p. 276.

Others have constructed research models specifically related to nursing. Wolfle and Bryant's research model (1978; described in Chapter III) was based on the dynamics of the social process of nursing education, while McCloskey's (1983) conceptual framework was an educational-social process interaction model. McCloskey looked at the relationships among individual characteristics of nurses, their educational preparation, employment setting, job skills (competence) and the impact of these variables on job effectiveness (performance). The design model tested in her research was: Job effectiveness is a function of formal education + continuing education + job skills + job responsibilities + academic aptitude (p. 54; see Chapter III for results).

While Bucher and Stelling discussed professional socialization in terms of symbolic interactionism, LaDuca (1980) used this theoretical approach to define the structure of competence in health professionals. For the evaluation of competence, others have talked of the need to

determine the "boundaries of competence" (McGaghie, 1980), and have described analogies between measures of cognitive knowledge and measures of the structure of the health care system itself (Greenlick, 1981). In addition to the structural approach to the assessment of quality of health care, Donabedian classifies two other approaches: evaluation of the process of care and evaluation of the outcomes of care (1976, p. 7).

The conceptual model for education of pediatric nurse practitioners, which was developed by the Association of Faculties of Pediatric Nurse Associate/Practitioner Programs, describes eight role characteristics that facilitate acquisition and application of critical educational program content within the practice setting. These characteristics are: accountability, advocacy, collaboration, competency, critical thinking, mutual decision making, responsiblity, and self-direction (1982, p. 8).

The aforementioned sources were the primary contributors for the conceptual model designed in this research. In addition, the model is based on empirical evidence cited in the literature review, and on the investigator's experience as a nurse practitioner and nurse practitioner educator.

Assumptions

The implicit assumptions of this research were: (a) the competence of any health professional is related to cognitive, affective, and psychomotor domains of behavior; (b) objective measures of competence/performance assess only the cognitive domain of behavior; (c) credentialing mechanisms, such as certification, are necessary to set standards of performance; (d) standard setting is an attempt to protect the public

from incompetent practitioners; and (e) objective performance standards do not necessarily assure that quality health care will be provided.

Furthermore: (f) objective performance standards serve to validate the level of academic/educational achievement of providers; (g) educational achievement is an outcome of formal (educational) and informal (experiential) learning, as well as other factors; and (h) level of educational achievement is the behavioral outcome of a cognitive psychological construct. Therefore, it should be possible to predict performance on a cognitive examination, if the relevant predictor variables are isolated.

Research Model

Empirical evidence from previous research suggests that the level of performance (academic achievement) of individuals is most consistently related to other cognitive measures, such as undergraduate grade point average or nursing state board examination scores. Unfortunately, those types of measures were not available for this research, and, therefore, were not included in the model (see Figure 1). Likewise, certain personality and demographic characteristics that have been shown to affect performance were not available for the sample used in this research.

Sociodemographic variables. Those sociodemographic variables that were available for the sample, and were included in the model, have all been related to cognitive performance in other studies. To elaborate, examination performance of candidates is directly related to their age, current work setting and job function, months of experience as a nurse practitioner and registered nurse, type of nurse practitioner preparation, highest level of education, and sex. There are also inter-

correlations between these independent variables, which may modify their effects on performance.

1. Age. Previous research has generally demonstrated that there are negative relationships between age and cognitive performance. That research includes studies by: AANA, 1983; Aldag and Rose, 1983; Conger and Fitz, 1963; Dawson-Saunders and Doolen, 1980; Dunn, 1981; Hopkins and Stanley, 1981; Johnson and Hutchins, 1966; Lavin, 1965; Mellsop, 1981; NAACOG Certification Corporation, 1980; Reed and Feldhusen, 1972; Reed, Feldhusen, and Van Mondfrans, 1973; and Tucker and McGaghie, 1982.

The investigator expected that the results of this research would be consistent with those previous findings; and, could be attributed to the effects of a timed test (Hopkins & Stanley, 1981), as well as other unknown factors related to the aging process. In addition, there are obvious intercorrelations between age (as a function of time) and months of experience as a registered nurse and nurse practitioner.

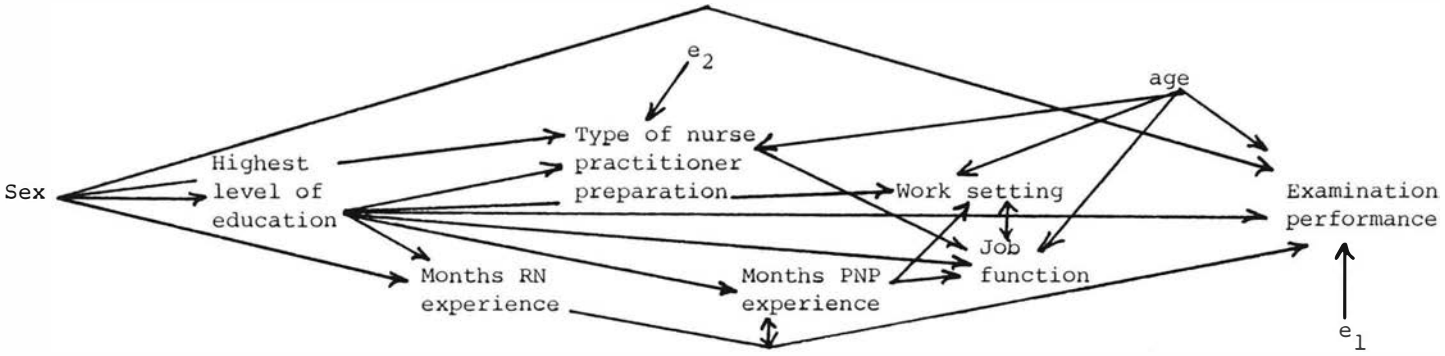
2. Current work setting and job function. It was expected that those candidates working as nurse practitioners in general pediatric primary care settings would have higher scores than candidates who were not functioning as nurse practitioners, or candidates who were employed in more specialized practice settings. These findings were expected because the examination is designed to assess knowledge in general pediatric primary care, and those candidates who were not practicing as nurse practitioners or were practicing in specialized settings necessarily narrowed their experiential knowledge base. This effect has been documented in other research (Fullerton & Thompson, 1983; NAACOG Certification Corporation, 1980).

There are also intercorrelations between job function and work

Figure 1
Research Model: Variables and Relationships Influencing the Examination Performance of Pediatric Nurse Practitioners

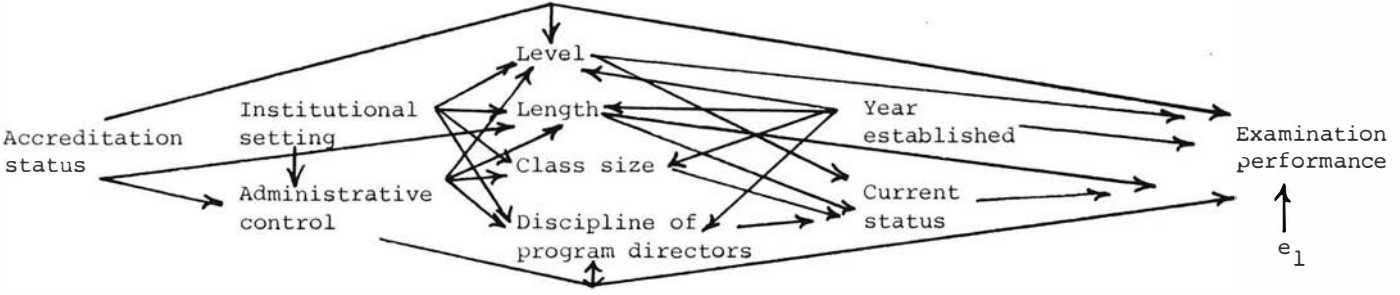
Independent Variables: Dependent Variable:

Sociodemographic Characteristics:



(e_1=other factors; e_2=program characteristics)

Program Characteristics:



(e_1=sociodemographic characteristics & other factors)

setting. Job function may be a consequence of the availability of employment in certain settings; and, conversely, certain settings may require that individuals perform particular functions (nurse practitioner or non-nurse practitioner skills).

These variables are affected by the candidate's length of experience as a nurse practitioner, type of nurse practitioner preparation, and highest level of education. Recent graduates tend to practice in nurse practitioner roles in general pediatric primary care settings. After several years in practice, they frequently narrow their focus to a particular subspecialty clinical area, or change their functional role (Dunn, 1981).

In regard to formal or informal preparation, the investigator expected that those candidates with informal preparation would be more likely to be employed as nurse practitioners in private practice settings, where they probably acquired their on-the-job training. Informal preparation is less transferable to other settings because of the specificity of training and experience (from one physician in one practice site) that could be expected. Additionally, informally prepared practitioners may have difficulty in obtaining positions as nurse practitioners in other settings, because of the current expectation of formal education and competition for positions with formally prepared practitioners.

Finally, level of education has been shown to affect decisions regarding employment setting and job function, with masters educated nurses less likely to be employed in nurse practitioner roles, and more likely to be employed in university/teaching settings (Sultz, Bullough et al., 1983).

3. Length of experience. Most previous research suggests that, above some minimum threshold, there are negative relationships between length of experience and cognitive performance (Downing & Maatsch, 1973; Dunn, 1981; Farrand et al., 1982; Maatsch, 1981; Mellsop, 1981; NAACOG Certification Corporation, 1980; Pawluk et al., 1976).

The investigator anticipated that this research would support those findings. That is, that candidates with several years of experience as nurse practitioners would achieve higher scores than those with less experience. However, a threshold would be reached at which the relationship became negative. The examination assesses candidates' knowledge of "textbook" primary care pediatrics; therefore, those with several years experience have an advantage. They should be able to recall and review the content from their programs and have other knowledge gained through experience, but have not been out of school long enough to be significantly affected by practice patterns outside academic settings. This is essentially a period of integration or consolidation (Bejar, 1983, p. 46).

Above the threshold, recall and review of program content is more difficult and the effects of experience gained in particular settings is more important. Length of experience as a nurse practitioner is also related to the candidate's age, length of experience as a RN, education, and sex. Relationships with age and length of RN experience are a function of time. Level of education influences graduates' functional roles and employment settings, and, therefore, the length of time they may be employed as a nurse practitioner. The gender (sex) of candidates also affects length of employment, due to competing family and child-rearing responsibilities.

In relation to length of RN experience, similar relationships with performance, as well as with age, sex, and educational level, could be expected. However, the intercorrelation with age may magnify negative effects above the threshold point.

4. Type of preparation. Other research suggests that formal versus informal preparation for a particular role may (NAACOG Certification Corporation, 1980, 1982) or may not (Maatsch, 1981) affect performance. It was anticipated that, for this sample, informally prepared candidates would not perform as well as formally prepared candidates. This could be attributed to the influence of structural variables in the formal programs (presumably more systematic and comprehensive methods of instruction, better quality of instruction, structured learning situation, formal evaluation of learning), and differences in the content presented. Formally prepared candidates may also have gained greater test sophistication (Hopkins & Stanley, 1981), as formal testing in the program is often similar to the format and content of the certification examination.

There are intercorrelations between type of nurse practitioner preparation and education and sex. It was expected that informally prepared candidates were more likely to be diploma level graduates. This expectation was based on the fact that most of these candidates received their on-the-job training in the early 1970s, before there were a large number of programs available, and that most nurses preparing for nurse practitioner roles at that time were diploma graduates (Sultz, Bullough et al., 1983).

In terms of relationships between preparation and gender (sex) of the candidate, this intercorrelation is probably a result of women's

competing roles within families. That is, early pediatric nurse practitioner programs were usually 16 weeks in length; required a full-time, 40 hour per week commitment (in class); and were not available in many geographic locations (Sultz, Bullough et al., 1983). Therefore, women with other obligations may have chosen to prepare informally, in a part-time practice setting more compatible with family responsibilities.

5. Level of education. Although some research indicates that there are no significant relationships between level of education and cognitive performance (Fullerton & Thompson, 1983; Hoekleman, 1975), others suggest the opposite (AANA, 1983; Dunn, 1981; Farrand et al., 1982; Fleming, 1979; IOM, 1983).

As with length of experience, it was expected that there was a particular threshold for educational level. Because of previous research, the investigator believed that that threshold would be at the baccalaureate level, and that candidates below and above that level of education would obtain lower scores. At the baccalaureate level, it was anticipated that candidates might have higher socioeconomic statuses and greater general ability than candidates whose highest level of education was a diploma or associate degree. The positive effects of higher socioeconomic status and more selective admissions procedures are well documented in general educational research (Lavin, 1965).

Above the baccalaureate level, it was expected that the positive effects of socioeconomic status and general ability might be reversed by the negative effect of lengthy postgraduate education (Mellsop, 1981), which is also related to candidate age. In addition, candidates' sex is related to level of education, for reasons previously mentioned.

6. Sex. Most nurses and nurse practitioners (especially in

pediatrics) are women. Gender is related to cognitive performance both directly and, in this research, indirectly. General educational research documents a direct relationship, with females the higher achievers (Lavin, 1965, pp. 43-44). In the research model, the effects of sex are also mediated by the candidate's level of education, type of preparation, and length of experience.

Social and economic factors have historically directed more women than men to nursing as an occupation. For women, the decision to enter nursing is influenced by perceptions that it is a choice compatible with childrearing (IOM, 1983, p. 90). In addition, being female also has an impact on an individual's ultimate level of educational preparation and labor force participation.

Persons who enroll in nursing programs tend to receive their education in or near their home community, and to subsequently practice there (IOM, p. 163). Thus, availability of particular types of educational programs in the community of residence is related to the type of basic preparation and decisions regarding further education.

Until the early 1970s, most RNs were diploma school graduates. With the growth of community colleges and greater access to baccalaureate programs in the 1970s, profound changes in nursing education patterns began to occur. In 1980, the distribution of RNs by highest level of education was: 54% diploma, 18% associate degree, and 20% bachelors degree. On the other hand, in 1981 the distribution for new graduates was: 17.2% diploma, 49.7% ADN, and 33.1% BSN (IOM, pp. 35, 55).

Although the number of masters and doctoral programs in nursing has increased dramatically in the past 10 years, access to them is limited because of their geographic locations. For example, in 1980, one-half

of all masters graduates completed programs in 7 states: California, Illinois, Massachusetts, New York, Ohio, Pennsylvania, and Texas. In addition, one half of all students enrolled in masters programs in nursing during the 1981-1982 academic year were part-time students, due to financial and family constraints (IOM, p. 141).

While the labor force participation of RNs is greater than for women in general (76.4% for licensed RNs in 1980), one third of those RNs are working part-time, and another 388,000 are inactive (no current license). There is a documented relationship between women's labor force participation and higher educational levels. Today RNs participate in the labor force at a rate similar to that for all college educated women (IOM, pp. 54, 202-203).

Educational program variables. As was true for the sociodemographic variables, certain characteristics of the nurse practitioner educational programs (in terms of context, structure, and process) were not available for the study population. Program variables that were available and were included in the model were derived from theoretical assumptions and sometimes scanty empirical evidence. To elaborate, the examination performance of candidates and the average performance of educational programs are directly related to programs': year of establishment, current status, educational level, length, class size, discipline of directors, administrative control, setting, and accreditation status. There are also relationships among these independent variables.

1. Year established. The year in which programs were established influenced their organization and content. Two years after the organization of the Association of Faculties of Pediatric Nurse Associate/Practitioner Programs (1975), comprehensive behavioral objectives,

curriculum content, and structural guidelines were published. Before this time, programs were based on very general guidelines that had been developed in 1971 (Guidelines, 1971).

In 1976, the Division of Nursing (DHEW) issued additional guidelines for those programs seeking grant support from the federal government. Because of these events during 1975 and 1976, the investigator anticipated that programs established after 1976 would produce graduates who would perform better than graduates of programs that were discontinued before that date.

The year the program was established also may affect the program's current status, level, length, class size, and the discipline of program directors. Programs that were established after 1976 are less likely to be discontinued than programs established before that time (Sultz, Bullough et al., 1983). Early programs were typically 16 week certificate programs with wide variations in class size and physician or physician and nurse directors. Programs established after 1976 were influenced by the AFPNA/PP and DHEW guidelines and the trend toward masters level preparation and nurse practitioner program directors (Sultz, Henry, Kinyon et al., 1983a, 1983b).

2. Current status. It was expected that candidates whose programs were still active at the time they took the exam would achieve higher scores than graduates of programs that had been discontinued. Other research has demonstrated this effect (AANA, 1983; Fleming, 1979), which is probably related to the quality of the educational program and its ability to obtain continued sources of funding. The intercorrelations that affect the year of establishment variable also affect current status. In addition, there are intercorrelations between current status

and year of establishment.

3. Educational level. Cognitive performance is influenced by the level of the program from which candidates graduated. This result has been shown in previous research (AANA, 1983; Dunn, 1981; Farrand et al., 1982; Fleming, 1979; NAACOG Certification Corporation, 1980).

The investigator expected that graduates of masters level programs would perform better than graduates of certificate level programs. This was expected because students in masters programs enter with a bachelors degree, while certificate students' prior education is varied. Therefore, this discrepancy is believed to be related to the educational backgrounds of masters program graduates, rather than differences in the content or quality of their programs.

The level of the educational program is affected by its setting and accreditation status. Today, all masters programs and most certificate programs are located in NLN accredited schools of nursing. Those programs that are not accredited, or are accredited by the ANA, are usually certificate programs that are located in schools of medicine or in other community settings (Sultz, Bullough et al., 1983; Sultz, Henry, Kinyon et al., 1983a).

4. Program length. It was anticipated that this variable would have a threshold below and above which graduates obtained lower scores. That threshold should be somewhere between 5-9 months. With shorter programs, there is difficulty presenting the necessary content and assimilating that content. Pediatric programs that are longer than 9 months are more likely to be masters level, and do not necessarily provide more actual hours of content or clinical practice (Sultz, Bullough et al., 1983).

In addition, shorter programs are more likely to be certificate level programs that are located outside of NLN accredited schools of nursing. Previous research on the effects of program length are conflicting, with some researchers demonstrating positive relationships (Fleming, 1979) and others demonstrating negative relationships (Mellsop, 1981).

5. Class size. Although previous research has reported no relationship between class size and performance (Martin et al., 1980), it should be related for this sample. The investigator believes that class sizes of 8-10 students assure lower student-faculty ratios and facilitate interaction between students and faculty and among students. The class size of programs is affected by the setting of the program and the administrative control of the program. Programs administered by schools of nursing in university or college settings usually have specific requirements in regard to minimum enrollments and faculty work load (and therefore, student-faculty ratios).

6. Discipline of directors. No research was located that described effects between performance and the discipline of program directors. It was expected that programs with nurse and physician co-directors would produce graduates who achieved higher examination scores. The examination content assesses candidates' knowledge of the medical and nursing components of the nurse practitioner role; therefore, programs administered and taught by co-directors should be better able to integrate those components of the role.

The discipline of directors is obviously related to the program's setting, administrative control, and accreditation status. While it was formerly common for programs controlled by schools of nursing in univer-

sity settings to have only a nurse director, there is now a trend toward co-directors in masters programs (Sultz, Bullough et al., 1983).

7. Administrative control. No research was found that looked at relationships between performance and administrative control of programs. Although the investigator anticipated that programs with co-directors would produce higher scoring graduates, joint administration at the institutional level is not likely to produce the same results. Programs that are jointly administered or are administered by medicine are not typically found at the masters level, in NLN accredited schools of nursing, or in university settings.

8. Accreditation status. The investigator expected that the accreditation status of the program would affect the performance of its graduates, and is an indirect measure of the quality of the program. It was anticipated that unaccredited programs would produce graduates who were lower achievers than NLN or ANA accredited programs. Also, graduates of NLN accredited programs should perform at higher levels than graduates of ANA accredited programs, because most ANA accredited programs are either under the administrative control of medicine or are located in settings other than universities.

9. Institutional setting. Based only on theoretical assumptions, it was expected that programs located in college or university settings should have graduates who perform at higher levels than other graduates. This should be true because academic settings have the advantages of education as a primary goal, full-time faculty, and generally greater access to clinical facilities. Other research (AANA, 1980; Fleming, 1979) also suggests that graduates of military specialty nursing programs tend to be high achievers. There are also relationships between

setting and accreditation status, as previously described.

Summary

In summary, the conceptual model for this research was derived from educational and social process models previously reported in the literature, and from empirical evidence in other research. Chapter V describes the research methodology, including design, sample, data base, and data analysis.

Chapter V. Methods

This research involved an analysis of data related to pediatric nurse practitioners, including their sociodemographic characteristics, characteristics of their educational programs, and their performance on a specialty certification examination. The methodological approach was descriptive and correlational.

Design

Data for six cohorts of examinees (1977-1982) were examined according to year of examination as well as in aggregate form. In this investigation, examination performance (standard composite scores) was the dependent variable. The independent variables were: (a) sociodemographic characteristics of examinees (sex; age; highest level of education; type of nurse practitioner preparation; months experience as a registered nurse; months experience as a nurse practitioner; and current work setting and job function), and (b) nurse practitioner program characteristics (level of the program; current program status; location and setting; administrative control; discipline of program directors; accreditation status; number of students per class; length in hours and weeks; and year program established).

Sample

The sample for this research was 3,206 pediatric nurse practitioners who took the specialty certification examination administered

by the National Board of Pediatric Nurse Practitioners and Associates between 1977 and 1982. All members of the sample were registered nurses with additional formal education, or its equivalent, to prepare them for expanded roles in the provision of primary health care to children. It was estimated that the sample represented between 55-58% of the total population of pediatric nurse practitioners.

Data Base

The investigator secured access to the data used in this research by submitting a formal proposal to the National Board. That proposal was approved by the Executive Board of that organization, contingent on the investigator's agreement to comply with the Board's policies regarding confidentiality and research and publication. (See appendix for November 2, 1982 letter of approval, National Board policies, and the investigator's agreement to comply.)

This section provides background information on the National Board examination, as well as information obtained and collected from the National Board and its testing agency, the National Board of Medical Examiners (NBME).

National board examination. The National Board was incorporated in 1975 as an independent agency composed of representatives of three organizations: NAPNAP, AAP, and the AFPNA/PP. The first National Qualifying Examination (NQE) for pediatric nurse practitioners was administered in 1977, and annually thereafter. By January 1983, the National Board had administered 3,387 examinations to 3,206 individuals (181 repeats) and had certified 2,747 candidates.

Over the 6 years in which the examination has been offered, the number of candidates tested and certified has declined. (See Table 4.)

There are several reasons why the 1977 cohort was at least 34% larger than other groups. The population of potential candidates was largest in 1977, with some practitioners out of school for as long as 10 years. Subsequent cohorts reflect the decreasing population of uncertified practitioners available for testing, and fewer graduates due to decreasing numbers of pediatric educational programs, increased length of programs, and fewer classes per year.

Table 4
National Qualifying Exam: Numbers Tested & Certified, 1977-1982

Year	# Tested	Reference Group[a]	#Repeats	#Certified
1977	943	942 (99.9%)	--	823
1978	587	488 (83.1%)	40	473
1979	620	503 (81.1%)	35	490
1980	481	365 (75.9%)	38	384
1981	464	345 (74.3%)	40	344
1982	292	260 (89.7%)	28	233
Totals:	3,387	2,903 (85.7%)	181	2,747

Note. Based on information obtained from the National Board and the NBME.
[a]Reference group=those formally prepared candidates being tested for the first time.

The NQE was developed with assistance from the NBME. The NBME was founded in 1915 as a voluntary and unofficial testing agency, and initially devoted its efforts to evaluation of medical student performance. In 1961, the American Board of Pediatrics initiated a consultative affiliation with the NBME for assistance with certification examination development for pediatricians. That relationship became the prototype for subsequent NBME affiliations with other specialty medical groups. In 1972, the NBME became involved in examination development for physicians' assistants, and in 1975 its relationship with the National

Board began (Report of the Committee on Goals, 1973, pp. 27-30).

The NQE evaluates the competency of nurses for entry into practice as pediatric nurse practitioners. The purpose of the examination is to measure candidates' knowledge of the nursing and medical components of the pediatric nurse practitioner role, and their ability to apply such knowledge in the provision of health care to children. Eligibility requirements include current licensure as a registered nurse and graduation from a formal educational program that has been approved by the National Board (National Board brochure, 1979).

Applicants who were not graduates of formal programs were considered on an individual basis from 1978-1981,* and had to submit evidence of equivalent training and practice under supervision. *(Because of extraordinary circumstances, one informally prepared candidate was permitted to sit for the NQE in 1977, and four informally prepared candidates were tested in 1982.)

A total of 303 informally trained nurses were tested during this period. Eighteen percent (18%; 54/303) of those candidates classified as informally prepared were actually graduates of formal programs that were not approved by the Board, either because the program had never sought approval or the program was deficient in total hours of pediatric content or clinical (N. Dickenson-Hazard, personal communication, August, 1983).

This is a norm-referenced examination. That approach is considered to be most useful in assessing a candidate's general knowledge or understanding of the subject area to be measured. It is designed to establish the examinee's relative standing in relation to the performance of other examinees, by comparing individual scores with the average per-

formance of an appropriate normative or reference group (Hughes, 1982). The reference group for each cohort of examinees was those formally prepared candidates who were being tested for the first time (see Table 4).

The blueprint used for test item development was a grid that divided four tasks in seven weighted categories for five pediatric age groups. The tasks were: data collection, assessment, management, and basic knowledge (techniques/principles). Categores included: history and physical examination, health maintenance, growth and development, common problems, culture, health care delivery, and pediatric drugs. The pediatric age groups were: unspecified, conception to newborn, infancy, childhood, and adolescence (N. Dickenson-Hazard, personal communication, November, 1982).

From 1977 to 1979 the examination format was 180 multiple choice questions (MCQs; one best answer) and 5 patient management problems (PMPs) containing 163 items. Because of the expense of using PMPs, they were discontinued. (PMPs are purported to measure skill at problem-solving. Holzemer, Farrand, and Schleutermann (1981), however, suggest that MCQ examinations are equal if not superior to PMPs in measuring the problem-solving skills of nurse practitioners.) The format between 1980 and 1982 was 275 multiple choice questions (N. Dickenson-Hazard, personal communication, November, 1982).

This is a timed, 4½ hour examination, but it cannot be considered a "speeded" test. If less than 85% of examinees have an opportunity to respond to all items, the test should be considered speeded. In such instances, internal consistency estimates should not be used because they tend to be spuriously high (Martuza, 1977, p. 131).

The maximum possible score is 800, and only correct answers are

counted. Individual raw scores are converted to Z scores (raw score minus mean/standard deviation). The negative numbers found with Z scores are eliminated by conversion to T scores (T=100Z + 500), so that the scaled scores have a mean of 500 and a standard deviation of 100. The pass/fail level is set one standard deviation below the mean (Kane, 1980b, pp. 84-86; Report of the NCHCA, 1981, p. 16).

Studies of the examination's validity and reliability have been conducted by psychometricians on the NBME staff. As with most other certification examinations in the health occupations, the NQE is said to have content validity. Other types of validity--construct, concurrent, predictive, differential, discriminant--have not been established. This is also typical, although the ANA certification examinations purport to have construct and concurrent validity as well as content validity (Report of the NCHCA, 1981, pp. 55-62).

Most certifying agencies also report that reliability has been established for their examinations (Report of the NCHCA, 1981, pp. 55-62). Table 5 presents the reliability, p, and r values for the NQE for 1977-1982. The statistics for each year were computed on that year's reference group (see Table 4), not on the total candidate population.

The item analysis for the examinations indicates that there is not a great deal of variation in the distribution of scores on the examination (homogeneous test) and that the examinee population (at least the formally prepared first-time takers) is relatively homogeneous. For example, the mean p, which is a measure of the level of difficulty of the exam (percent answering correctly at a standard score of 500), is consistently high, demonstrating that most of these examinees answered

Table 5

Performance of the 1977-1982 National Qualifying Examinations

Examination section & year	k[a]	Mean p	Mean r	Reliability alpha	Reliability Mosier
MCQ					
1977	176	.70	.27	.84	.86
1978	173	.70	.23	.78	.81
1979	173	.67	.26	.82	.85
1980	258	.65	.22	.84	--
1981	253	.66	.24	.87	--
1982	259	.71	.24	.86	--
PMP					
1977	136	.80	.26	.72	.86
1978	255	.79	.23	.71	.81
1979	162	.75	.24	.69	.85

Note. Information obtained from the NBME.

[a]k=number of test items on which statistics computed; numbers are not consistent with the total number of questions on each exam because some MCQ items were discarded and not all PMP options were gradable.

the majority of test items correctly: overall average 68% MCQ and 78% PMP. With norm-referenced examinations, you would like to have a mean p of about .5; however, p values in the .3 to .7 range are acceptable (Hopkins & Stanley, 1981, p. 282; Martuza, 1977, pp. 178-179).

Likewise, the mean r values (item-test point biserial correlations; Pearson r used to measure the degree of correlation between test items and the total test), which are influenced by the homogeneity of the examinee population and the homogeneity of the test items, are consistently low. Martuza (1977) considers this value to be the single most important characteristic for norm-referenced tests. Because this is a measure of the extent to which the examination items discriminate between high and low scoring candidates, low values mean that there is limited discrimination between high and low scorers in the reference group (presumably because of a homogeneous population and a homogeneous test).

(The maximum r value is 1: all high scorers answer each item correctly, all low scorers answer incorrectly.) (NAACOG Certification Corporation, 1980, p. 3; Martuza, 1977, p. 180)

The high alpha reliability values show that the exams are internally consistent. That is, that the test items are measuring the same content domain; therefore, the examinations are relatively homogeneous. This homogeneity also influenced the values for the biserial correlations (mean r) (Martuza, 1977, p. 128).

Mosier reliability was calculated for the examinations between 1977-1979, when they were composed of both MCQs and PMPs. This is a measure of the reliability of composite scores, according to a weighted average of the reliability of the components. The weighting for the multiple choice and patient management sections was 3 MCQ + PMP/4=total test composite (NBME internal document, 1983). If the components or sections of an examination are not related, the composite will be low. If the components are related (intercorrelated), the composite reliability will be greater than the average of the component reliabilities (Guilford & Fruchter, 1973, p. 436). As seen in Table 5, the reliabilities of the MCQ and PMP sections of the examinations, and their weighted component averages (1977--.81; 1978--.76; 1979--.79) are lower than the Mosier values, indicating that the two sections of the examination were intercorrelated.

The values for the mean p, mean r, and alpha reliability are similar to those for the NAACOG Certification Corporation's 1980 and 1981 examinations for obstetric-gynecologic nurse practitioners (the only comparable data found by the investigator). For those examinations, p values were .66 and .62; r values were .29 and .22; and, re-

liability estimates were .87 and .89, for 1980 and 1981 respectively (NAACOG Certification Corporation, 1980, p. 3; 1982, p. 2).

To summarize these findings, these examinations are internally consistent and measure a particular content domain--that knowledge determined as relevant for minimally competent practice as a pediatric nurse practitioner. The population of formally prepared candidates taking the examinations for the first time and the examinations themselves are relatively homogeneous, with restriction of variance in the distribution of scores.

NBME tape. Data regarding sociodemographic characteristics of 3,206 examinees and their scaled composite scores were obtained from the NBME, Philadelphia. That information was coded and placed on a computer tape that the investigator purchased, subsequent to editing by the NBME staff. Complete data were not available for all examinees: Some candidates neglected to provide all information that was requested by the National Board on the form completed at the time they applied to take the examination. (Detailed information about the data available and data processing for analysis is contained in the appendix.) The NBME tape was transferred to a permanent computer file by the investigator, which was created by using the Wylbur^{R} data entry system and the Statistical Analysis System (SAS) program for data analysis.

Collected data. Information regarding 114 educational programs was collected by the investigator from files at the National Board's Rockville, Maryland office between January 17-19, 1983. Two sources of file information were used: data submitted by program directors when they sought initial approval of their programs by the National Board, and data obtained by the National Board in its March, 1982 survey of program

directors. In that survey, current information was obtained for 53 programs (National Board Survey Results, internal document, January 25, 1982).

The National Board classifies educational programs according to their current activity status and status with the Board. That is, programs are classified as: currently active and approved by the Board, inactive/discontinued before 1976 but retroactively approved by the Board, inactive/discontinued after 1976 and approved by the Board, and active but not approved by the Board (candidates applied for individual consideration). Complete information was not available for all programs, particularly those in the discontinued/retroactively approved category.

In addition, the information available for informally prepared candidates was limited. Although the investigator could identify informally prepared candidates according to their program code on the NBME tape, no other information regarding specific characteristics of their training was avialable for analysis. As previously described, summary information was obtained for those candidates classified as informally prepared who were actually graduates of formal programs that were not approved by the Board (N. Dickenson-Hazard, personal communication, August, 1983).

Supplementary sources were used to obtain information about the accreditation status and level of programs. The sources regarding accreditation status were: (a) a listing of baccalaureate and masters degree programs in nursing accredited by the National League for Nursing (NLN, 1982), and (b) a directory of nurse practitioner certificate programs accredited by the American Nurses Association (ANA, 1982).

These documents were also used to determine whether programs were certificate or masters level, as was (c) a directory of expanded role programs published by DHHS, Division of Nursing (1982).

All data regarding educational program characteristics were coded by the investigator and a computer file was created, again using the Wylbur^{R} data entry system and the SAS program for data analysis. This data file was then added to the file created for the NBME tape data, and the two files were later merged for some of the analyses.

Confidentiality. As previously mentioned, the investigator agreed to comply with the National Board's policies on confidentiality and research and publication in order to acquire access to this data base. The confidentiality of examinees was protected by removing identifying information, such as name and social security number, from the NBME tape before it was sent to the investigator. Individual examinees were identified only by a code number on the edited tape; therefore, no sociodemographic characteristic or examination score can be connected with a particular person. Likewise, the investigator grouped educational programs according to program codes and other variables, so that no particular program could be linked with examination scores or individual examinees.

No other person or institution has been or will be granted access to this data base without the express consent of the National Board; and, the NBME tape will be returned to the National Board upon completion of this research. Fianlly, the investigator will prepare a report of this research for the National Board, and will grant them the right to review any manuscript subsequently prepared for publication.

Data Analysis

Three data files were used in this analysis. They included: (a) an examinee file containing the sociodemographic variables and examination scores; (b) a program file containing the educational program variables and average program scores; and (c) a merged examinee file containing sociodemographic variables, program variables, and examination scores for each examinee. The SAS program for computer analysis was used to examine the data by descriptive and multivariate statistical techniques.

In addition to aggregate analysis, the examinee and merged files were subdivided. For those files, subsets were created for each cohort of examinees (by year of examination) and for the first-time takers and repeaters. Cohort as well as aggregate analysis was performed to determine whether there were significant effects related to group membership (Cronbach, 1982, p. 186). Data for first-time takers and repeaters were analyzed separately because of evidence from previous research that repeaters are likely to be a norm group different than the group of candidates taking the examination for the first time, with those differences affecting both outcomes and interpretation of results (Fleming, 1979; Fullerton & Thompson, 1983; Mellsop, 1981).

Preliminary analyses involved construction of frequency distributions for each independent variable to describe the examinee population and characteristics of the educational programs. Frequency distributions were also constructed for the dependent variable, to describe variations in examination scores. Aggregate and cohort analyses were performed, and crosstabulations and analysis of variance were used to identify differences between groups. Correlational analyses were then

performed to examine relationships between each independent variable and the dependent variable, and to explore intercorrelations between the independent variables.

Several regression analyses were prepared. Equations were constructed to examine the effects of the sociodemographic variables on examination performance, for the aggregate as well as the six cohorts. Other equations were constructed to determine the effects of program variables on average examination performance by program. A final analysis regressed the merged sociodemographic and program variables on examination performance, again for the aggregate and six cohorts. In constructing these regression equations, dummy variables were created for those independent variables that were not measured at the interval level (Nie, Hull, Jenkins, Steinbrenner, & Brent, 1975, pp. 373-374).

The results of these analyses are presented in Chapters VI and VII. Chapter VI discusses the descriptive analyses and Chapter VII describes the detailed multivariate analyses.

Chapter VI. Descriptive Analysis

This chapter describes the sociodemographic characteristics, educational program characteristics, and examination scores of pediatric nurse practitioners who took the National Board's qualifying examination (NQE) between 1977-1982. The sample included 3,206 individuals, who were tested a total of 3,387 times (181 repeats).

Sociodemographic Characteristics

The sociodemographic variables available for examinees included: NQE status, examination year, sex, age, RN experience, PNP experience, highest level of education, type of PNP preparation, current function, and employment setting.

Variable Definitions

The definitions that follow include both continuous and discrete variables. For some of the analyses the continuous variables were collapsed and assigned a numerical scale. Likewise, dummy variables were created for discrete variables in some of the analyses.

1. NQE status. Examinees were classified according to whether they were taking the examination for the first time (1) or were repeating the examination (0).

2. Examination year. The year in which examinees took the NQE was coded as 1977, 1978, 1979, 1980, 1981, or 1982.

3. Sex. Categories for this variable were female (1) or male (0).

4. Age. Birthdates available for examinees were converted to age in years. These ages were collapsed into 5 age groups (1-5; see Table 7) for some of the analyses.

5. RN experience. This was the number of months of employment as a registered nurse, excluding time employed as a nurse practitioner. Experience was collapsed into 5 categories for some analyses (scale 1-5; see Table 7). Years as well as months are given to make interpretation easier.

6. PNP experience. This variable described the number of months of employment as a nurse practitioner, exclusive of other RN experience. For some analyses, PNP experience was collapsed into 7 categories (scale 1-7; see Table 7) and was expressed in years to facilitate interpretation.

7. Highest level of education. Examinees were classified in 9 categories, according to whether their highest level of education was a non-nursing associate degree (1), associate degree in nursing (2), diploma in nursing (3), non-nursing bachelors degree (4), bachelors degree in nursing (5), non-nursing masters degree (6), masters degree in nursing (7), non-nursing doctorate (8), or doctorate in nursing (9). It should be noted that the numerical scale (1-9) is an approximate proxy for number of years of education. Although information about examinees' basic nursing education was obtained, it could not be used for analysis (see appendix for further explanation).

8. PNP preparation. Examinees were classified as formally prepared (1)--graduates of formal nurse practitioner programs--or informally prepared (0)--on-the-job training or graduates of programs not approved by the National Board.

9. Current function. This variable classified examinees according

to whether they were currently performing the functions of a nurse practitioner (1) or were not performing those functions (0).

10. Employment setting. There were 9 categories of current employment setting for the examinees. Seven (7) settings were for those functioning as nurse practitioners, and 2 settings were for those employed as RNs or not employed (see Table 7). Settings were collapsed into 2 categories for some analyses (outpatient clinic=1; other setting=0).

Table 6 provides information on the sociodemographic variables for the total population tested, in terms of available and missing data (see appendix for further discussion of missing data).

Table 6
Sociodemographic Variables: Data Available and Missing, 1977-1982 (N=3,387)

Variable Name	Data Available	Data Missing[a]
NQE status	3,387	0
Exam year	3,387	0
Sex	3,385	2
Age	3,374	13
RN experience	3,204	183
PNP experience	3,349	38
Highest education	3,380	7
PNP preparation	3,387	0
Current function	3,372	15
Employment setting	3,381	6

[a]All missing data were confined to first-time takers except RN experience, where 10 missing entries were for repeaters.

Examinee Profile by Person

Of the 3,387 examinees tested, 3,206 were taking the examination for the first time and 181 were repeats. The profiles for first-time takers and repeaters are described separately.

Table 7
First-Time Examinees: Frequencies & Categories for Sociodemographic Variables (N=3,206)

Variable Name (Scale)	Frequency	Percent	Mean
Sex	N=3,204		
female (1)	3,148	98.2	
male (0)	56	1.8	
Age: years	N=3,193		34.6
20-24 (1)	95	3.0	
25-34 (2)	1,785	55.9	
35-44 (3)	872	27.3	
45-54 (4)	368	11.5	
55+ (5)	73	2.3	
RN experience: months/years	N=3,033		96.8/8.1
less than 12/1 (1)	186	6.1	
13-60/1-5 (2)	1,162	38.3	
61-120/5-10 (3)	886	29.2	
121-180/10-15 (4)	439	14.5	
181+/15+ (5)	360	11.9	
PNP experience: months/years	N=3,168		24.9/2.1
none (1)	423	13.4	
less than 12/1 (2)	990	31.2	
13-24/1-2 (3)	599	18.9	
25-36/2-3 (4)	444	14.0	
37-48/3-5 (5)	266	8.4	
49-60/4-5 (6)	206	6.5	
61+/5+ (7)	240	7.6	
Highest education: by degree	N=3,199		
non-nursing associate (1)	5	0.2	
nursing associate (2)	236	7.4	
nursing diploma (3)	1,020	31.9	
non-nursing bachelors (4)	131	4.1	
nursing bachelors (5)	1,137	35.5	
non-nursing masters (6)	111	3.4	
nursing masters (7)	550	17.2	
non-nursing doctorate (8)	5	0.2	
nursing doctorate (9)	4	0.1	
PNP preparation	N=3,206		
formal (1)	2,903	90.6	
informal (0)	303	9.4	
Current function	N=3,191		
PNP (1)	2,620	82.1	
non-PNP (0)	571	17.9	
Employment setting	N=3,200		
private physician (0)	732	22.9	
faculty member (0)	110	3.4	
community health (0)	323	10.1	
outpatient clinic (1)	1,050	32.8	
school system (0)	185	5.8	
hospital inpatient (0)	94	2.9	
other setting, NP (0)	230	7.2	
employed as RN (0)	329	10.3	
not employed (0)	147	4.6	

First-time takers. Most first-time takers were tested only once (3,055/3,206=95.3%); 151 first-time takers later repeated the exam (4.7%). As previously mentioned, the largest number of examinees was tested in 1977 (29.4%), with subsequent cohorts reflecting smaller numbers of uncertified practitioners available for testing, and smaller numbers of program graduates due to a decrease in the number of programs and increase in program length.

First-time takers were usually females between the ages of 25-34 years (mean=34.6 years; interval level measurement). They typically had between 1-5 years of RN experience (mean=8.1 years; interval level), and between 1-24 months of experience as PNPs (mean=2.1 years; interval level). Most examinees had bachelors degrees and received their nurse practitioner preparation in a formal educational program. They were currently functioning as nurse practitioners, usually in a hospital or other outpatient clinic setting (see Table 7 for categorical breakdowns).

Repeaters. Between 1978-1982 the NQE was repeated 181 times by 151 people. Those examinees who repeated once number 125, while 22 people repeated twice and 4 people repeated three times. The number of repeats is fairly evenly distributed over the 5 examination years.

This profile describes both the repeaters (individuals) and the repeats (total entries for repeaters). (See Table 8.) In regard to repeaters, most were females between the ages of 35-44 years (mean=40.6 years). Typically, they had more than 10 years of experience in nursing (mean=11.1 years) and more than 24 months of experience as nurse practitioners (mean=3.5 years). Repeaters usually had diplomas or bachelors degrees and received their nurse practitioner preparation in a formal program. These examinees were functioning as nurse practitioners, most

Table 8

Repeating Subset: Frequencies & Categories for Sociodemographic Variables of Individual Repeaters (N=151) & Repeating Group (N=181)

Variable Name (Scale)	#Repeaters (%)	#Repeats (%)	Mean
Sex	N=151	N=181	
female (1)	150 (99.3)	178 (98.3)	
male (0)	1 (0.7)	3 (1.7)	
Age: years	N=151	N=181	40.6
20-24 (1)	2 (1.3)	2 (1.1)	
25-34 (2)	49 (32.5)	52 (28.7)	
35-44 (3)	55 (36.4)	66 (36.5)	
45-54 (4)	33 (21.8)	48 (26.5)	
55+ (5)	12 (8.0)	13 (7.2)	
RN experience: months/years	N=141	N=171	133.7/
less than 12/1 (1)	9 (6.4)	10 (5.9)	11.1
13-60/1-5 (2)	38 (26.9)	45 (26.3)	
61-120/5-10 (3)	28 (19.9)	33 (19.3)	
121-180/10-15 (4)	28 (19.9)	33 (19.3)	
181+/15+ (5)	38 (26.9)	50 (29.2)	
PNP experience: months/years	N=151	N=181	41.6/
none (1)	8 (5.3)	8 (4.4)	3.5
less than 12/1 (2)	17 (11.3)	18 (9.9)	
13-24/1-2 (3)	37 (24.5)	39 (21.6)	
25-36/2-3 (4)	28 (18.5)	33 (18.2)	
37-48/3-4 (5)	19 (12.6)	25 (13.8)	
49-60/4-5 (6)	14 (9.3)	17 (9.4)	
61+/5+ (7)	28 (18.5)	41 (22.7)	
Highest education: by degree	N=151	N=181	
non-nursing associate (1)	1 (0.7)	1 (0.6)	
nursing associate (2)	18 (11.9)	20 (11.0)	
nursing diploma (3)	58 (38.4)	63 (34.8)	
non-nursing bachelors (4)	18 (11.9)	22 (12.1)	
nursing bachelors (5)	40 (26.5)	59 (32.6)	
non-nursing masters (6)	8 (5.3)	8 (4.4)	
nursing masters (7)	7 (4.6)	7 (3.9)	
non-nursing doctorate (8)	1 (0.7)	1 (0.6)	
nursing doctorate (9)	0	0	
PNP preparation	N=151	N=181	
formal (1)	134 (88.7)	161 (88.9)	
informal (0)	17 (11.3)	20 (11.1)	
Current function	N=151	N=181	
PNP (1)	133 (33.1)	163 (90.1)	
non-PNP (0)	18 (11.9)	18 (9.9)	
Employment setting	N=151	N=181	
private physician (0)	27 (17.8)	37 (20.4)	
faculty member (0)	2 (1.3)	2 (1.1)	
community health (0)	16 (10.6)	22 (12.2)	
outpatient clinic (1)	71 (47.0)	81 (44.7)	
school system (0)	9 (6.0)	11 (6.1)	
hospital inpatient (0)	9 (6.0)	10 (5.5)	
other setting, NP (0)	3 (2.0)	3 (1.7)	
employed as RN (0)	11 (7.3)	12 (6.6)	
not employed (0)	3 (2.0)	3 (1.7)	

commonly in hospital or other outpatient clinic settings. Figures for the repeating subset are similar, and are presented in Table 8.

Comparisons between subsets. There were significant differences between first-time takers and the repeating subset on 6 of 8 sociodemographic variables; there were no differences according to sex or type of nurse practitioner preparation (see Table 9).

Table 9
Sociodemographic Variables: Crosstabulations for First-Time Taker & Repeater Subsets (N=3,387)

Variable Name	N	chi square	df	p
Sex	3,385	0.91	1	.9280
Age group	3,374	77.82	4	.0001
RN experience group	3,204	52.48	4	.0001
PNP experience group	3,349	93.51	6	.0001
Highest education	3,380	70.69	8	.0001
PNP preparation	3,387	0.51	1	.4762
Current function	3,387	7.51	1	.0061
Employment setting	3,381	29.32	8	.0006

On the average, repeaters were 6 years older than first-time takers, with a larger proportion in the over 44 age groups. The largest number of repeats was made by examinees in the 45-54 year age group. Repeaters had more experience as RNs (average +3 years) and as PNPs (average +1.4 years). Those examinees who repeated most frequently had more than 15 years of experience in nursing and more than 5 years of experience as nurse practitioners.

In terms of highest education, there were proportionally more diploma and associate degree nurses and fewer masters degree nurses in the repeater subset. Diploma nurses repeated the examination the largest number of times. All those examinees who repeated the NQE more than once were functioning as NPs; there were fewer non-PNPs in the repeater subset than in the first-time taker subset.

When these subsets were compared according to the employment setting of examinees, there were proportionally fewer candidates in the repeater subset who were employed as faculty members, in "other" settings, or were not employed. Finally, there was a large percentage of repeaters employed in inpatient hospital settings, but examinees employed in community health settings repeated the NQE most frequently.

Comparisons among cohorts. Because of the small number of repeaters, crosstabulations for that subset on sociodemographic variables by exam year produced values that may not be valid. There were, however, significant differences in NQE status according to exam year (see Table 10). The cohort analyses were performed on the total examinee population, and there were significant differences between groups for all sociodemographic variables except sex (see Table 10). Further, there were significant differences between group means when continuous variables (age, RN experience, PNP experience) were examined by analysis of variance (see Table 11).

Over the 6 examination years, the mean age of candidates has decreased. Duncan's multiple range test for analysis of variance (Daniel, 1974, pp. 194-195; Duncan, 1955) indicated that those examinees who were tested prior to 1980 were significantly older than those tested in 1980-1982. (Duncan's multiple range test is hereafter called Duncan's test.)

Likewise, the average number of months of RN and PNP experience has decreased over time. In the analysis of variance for RN experience, Duncan's test showed that the 1977 cohort had significantly more experience than those tested in later years. For PNP experience, Duncan's test demonstrated that the 1982 cohort had significantly less experience

than those in previous years.

Table 10
Sociodemographic Variables: Crosstabulations for Examinees by Year of Examination (N=3,387)

Variable Name	N	chi square	df	p
NQE status	3,387	82.32	5	.0001
Sex	3,385	2.77	5	.7359
Age group	3,374	77.19	20	.0001
RN experience group	3,204	50.51	20	.0002
PNP experience group	3,349	353.68	30	.0001
Highest education	3,380	475.86	40	.0001
PNP preparation	3,387	231.10	5	.0001
Current function	3,387	94.88	5	.0001
Employment setting	3,381	278.33	45	.0001

Table 11
Selected Sociodemographic Variables: Means & F Values by Exam Year

Year: (N=)	All (3,387)	1977 (943)	1978 (587)	1979 (620)	1980 (481)	1981 (464)	1982 (292)	F
Variable:								
Age: years	34.9	35.6	36.3	35.1	33.8	34.0	32.8	11.63*
RN experience: mos.	98.7	110.1	92.6	93.2	98.7	94.0	93.2	5.50*
PNP experience: mos.	25.8	27.7	27.5	24.2	24.1	23.8	18.7	3.93*

*$p < .05$

Also, the average level of education of examinees increased between 1977-1982. In 1977-1979 the typical level of education was a non-nursing bachelors degree; for 1980-1982 it was a bachelors degree in nursing. The proportion of the examinee population who were informally prepared PNPs increased gradually between 1977-1981. In 1981, the last year in which informally prepared candidates were permitted to sit for the exam, 20.5% of all those tested were informally prepared. (As mentioned previously, due to extraordinary circumstances one informally prepared candidate was permitted to take the exam in 1977, and four in 1982.)

There was also a steadily increasing proportion of the population of examinees who were not functioning as nurse practitioners; by 1982, 33.7% of those examined were in this category. Finally, between 1977-1982 there was evidence of movement out of community health and private physician employment settings to other settings. There were also decreases in the proportion of examinees employed in outpatient and "other" settings as NPs, and an increase in the proportion of candidates who were not employed.

Examinee Profile by Program

For this profile, examinees were described in terms of three major program characteristics: formal or informal program; certificate or masters level program; and active or inactive program. This was based on the characteristics of the first-time takers only, therefore, there was no subset comparison. In addition, since this was an aggregate-level analysis, there was no cohort (by exam year) description.

Formal versus informal program. The average informally prepared examinee (N=303) was about 2 years older than the average formally prepared examinee (N=2,903) and had more experience as a RN (+1.6 months) and PNP (+18.4 months). While the informally prepared examinee typically had a non-nursing bachelors degree, the formally prepared candidate had a bachelors degree in nursing. Type of preparation, however, had no effect on whether the examinee was functioning as a nurse practitioner (both were) or on the employment setting (both were likely to be in some setting other than an outpatient clinic). Crosstabulations indicated that the differences in age, PNP experience, and education were significant (see Tables 12 and 13).

Table 12
Sociodemographic Variables: Crosstabulations by Type of Nurse Practitioner Preparation (N=3,206)

Variable Name	N	chi square	df	p
Sex	3,204	2.29	1	.1304
Age group	3,193	14.31	4	.0064
RN experience group	3,033	8.43	4	.0771
PNP experience group	3,168	176.50	6	.0001
Highest education	3,017	89.53	8	.0001
Current function	3,193	3.32	1	.0683
Employment setting	3,206	1.41	1	.2348

Table 13
Selected Sociodemographic Variables: Means & F Values by Type of Nurse Practitioner Preparation

Type of Preparation:	All	Formal	Informal	F
Variable:				
Age: years	34.6	34.4	36.7	21.40*
RN experience: mos.	96.8	96.6	98.2	0.09
PNP experience: mos.	24.3	22.5	40.9	162.32*

*$p < .05$

Active versus inactive program. Like the informally prepared examinee, graduates of inactive programs were older (+3.4 years) and had more RN (+24.5 months) and PNP (+13.5 months) experience than examinees whose programs were still active. There were no apparent differences in sex, education, function, or employment setting. As Table 14 indicates, however, there were significant differences in these areas. The inapparent differences were related to the fact that inactive program graduates had lower levels of education, were more likely to be women, and were more likely to be functioning as nurse practitioners in outpatient clinic settings.

Table 14

Sociodemographic Variables: Crosstabulations by Program Status

Variable Name	N	chi square	df	p
Sex	2,902	6.90	1	.0086
Age group	2,892	123.38	4	.0001
RN experience group	2,750	62.25	4	.0001
PNP experience group	2,871	224.41	6	.0001
Highest education	2,732	59.84	8	.0001
Current function	2,892	40.97	1	.0001
Employment setting	2,903	24.72	1	.0001

Table 15

Selected Sociodemographic Variables: Means & F Values by Program Status

Program Status:	All	Active	Inactive	F
Variable:				
Age: years	34.4	33.2	37.1	157.27*
RN experience: mos.	96.6	89.1	113.6	50.06*
PNP experience: mos.	22.5	18.3	31.8	225.05*

*$p < .05$

Masters versus certificate program. Compared to the masters program graduate, the average certificate program graduate had less education (BSN versus MSN), was older (+4.1 years), and had more RN (+31 months) and PNP (+8.8 months) experience. Graduates of both types of programs were typically functioning as nurse practitioners and were employed in some setting other than an outpatient clinic. There were significant differences between these groups in regard to age, experience, education, and function (certificate graduates were more likely to be functioning as nurse practitioners). (See Tables 16 and 17.)

Table 16

Sociodemographic Variables: Crosstabulations by Program Level

Variable Name	N	chi square	df	p
Sex	2,902	2.09	1	.1484
Age group	2,892	66.59	4	.0001
RN experience group	2,750	48.62	4	.0001
PNP experience group	2,871	57.78	6	.0001
Highest education	2,732	241.66	8	.0001
Current function	2,892	31.36	1	.0001
Employment setting	2,903	0.68	1	.4094

Table 17

Selected Sociodemographic Variables: Means & F Values by Program Level

Program Level:	All	Masters	Certificate	F
Variable:				
Age: years	34.4	30.8	34.9	78.13*
RN experience: mos.	96.6	69.2	100.3	39.60*
PNP experience: mos.	22.5	14.8	23.6	44.41*

*$p < .05$

Correlations between Sociodemographic Variables

Correlations between sociodemographic variables were performed by subsets, cohorts, and major program variables. These Pearson correlations were based on the maximum available data for each pair of variables examined.

Examinee subsets. Zero-order correlations between sociodemographic variables for first-time takers and repeaters are presented in Tables 18 and 19. For first-time takers, there were a number of correlations that were significant at $p=.05$ or less; however, only two relationships were of moderate or high magnitude. Those were: a high magnitude, positive correlation between examinee age and RN experience, and a moderate magnitude, positive relationship between highest level of education and year of examination. In the repeating subset, there were fewer significant relationships, and only one of moderate or high magnitude:

the positive correlation between age and RN experience.

Table 18
First-Time Takers: Zero-Order Correlations between Sociodemographic Variables (N=2,880-3,206)

Variable Name:	Age (X2)	Experience: RN (X3)	Experience: PNP (X4)	Highest education (X5)	PNP preparation (X6)	Current function (X7)	Employment setting (X8)	Exam year (X9)
Sex (X1)	.02	.05*	.00	.02	-.03	-.02	-.02	.02
X2		.61*	.30*	-.01	-.08*	.10*	-.02	-.15*
X3			.07*	.00	.00	.08*	-.01	-.07*
X4				.01	-.22*	.26*	.14*	-.14*
X5					.16*	.01	.01	.45*
X6						-.03	.02	-.15*
X7							.31*	-.17*
X8								-.06*

*p<.05

Table 19
Repeating Subset: Zero-Order Correlations between Sociodemographic Variables (N=160-181)

Variable Name:[a]	X2	X3	X4	X5	X6	X7	X8	X9
X1	-.04	.14	-.14	.04	-.05	-.04	-.14	-.01
X2		.61*	.19*	-.06	-.23*	-.02	-.04	-.05
X3			.03	.02	-.06	.02	-.08	-.09
X4				-.15	-.14	.26*	.02	.22*
X5					.12	-.13	-.03	-.02
X6						-.06	.07	-.18*
X7							.22*	-.04
X8								-.12

[a]Definitions given in Table 18
*p<.05

Examinee cohorts. When correlations between sociodemographic variables were examined by year of examination, some significant relationships were found. These relationships, however, were not consistent across years.

The most consistent relationships were between examinee age, RN experience, PNP experience, and other sociodemographic variables. In

addition, while age was consistently related (positively) to RN experience and PNP experience, these two types of nursing experience were related to each other only in 1978 and 1979. For highest level of education, most significant correlations occurred in 1977, 1981, and 1982--generally negative relationships of low to moderate magnitude.

PNP preparation was most consistently related to examinee age (negative and low magnitude) and months of PNP experience (negative, moderate magnitude). The current function of examinees was also related to their PNP experience (positive, low to moderate magnitude) and their employment setting (positive, high magnitude).

Average program examinee. Correlational analyses were also performed between sociodemographic variables and three major program variables: formal versus informal program; active versus inactive program; and masters versus certificate program. Controlling for these program variables, however, made almost no difference in the direction, magnitude, or significance of relationships previously observed in the zero-order correlations.

The exception was a change in the relationship between PNP experience and highest education, with type of preparation controlled. The zero-order correlation was non-significant (r=.01). For formally prepared examinees, however, the correlation was significant ($p<.05$) and of low magnitude (r=.08). On the other hand, PNP experience was negatively related to level of education (r=-.19; $p<.05$) for the informally prepared examinees. In other words, among the examinees who were informally prepared, those with more PNP experience tended to have fewer years of education (according to degree).

Educational Program Characteristics

The characteristics of the nurse practitioner educational programs from which examinees graduated included: status, level, location, setting, administrative control, discipline of director(s), accreditation status, year established, class size, and hours and weeks in length.

Variable Definitions

As was true for sociodemographic variables, program variables were both continuous and discrete. For some of the analyses the continuous variables were collapsed and assigned a numerical scale. In addition, dummy variables were created for discrete variables in some analyses.

1. Status. Programs were classified as active (1) or inactive (0)--discontinued.

2. Level. There were 2 categories for program level: masters (1) and certificate (0). (See appendix for explanation of classification used for those programs that were both certificate and masters or had progressed from certificate to masters.)

3. Location. This variable described the state in which programs were located, and was used for descriptive purposes only.

4. Setting. There were 5 types of institutional settings (sponsors) for programs: hospitals, community agencies, colleges and universities, military facilities, and joint sponsorship (between hospitals, universities, military facilities, and/or community agencies). For some analyses, this variable was collapsed into 2 categories: college/university/joint (1) and other (0).

5. Administrative control. The administrative control of programs was vested in nursing, medicine, or was jointly held. This variable

was collapsed into 2 categories--nursing (1), other (0)--for some analyses.

6. Discipline of director(s). Program directors were either nurses, physicians, or nurses and physicians (joint co-directors). In some analyses, 2 categories were used: joint (1) and other (0).

7. Accreditation status. Programs were accredited by the ANA, the NLN, by both ANA and NLN, or were not accredited. Two categories were used for some analyses: accredited (1) or not accredited (0).

8. Year established. This variable described the year in which the program was established, and was collapsed into 3 categories (1966-1970; 1971-1975; 1976+) for some analyses.

9. Class size. This described the average number of students per class in the educational programs. Class size was collapsed into 3 categories (less than 8; 8-10; more than 10) for some analyses.

10. Length. Program length was subdivided by number of hours and number of weeks. For some analyses, each of these subdivisions was collapsed into 4 categories (see Table 20). In the discussion to follow, hours and weeks are converted to months to make interpretation easier.

Table 20 provides information on the educational program variables, in terms of available and missing data for each program (N=114) and each examinee (N=3,064: 2,903 first-time takers and 161 repeats). Note that these data do not include those examinees who were informally prepared (N=303).

Table 20
Educational Program Variables: Data Available & Missing, 1977-1982, by Program (N=114) & Examinee (N=3,064)

Variable Name	Programs		Examinees	
	Available	Missing	Available	Missing
Status	114	0	3,064	0
Level	114	0	3,064	0
Setting	114	0	3,064	0
Administration	103	11	2,955	109
Directors	101	13	2,940	124
Accreditation	113	1	3,051	13
Year established	90	24	2,652	412
Class size	76	38	2,590	474
Length: hours	62	52	1,834	1,230
weeks	87	27	2,715	349

Program Profile by Program

This section describes the characteristics of 114 formal educational programs from which examinees graduated (see Table 21). Although 63.6% of these programs were active (operational) at the time examinees were tested, 26 had become inactive by 1982 (52% of the programs classified as inactive in Table 21). Therefore, the overall percentage of active programs has decreased dramatically.

In regard to educational level, most programs were certificate level. Several programs became masters level between 1976-1982 (see appendix for explanation of their classification for analysis), and 2 additional programs became masters level in 1982-1983 (does not affect these data).

The typical nurse practitioner program was sponsored by a college or university, with administrative control vested in nursing, but with nurse and physician co-directors. These programs were usually established between 1971-1975, and were accredited by the National League for Nursing (NLN). For most programs, class sizes averaged 8-10 students (mean=9.2; measured at interval level); length was usually 4-9

months (mean=864.5 hours over 38 weeks; interval level measurements).

These programs were geographically located in 34 states, the District of Columbia, and West Germany (military facility). The largest proportion of programs were in the northeastern (N=35) and western (N=31) states, with fewer programs in the south (N=25) and midwest (N=23). California had the largest number of programs (18), followed by New York (11). There were 16 states without programs represented in this sample: Alaska, Delaware, Georgia, Hawaii, Idaho, Kentucky, Montana, Nebraska, New Hampshire, New Mexico, Nevada, North Dakota, Rhode Island, South Dakota, Vermont, and Wyoming.

Finally, the number of examinee-graduates per program ranged from 1 (11 programs) to 159 (1 program). Further analysis examined these programs by two major characteristics: their educational level and current status.

Educational level. When the characteristics of masters and certificate level programs were examined separately, some differences emerged (see Table 22). All masters programs were located in collegiate schools of nursing, and all but one was NLN accredited. In contrast, about one-half of the certificate programs were located in colleges or universities and the same number were not accredited. While virtually all masters programs were operational in 1983, less than half of the certificate programs were. The administrative control and direction of masters programs were typically vested in nursing; for certificate programs, there was greater diversity in program administration and direction. Also, the masters programs had larger class sizes and were longer in weeks compared to the average certificate program. As Table 23 indicates, these differences were statistically significant.

Table 21

Educational Programs: Frequencies & Categories for Program Variables (N=114)

Variable Name (Scale)	Frequency	Percent	Mean
Status	N=114		
active (1)	64	56.1	
inactive (0)	40	43.9	
Level	N=114		
masters (1)	24	21.1	
certificate (0)	90	78.9	
Setting	N=114		
college/university (1)	70	61.4	
hospital (0)	13	11.4	
community agency (0)	10	8.8	
military facility (0)	5	4.4	
joint sponsors (1)	16	14.0	
Administration	N=103		
nursing (1)	50	48.5	
medicine (0)	22	21.4	
joint (0)	31	30.1	
Directors	N=101		
joint (1)	42	41.6	
nursing (0)	51	50.5	
medicine (0)	8	7.9	
Accreditation	N=113		
NLN (1)	62	54.9	
ANA (1)	3	2.6	
NLN + ANA (1)	2	1.8	
not accredited (0)	46	40.7	
Year established	N= 90		
1966-1970 (1)	13	14.4	
1971-1975 (2)	59	65.6	
1976+ (3)	18	20.0	
Class size	N= 76		9.2
less than 8 (1)	22	28.9	
8-10 (2)	32	42.2	
more than 10 (3)	22	28.9	
Length: hours/months	N= 62		854.5/ 5.4
less than 640/4 (1)	22	35.5	
640-1440/4-9 (2)	35	56.5	
1441-1920/9-12 (3)	2	3.2	
1920+/12+ (4)	3	4.8	
Length: weeks/months	N= 87		38.1/ 9.5
16/4 (1)	10	11.5	
17-36/4-9 (2)	37	42.5	
37-48/9-12 (3)	20	20.0	
48+/12+ (4)	20	20.0	

Table 22

Educational Program Variables by Program Level

Program Level:	Masters (%)	Certificate (%)	Mean
Variable Name:			
Status (N=114)	N=24	N=90	
active	23 (95.8)	41 (45.6)	
inactive	1 (4.2)	49 (54.4)	
Setting (N=114)	N=24	N=90	
college/university	24 (100)	46 (51.1)	
hospital	0	13 (14.4)	
community agency	0	10 (11.1)	
military facility	0	5 (5.6)	
joint sponsors	0	16 (17.8)	
Administration (N=103)	N=24	N=79	
nursing	22 (91.7)	28 (35.5)	
medicine	0	22 (27.8)	
joint	2 (8.3)	29 (36.7)	
Directors (N=101)	N=24	N=77	
joint	6 (25.0)	36 (46.7)	
nursing	18 (75.0)	33 (42.9)	
medicine	0	8 (10.4)	
Accreditation (N=113)	N=24	N=89	
NLN	23 (95.8)	39 (43.8)	
ANA	0	3 (3.4)	
NLN + ANA	0	2 (2.2)	
not accredited	1 (4.2)	45 (50.6)	
Year established (N=90)	N=23	N=67	
1966-1970	1 (4.3)	12 (17.9)	
1971-1975	16 (69.6)	43 (64.2)	
1976+	6 (26.1)	12 (17.9)	
Class size (N=76)	N=19	N=57	11.5-M
less than 8	2 (10.5)	20 (35.1)	8.5-C
8-10	6 (31.6)	26 (45.6)	
more than 10	11 (57.9)	11 (19.3)	
Length: hrs/mos (N=62)	N=13	N=49	812.2-M
less than 640/4	4 (30.8)	20 (40.8)	878.4-C
640-1440/4-9	9 (69.2)	24 (49.0)	
1441-1920/9-12	0	2 (4.1)	
1920+/12+	0	3 (6.1)	
Length: wks/mos (N=87)	N=21	N=66	48.3-M
16/4	1 (4.8)	9 (13.6)	34.9-C
17-36/4-9	5 (23.8)	32 (48.5)	
37-48/9-12	3 (14.3)	17 (25.8)	
48+/12+	12 (57.1)	8 (12.1)	

Table 23

Educational Program Variables: Crosstabulations by Program Level

Variable Name	N	chi square	df	p
Status	114	19.45	1	.0001
Setting	114	10.04	1	.0015
Administration	103	21.26	2	.0001
Directors	101	13.65	2	.0011
Accreditation	113	18.93	1	.0001
Year established	90	2.83	2	.2431
Class size	76	10.97	2	.0041
Length: hours	62	2.39	3	.4962
weeks	87	18.33	3	.0004

Current status. Programs were also examined according to whether they were currently active or not (see Table 24). As with program level, there were differences in the characteristics of active and inactive programs.

Most active programs were in NLN accredited colleges or universities, whereas inactive programs were typically not accredited and were more diverse in sponsorship. In active programs, administrative control tended to be vested in nursing, although they were as likely to have joint co-directors as a nurse director. Inactive programs, on the other hand, were evenly divided in terms of administrative control and usually had joint co-directors. Active programs had larger class sizes than inactive programs, and were longer in hours and weeks.

Crosstabulations indicated that these differences were significant (see Table 25). Although the crosstabulation analysis for length in hours demonstrated significant differences, analysis of variance showed no significant differences between average hours for active and inactive programs (N=62; F=0.39; p=.5364).

Table 24

Educational Program Variables by Current Status of Program

Current Status: Variable Name:	Active (%)	Inactive (%)	Mean
Setting (N=114)	N=64	N=50	
college/university	49 (76.6)	21 (42.0)	
hospital	3 (4.7)	10 (20.0)	
community agency	3 (4.7)	7 (14.0)	
military facility	3 (4.7)	2 (4.0)	
joint sponsors	6 (9.3)	10 (20.0)	
Administration (N=103)	N=60	N=43	
nursing	37 (61.7)	13 (30.2)	
medicine	8 (13.3)	14 (32.6)	
joint	15 (25.0)	16 (37.2)	
Directors (N=101)	N=61	N=40	
joint	28 (45.9)	23 (57.5)	
nursing	31 (50.8)	11 (27.5)	
medicine	2 (3.3)	6 (15.0)	
Accreditation (N=113)	N=62	N=51	
NLN	44 (71.0)	17 (33.3)	
ANA	3 (4.8)	0	
NLN + ANA	2 (3.2)	0	
not accredited	13 (21.0)	34 (66.7)	
Year established (N=90)	N=58	N=32	
1966-1970	6 (10.3)	7 (21.9)	
1971-1975	38 (65.5)	21 (65.6)	
1976+	14 (24.1)	4 (12.5)	
Class size (N=76)	N=46	N=30	10.0-A
less than 8	9 (19.6)	13 (43.3)	8.1-I
8-10	20 (43.5)	12 (40.0)	
more than 10	17 (36.9)	5 (16.7)	
Length: hrs/mos (N=62)	N=43	N=19	887.1-A
less than 640/4	13 (30.3)	11 (57.9)	813.5-I
640-1440/4-9	28 (65.1)	5 (26.3)	
1441-1920/9-12	1 (2.3)	1 (5.3)	
1920+/12+	1 (2.3)	2 (10.5)	
Length: wks/mos (N=87)	N=54	N=33	40.3-A
16/4	3 (5.6)	7 (21.2)	34.5-I
17-36/4-9	26 (48.1)	11 (33.3)	
37-48/9-12	7 (13.0)	13 (39.4)	
48+/12+	18 (33.3)	2 (6.1)	

Table 25

Educational Program Variables: Crosstabulations by Current Status

Variable Name	N	chi square	df	p
Setting	114	8.40	1	.0037
Administration	103	8.35	2	.0154
Directors	101	1.16	1	.2815
Accreditation	113	19.15	1	.0001
Year established	90	3.29	2	.1926
Class size	76	6.18	2	.0455
Length: hours	62	8.52	3	.0344
weeks	87	18.28	3	.0004

Program Profile by Examinee

The program profile by examinee is described according to subsets (first-time takers and repeaters) and cohorts.

Comparisons between subsets. There were 2,903 formally prepared first-time takers and 134 formally prepared repeaters (who repeated the examination a total of 161 times). When the first-time takers and repeaters were compared, there were significant differences for 5 of 10 program variables (see Table 26). There were no differences between these subsets in regard to setting, accreditation status, length in weeks, class size, or year established (see Tables 26, 27, and 28).

First-time takers were more likely to graduate from active programs that were administered by nursing, with joint co-directors or a nurse director. These programs were shorter in number of total hours than those of repeaters, and there were fewer masters program graduates in the repeater subset.

Table 26
Educational Program Variables: Crosstabulations for First-Time Takers & Repeaters (N=3,604)

Variable Name	N	chi square	df	p
Status	3,064	14.01	1	.0002
Level	3,064	11.56	1	.0007
Setting[a]	3,064	0.01	1	.9436
Administration[a]	2,955	5.05	1	.0246
Directors[a]	2,940	6.90	1	.0086
Accreditation[a]	3,051	0.52	1	.4711
Year established	2,652	1.12	2	.5720
Class size	2,590	0.52	2	.7707
Length: hours	1,834	8.00	3	.0460
weeks	2,715	2.09	3	.5535

[a]Variables collapsed into 2 categories (see Table 21) due to small cell frequencies.

Table 27

Program Graduates: Frequencies & Categories for Program Variables of First-Time Takers (N=2,903)

Variable Name	Frequency	Percent	Mean
Status	N=2,903		
active	1,997	68.8	
inactive	906	31.2	
Level	N=2,903		
masters	344	11.8	
certificate	2,559	88.2	
Setting	N=2,903		
college/university	2,836	63.3	
hospital	229	7.9	
community agency	85	2.9	
military facility	201	6.9	
joint sponsors	552	19.0	
Administration	N=2,812		
nursing	1,200	42.7	
medicine	606	21.5	
joint	1,006	35.8	
Directors	N=2,791		
joint	1,469	52.6	
nursing	1,219	43.7	
medicine	103	3.7	
Accreditation	N=2,919		
NLN	1,584	54.2	
ANA	145	5.0	
NLN + ANA	230	7.9	
not accredited	960	32.9	
Year established	N=2,513		
1966-1970	555	22.1	
1971-1975	1,723	68.6	
1976+	235	9.3	
Class size	N=2,460		10.3
less than 8	342	13.9	
8-10	1,208	49.1	
more than 10	910	37.0	
Length: hrs/mos	N=1,745		782.0/
less than 640/4	663	38.0	4.9
640-1440/4-9	987	56.6	
1441-1920/9-12	55	3.1	
1920+/12+	40	2.3	
Length: wks/mos	N=2,569		36.1/
16/4	212	8.2	9.0
17-36/4-9	1,547	60.2	
37-48/9-12	382	14.9	
48+/12+	428	16.7	

Table 28

Program Graduates: Frequencies & Categories for Program Variables of Repeaters (N=134) & Repeats (N=161)

Variable Name	#Repeaters (%)	#Repeats (%)	Mean
Status	N=134	N=161	
active	71 (53.0)	88 (54.7)	
inactive	63 (47.0)	73 (45.3)	
Level	N=134	N=161	
masters	5 (3.7)	5 (3.1)	
certificate	129 (96.3)	156 (96.9)	
Setting	N=134	N=161	
college/university	74 (55.2)	88 (54.7)	
hospital	14 (10.5)	17 (10.6)	
community agency	5 (3.7)	6 (3.7)	
military facility	6 (4.5)	6 (3.7)	
joint sponsors	35 (26.1)	44 (27.3)	
Administration	N=118	N=143	
nursing	46 (39.0)	53 (37.1)	
medicine	17 (14.4)	23 (16.1)	
joint	55 (46.6)	67 (46.8)	
Directors	N=124	N=149	
joint	75 (60.5)	92 (61.7)	
nursing	45 (36.3)	52 (34.9)	
medicine	4 (3.2)	5 (3.4)	
Accreditation	N=106	N=132	
NLN	60 (46.6)	64 (48.5)	
ANA	11 (10.4)	15 (11.4)	
NLN + ANA	10 (9.4)	13 (9.8)	
not accredited	25 (23.6)	40 (30.3)	
Year established	N=117	N=139	
1966-1970	27 (23.1)	36 (25.9)	
1971-1975	78 (66.7)	91 (65.5)	
1976+	12 (10.2)	12 (8.6)	
Class size	N=110	N=130	10.6
less than 8	16 (14.5)	18 (13.8)	
8-10	51 (56.4)	60 (46.2)	
more than 10	43 (39.1)	52 (40.0)	
Length: hrs/mos	N= 71	N= 89	840.7/
less than 640/4	20 (28.2)	28 (31.5)	5.2
640-1440/4-9	46 (64.8)	53 (59.5)	
1441-1920/9-12	2 (2.8)	2 (2.3)	
1920+/12+	3 (4.2)	6 (6.7)	
Length: wks/mos	N=121	N=146	36.4/
16/4	6 (5.0)	8 (5.5)	9.1
17-36/4-9	75 (62.0)	91 (62.3)	
37-48/9-12	20 (16.5)	25 (17.1)	
48+/12+	20 (16.5)	22 (15.1)	

Comparisons among cohorts. As with the sociodemographic variables, cohort analysis was performed on the total population of formally prepared candidates (N=3,064). There were significant differences between groups by year of examination for 9 of 10 program variables (see Table 29). There were no significant differences according to administration of the programs.

Table 29
Educational Program Variables: Crosstabulations for Examinees by Year of Examination (N=3,064)

Variable Name	N	chi square	df	p
Status	3,064	143.99	5	.0001
Level	3,064	204.60	5	.0001
Setting	3,064	35.19	5	.0001
Administration	2,955	4.60	5	.4661
Directors	2,940	31.77	5	.0001
Accreditation	3,051	38.13	5	.0001
Year established	2,652	178.61	10	.0001
Class size	2,590	75.92	10	.0001
Length: hours	1,834	65.91	15	.0001
weeks	2,715	79.88	15	.0001

Table 30
Selected Program Variables: Means & F Values by Year of Examination

Year:	All	1977	1978	1979	1980	1981	1982	F
(N=)	(3,064)	(942)	(530)	(540)	(395)	(369)	(288)	
Variable:								
Class size	10.3	9.7	10.3	10.4	11.1	10.4	10.6	11.0*
Length: hours	784.8	737.7	774.2	761.6	785.4	789.2	859.6	8.4*
Length: weeks	36.1	34.2	35.5	36.8	36.4	37.6	39.2	8.7*

*$p < .05$

Over the 6 years in which the examination has been given, a steadily increasing proportion of the examinees were graduates of programs whose status was active. Most inactive program graduates (41.3%) took the exam in 1977. In regard to program level, there was a gradual increase

in the number of examinees who were graduates of masters level programs (4.8% in 1977; 28.5% in 1982). The largest proportion of certificate level graduates (33.0%) took the exam in 1977.

Incremental increases were also evident for the proportion of programs sponsored by colleges or universities (76.6% in 1977; 85.4% in 1982). Although the proportion of examinees whose programs were administered by nursing remained relatively constant (44-47%) between 1977-1982, those whose programs were administered by medicine have decreased and those with joint administration have increased. With the exception of 1977, when most graduates had a director who was either a nurse or a physician (56.5%), graduates typically had joint co-directors (51-56%).

The proportion of graduates of accredited programs has gradually increased (65.7% in 1977; 79.3% in 1982); graduates of unaccredited programs were most likely to take the exam in 1977 (37%). As might be expected, most graduates' programs were established between 1971-1975, and their proportion in each exam year was relatively constant (65.7-71.4%). Forty-two percent (42%) of graduates whose programs were established between 1966-1970 took the exam in 1977.

Although the average class size of programs appears to have changed very little (mean range=7.7-11.1), analysis of variance and Duncan's test indicated that there were significant differences in means between the 1977 group, the 1980 group, and the other 4 years (see Table 30).

Also over this 6 year period, the length of examinees' programs in both hours and weeks has increased. For length in hours, Duncan's test showed that there were significant differences in means between

those tested prior to 1981 and those tested in 1981-1982. The analysis for length in weeks indicated that the means for 1977 and for 1982 were significantly different than the means for 1978-1981.

Correlations between Educational Program Variables

Zero-order correlations were performed for program variables by program (N=114) and by examinee-graduate (N=3,064). The examinee-graduate correlations were subdivided by subset and cohort.

Educational programs. Correlations between program variables are presented in Table 31. Those variables not measured at the interval level were collapsed into dichotomous variables as previously described (see Table 21).

Those variables most consistently related to other program variables were length in weeks (usually positive, moderate magnitude) and educational level (usually positive, moderate magnitude). The significant relationships of the highest magnitude, however, were between programs' accreditation status and their settings and administration.

Table 31

Educational Program Variables: Zero-Order Correlations by Program (N=47-114)

Variable Name	Level (X2)	Setting (X3)	Administration (X4)	Directors (X5)	Accreditation (X6)	Length: Hours (X7)	Length: Weeks (X8)	Class size (X9)	Exam Year (X10)
Status (X1)	.41*	.27*	.29*	-.11	.42*	.08	.20	.26*	.24*
X2		.30*	.46*	-.27*	.42*	-.06	.41*	.37*	.11
X3			.43*	-.16	.75*	.04	.26*	.18	.13
X4				-.51*	.62*	.07	.33*	.26*	.23*
X5					-.16	-.18	-.25*	-.02	.01
X6						.14	.34*	.31*	.19
X7							.38*	-.17	.25*
X8								.23*	.19
X9									-.06

*p < .05

Examinee subsets. Correlations between program variables for first-time takers (N=2,903) and repeats (N=161) are given in Tables 32 and 33. For the first-time takers, there were significant relationships between most program variables. Those most consistently related to other variables were current status, level, accreditation, and length. These were usually positive relationships of low to high magnitude. The correlation with the highest magnitude (positive) was between accreditation status and program setting.

For the repeats, there were fewer significant relationships. The only relationship that was relatively consistent was between program status and other variables, although the highest magnitude correlation was a negative one between program administration and program directors. (The obvious relationship between hours and weeks has been excluded from discussion.)

Table 32
First-Time Takers: Zero-Order Correlations between Program Variables (N=1,517-2,903)

Variable Name	Level (X2)	Setting (X3)	Administration (X4)	Directors (X5)	Accreditation (X6)	Length: Hours (X7)	Length: Weeks (X8)	Class size (X9)	Exam Year (X10)
Status X1	.24*	.16*	.22*	.06*	.42*	.06*	.12*	.26*	.26*
X2		.17*	.27*	-.11*	.24*	.09*	.34*	.10*	.09*
X3			.27*	-.04*	.81*	.12*	.15*	.06*	-.03
X4				-.51*	.46*	.04	.21*	.28*	.29*
X5					-.07*	-.19*	-.15*	-.08*	.02
X6						.13*	.21*	.18*	.16*
X7							.52*	-.21*	.21*
X8								-.01	.22*
X9									.02

*p<.05

Table 33
Repeating Subset: Zero-Order Correlations between Program Variables (N=72-161)

Variable Name	Level (X2)	Setting (X3)	Administration (X4)	Directors (X5)	Accreditation (X6)	Length: Hours (X7)	Length: Weeks (X8)	Class size (X9)	Exam Year (X10)
Status (X1)	.16*	.16*	.26*	.10	.46*	.14	.13	.23*	.18*
X2		.08	.09	-.08	.13	-.07	.28*	.02	-.02
X3			.22*	-.08	.80*	.20	.06	.04	-.10
X4				-.50*	.39*	.14	.13	.09	.19*
X5					-.06	-.02	.04	.12	.14
X6						.17	.22*	.09	.05
X7							.51*	-.35*	.14
X8								.09	.14
X9									.03

*$p < .05$

Examinee cohorts. When educational program variables were evaluated by year of examination, some consistent relationships across years were found. (All relationships reported were significant at $p=.05$ or less.)

Those variables most consistently related to other program variables across years were status, administration, directors, and accreditation. Significant relationships of high magnitude were found between accreditation and status, setting, and directors (r range=.37 to .83). There were also consistent relationships between programs' administration and directors (r range=-.40 to -.70); and, obviously, between length in hours and weeks (r range=.48 to .63).

For 3 years there were other relationships of moderate to high magnitude; these correlations were of higher magnitude than those for the subsets (Table 32). These relationships were: in 1977, between status and class size (r=.42); in 1981, between educational level and

1982, there was a significant relationship between educational level and administration (r=.50).

Examination Performance

The performance or criterion measure used in this research was the examination scores of certification candidates.

Variable Definition

As described in Chapter IV, individual raw scores were converted to T scores, so that the mean standard score of formally prepared first-time takers was 500, with a standard deviation (SD) of ± 100. These were composite scores: data made available to the investigator included no breakdown between scores on multiple choice questions and patient management problems (1977-1979), between questions designed to measure different types of cognitive performance (for example, comprehension versus application), or between questions according to the blueprint used for test item development (tasks, age groups, areas of content).

For most analyses, examination scores are presented as continuous variables (rounded to the closest whole number). In some instances, however, scores were collapsed into 3 categories and assigned a numerical scale. Those categories and scales were: low (1)--scores 1 standard deviation below the mean (of the total sample); average (2)--scores within mean range; and high (3)--scores 1 standard deviation above the mean.

Profile by Examinee

Examination scores were analyzed for examinees according to subsets (first-time takers and repeaters) and cohorts (year of exam).

Subset comparisons. Of the 3,387 candidates tested, 3,206 were first-time takers and 181 were repeats. The mean examination score for the total sample was 489: 495 for first-time takers (SD= 101; range=35-760) and 369 for repeats (SD=88; range=130-640). Analysis of variance revealed that these differences in average score were significant (N=3,387; F=271.44; p=.0001).

Cohort comparisons. There were no significant differences in average score by year of examination. Table 34 presents the mean scores, standard deviations, range, number tested, and F values by year for the total population and for the subsets.

While there were no significant differences among the first-time taker subset or the repeating subset, there were significant differences between these subsets for each year.

Table 34
Mean Examination Scores by Year of Exam & NQE Status

Year	Mean Score	SD	Range	N	F
1977-total	501	99	35-730	943	
1978-total	487	106	60-760	587	
first-time takers	495	100	60-760	547	55.31*
repeats	472	105	260-640	40	
1979-total	487	105	70-740	620	
first-time takers	493	103	70-740	585	40.42*
repeats	380	84	180-525	35	
1980-total	482	106	175-720	481	
first-time takers	491	102	175-720	443	43.60*
repeats	378	86	200-560	38	
1981-total	479	107	155-760	464	
first-time takers	489	103	155-760	424	55.63*
repeats	365	77	170-525	20	
1982-total	485	108	130-720	292	
first-time takers	499	99	175-720	264	62.34*
repeats	345	86	130-510	28	

*p<.05

The average performance of the repeating subset was dramatically affected by the lower scores of multiple-repeaters (those persons who repeated the exam more than once). For those 22 candidates who repeated the exam twice (total of 3 tests), the score range was 200-480. These examinees (12/22=54.5%) tended to show inconsistent performance across examinations; a smaller proportion either improved their performance (5/22=22.7%) or their performance declined (5/22=22.7%) with repeats. For the 4 candidates who repeated the exam three times (total of 4 tests), the score range was 60-410. Unlike the other multiple-repeaters, however, this group improved their performance with each successive retake. Multiple-repeaters usually retook the examination the year following a failure (1978 for 1977 failure, and so on). It should be noted that only 23% (6/26) of this group eventually received an examination score that was within mean range (passing).

Sociodemographic comparisons. Crosstabulation analysis demonstrated significant differences in examination scores by sociodemographic variables. For these analyses, scores were collapsed into 3 categories as described earlier. There were no significant differences in scores for first-time takers according to examinee sex, PNP experience, or function (see Table 35).

For age group, there were proportionally more low scorers in the over 35 age groups and more high scorers in the 25-34 age group. Analysis of variance and Duncan's test indicated that there were also significant differences in mean scores by age group for both first-time takers and the repeating subset (see Table 36). For first-time takers, those in the 25-34 age group had higher mean scores (mean=514) than those in the 20-24 (mean=483) or 35-44 (mean=483) age groups. There

Table 35

Examination Scores:[a] Crosstabulations by Sociodemographic Variables

Variable Name	N[b]	chi square	df	p
Sex	3,204	0.08	1	.9600
Age group	3,193	148.98	8	.0001
RN experience group	3,033	96.82	8	.0001
PNP experience group	3,198	20.21	12	.1236
Highest education	3,194	234.70	14	.0001
PNP preparation	3,206	40.78	2	.0001
Current function	3,191	0.60	2	.7420
Employment setting	3,200	58.36	18	.0001

[a]Scores collapsed into 3 categories: (1) low, (2) average, (3) high

[b]First-time takers only; small cell frequencies for repeater subset produced data that may be invalid.

Table 36

Examination Scores:[a] Analysis of Variance for Sociodemographic Variables by NQE Status

Variable Name	N	df	F	p
Sex				
first-time taker	3,204	1	0.03	.8708
repeater	181	1	1.80	.1809
Age group				
first-time taker	3,193	4	53.56	.0001
repeater	181	4	7.97	.0001
RN experience group				
first-time taker	3,033	4	27.41	.0001
repeater	171	4	2.46	.0474
PNP experience group				
first-time taker	3,198	6	3.03	.0037
repeater	181	6	3.32	.0041
Highest education				
first-time taker	3,194	7	43.18	.0001
repeater	181	6	2.27	.0393
PNP preparation				
first-time taker	3,206	1	63.64	.0001
repeater	181	1	1.13	.2898
Current function				
first-time taker	3,191	1	0.26	.6092
repeater	181	1	0.48	.4898
Employment setting				
first-time taker	3,200	8	8.56	.0001
repeater	181	8	0.32	.9662

[a]Scores measured at the interval level, not categories

were also differences in mean scores for the 45-54 age group (mean= 459) and for the over 55 age group (mean=392).

For the repeater subset, Duncan's test showed that the mean scores for the 3 youngest age groups (mean range=380-462) were significantly different than those for the 2 older age groups (mean range= 312-328).

Crosstabulation analysis also indicated differences in scoring according to RN experience, with a larger proportion of those with more months of experience achieving lower scores. Analysis of variance also revealed differences in mean scores according to RN experience, for both first-time takers and repeaters. For first-time takers, Duncan's test demonstrated that the mean scores of examinees with 13-60 months (1-5 years) of experience (mean=511) and 61-120 months (5-10 years) of experience (mean=504) were different than those with less than 12 months (mean=484) and those with 121-180 months (10-15 years; mean= 483). They were also different than those with more than 181 months of experience (mean=453). For the repeaters, those with more than 181 months achieved significantly lower average scores (mean=341) than examinees with less experience (mean range=367-398).

For PNP experience, there were no significant differences on crosstabulation analysis; however, analysis of variance did show differences. For first-time takers, Duncan's test demonstrated that the mean scores of those with no experience to 60 months (5 years) of experience were alike (mean range=491-504), but the mean for those with more than 60 months experience was different (mean=477). In the repeater subset, both those with no experience and those with more than 60 months scored alike (means=294 and 340, respectively), as did those with between 1-60

months (mean range=365-408).

With regard to highest level of education, those with masters or higher degrees tended to have a larger proportion of high scorers while those with less than bachelors degrees were over-represented in the low scoring group. There were also significant differences in mean scores according to educational level. Analysis of variance and Duncan's test showed differences between those with doctorates in nursing (mean=441), associate degrees in nursing (mean=443), and non-nursing associate degrees (mean=432), compared to those with other types of educational preparation (means=536-MSN; 514-BSN; 502-MS; 476-BS; 475-non-nursing doctorate; 468-diploma). For the repeating subset, there was a significant difference in the mean score of the doctorally prepared repeaters (mean=170) versus all other educational groups (mean range=348-414).

There were also significant differences in scoring for first-time takers according to their type of nurse practitioner preparation. Formally prepared examinees achieved higher average scores (mean=500) than did the informally prepared examinees (mean=452). There were no differences for the repeating subset on analysis of variance.

For employment setting, crosstabulations revealed significant differences in scoring for the first-time takers, with faculty members and the unemployed over-represented in the high scoring group and those employed in school systems and hospital inpatient settings over-represented in the low scoring group. With analysis of variance, there were no significant differences in the average scores of the repeaters, but there were differences among first-time takers. Duncan's test indicated that those in faculty settings and the unemployed scored alike (means=535 and 527, respectively), those in school systems and inpatient set-

tings scored alike (means=466 and 450, respectively), and those in all other settings scored alike (mean range=483-503).

Educational program comparisons. Examination scores were also analyzed according to the characteristics of examinees' educational programs (merged file). For this analysis, there was no cohort analysis, but there was analysis by subsets.

Crosstabulations showed significant differences in scoring categories for 5 of 10 program variables (see Table 37). There were no significant differences for setting, program directors, year established, or length in hours or weeks.

There were differences by program status, with graduates of inactive programs over-represented in the low scoring group. Analysis of variance indicated no significant differences in mean score for the repeater subset; however, there were differences for the first-time takers (see Table 38). For first-time takers, graduates of active programs had higher mean scores (mean=508) than graduates of inactive programs (mean=483).

In regard to educational level, there were proportionally more masters program graduates in the high scoring group. Analysis of variance and Duncan's test showed that masters graduates had significantly higher mean scores (mean=541) than certificate graduates (mean=495), among first-time takers. There were no differences for repeaters.

There were also differences in scoring categories according to program administration: graduates of programs administered by nursing were over-represented in the high scoring group. Analysis of variance demonstrated that there were significant differences in mean scores for the first-time takers. Graduates of nursing-administered programs

Table 37
Examination Scores:[a] Crosstabulations by Program Variables

Variable Name	N[b]	chi square	df	p
Status	2,903	29.45	2	.0001
Level	2,903	63.02	2	.0001
Setting	2,903	5.34	2	.0692
Administration	2,812	9.81	2	.0074
Directors	2,791	4.40	2	.1107
Accreditation	2,919	8.59	2	.0136
Year established	2,513	6.73	4	.1510
Class size	2,460	13.19	4	.0104
Length: hours	1,745	2.27	6	.8937
weeks	2,569	8.92	6	.1784

[a]Scores collapsed into 3 categories: (1) low, (2) average, (3) high
[b]First-time takers only; small cell frequencies for the repeating subset produced data that may not be valid.

Table 38
Examination Scores:[a] Analysis of Variance for Program Variables by NQE Status

Variable Name	N	df	F	p
Status				
first-time takers	2,903	1	41.07	.0001
repeaters	160	1	2.60	.1087
Level				
first-time takers	2,903	1	68.20	.0001
repeaters	160	1	2.32	.1295
Setting				
first-time takers	2,903	1	1.37	.2421
repeaters	160	1	2.02	.1567
Administration				
first-time takers	2,812	1	6.98	.0083
repeaters	143	1	2.37	.1257
Directors				
first-time takers	2,791	1	5.85	.0156
repeaters	149	1	1.88	.1723
Accreditation				
first-time takers	2,919	1	9.08	.0026
repeaters	131	1	0.35	.5535
Year established				
first-time takers	2,513	2	1.94	.1441
repeaters	139	2	1.54	.2185
Class size				
first-time takers	2,460	2	2.52	.0810
repeaters	130	2	0.15	.8611
Length: hours				
first-time takers	1,745	3	0.91	.4384
repeaters	89	3	0.63	.6036
Length: weeks				
first-time takers	2,569	3	3.24	.0211
repeaters	146	3	4.18	.0074

[a]Scores measured at the interval level, not categories

had higher average scores (mean=506) than graduates of medicine or jointly-administered programs (mean=495).

Although crosstabulation revealed no significant differences in scoring categories by discipline of program directors, there were significant differences in mean scores, for first-time takers. Duncan's test showed that jointly directed program graduates achieved a mean score of 496, which was significantly lower than the mean score of graduates whose programs were directed by a nurse or by a physician (mean=506).

The accreditation status of programs from which examinees graduated made a significant difference in their scoring category, with NLN accredited program graduates over-represented in the high scoring group. Non-accredited program graduates and ANA accredited program graduates were over-represented in the low scoring group. Analysis of variance indicated differences in mean scores for first-time takers, according to whether their program was accredited (mean=507) or not (mean=494).

Crosstabulation analysis revealed that there were significant differences in scoring categories by class size. There were proportionally larger numbers of graduates in the low scoring groups whose class sizes were less than 8, and larger numbers of graduates in the high scoring group whose class sizes were greater than 10 students. There were, however, no significant differences in the mean scores of examinees according to class size (analysis of variance).

Finally, crosstabulations demonstrated no significant differences according to length of the program in weeks. Analysis of variance and Duncan's test, for both first-time takers and repeaters, did show differences in mean scores. For first-time takers, those examinees

whose programs were 16 weeks (4 months; mean=519) scored higher than those whose programs were over 16 weeks (mean range=498-507). For the repeating subset, those whose programs were 16 weeks (mean=292) and 37-48 weeks (9-12 months; mean=339) scored alike, and those whose programs were 17-36 weeks (4-9 months; mean=376) and more than 48 weeks (12 months; mean=401) scored alike.

Profile by Program

An analysis of variance was also performed for program variables and average program score (program file). This analysis indicated that there were no significant differences in mean scores for 9 of 10 program variables. Educational level of the program was the exception, with an average score of 521 for masters programs (N=24) and 489 for certificate programs (N=90); N=114, F=6.26, p=.0138.

Supplementary analysis. Over the past 18 years, educational programs have undergone a number of changes in relation to their level, administrative structure, and content. For that reason, the investigator was particularly interested in the performance of graduates of programs that, although sponsored by the same institution, had changed sufficiently to require a separate program review and code by the National Board.

To determine whether these changes over the years in same-sponsored programs made a difference in the average performance of examinee-graduates, an analysis of variance and Duncan's test was performed. There were 32 programs (28% of total number) sponsored by 12 different universities that were involved in this analysis. Of the 32, 10 were no longer operational (all certificate level), 6 had never been formally reviewed or approved by the National Board (4 masters; 2 certificate),

and 16 were currently active (7 certificate; 9 masters). For most of these programs, the program changes were related to progression from certificate to masters level.

With one exception, analysis of variance revealed no significant differences in mean scores. The exception was a university that had jointly sponsored its original certificate program with a hospital and a community agency. The graduates of the original program achieved significantly lower average scores (mean=364) than graduates of the 3 other programs that the university subsequently sponsored (1 certificate, 2 masters; mean range=493-527).

These results suggested to the investigator that perhaps differences between programs with different sponsors were more important than within program (same-sponsor) differences. That is, it is the characteristics of the individuals who are attracted to a particular program in the first place that make a difference in their performance as graduates.

Correlations between Variables

Correlational analyses were performed between examination scores and sociodemographic and program variables for examinees, by subset (see Tables 39 and 40). They were also done for sociodemographic variables controlling for three major program variables, and for program variables by program.

As Tables 39 and 40 indicate, there were fewer significant relationships between scores and other variables for the repeating subset than for first-time takers. For the repeating group, the only significant relationship of moderate or high magnitude was the negative correlation between score and age (r=-.40). For the first-time takers, all

significant correlations were or low magnitude. The highest magnitude correlation was the positive one between score and highest education (r=.28).

Table 39
Zero-Order Correlations between Examination Scores & Sociodemographic Variables by NQE Status

Variable Name	First-Time Takers (N=3,190-3,206)	Repeats (N=181)
Sex	.00	.10
Age	-.21*	-.40*
RN experience	-.14*	-.23*
PNP experience	-.07*	-.13*
Highest education	.28*	.08
PNP preparation	.14*	.08
Current function	.01	.05
Employment setting	.00	.04
Exam year	-.03	-.10

*p<.05

Table 40
Zero-Order Correlations between Examination Scores & Program Variables by NQE Status

Variable Name	First-Time Takers (N=2,745-2,903)	Repeats (N=89-161)
Status	.12*	.13
Level	.15*	.12
Setting	.02	-.11
Administration	.05*	-.13
Directors	-.05*	-.11
Accreditation	.06	-.05
Year established	.00	.09
Class size	-.01	-.05
Length: hours	-.02	-.09
weeks	.00	.04

*p<.05

Major program variables controlled. Correlations between scores and sociodemographic variables, with 3 major program variables controlled, produced the following results. With PNP preparation (formal versus informal) controlled, there were differences in the magnitude of correlations between scores and age, RN experience,

and PNP experience compared to the zero-order correlations (see Table 39). For first-time takers, the zero-order correlation was r=-.21 between score and age. When PNP preparation was controlled, it became obvious that the negative impact of age was more important for the informally prepared examinee than for the formally prepared examinee (see Table 41). This was also true for the repeating subset.

In relation to RN experience, for first-time takers there was no important effect with PNP preparation controlled. For the informally prepared repeaters, however, the negative relationship between score and RN experience was not significant. The effect on PNP experience, for first-time takers, was to eliminate the significance for formally prepared examinees and to increase the magnitude of the correlation for those who were informally prepared.

With program status (active versus inactive) controlled, the only important changes occurred in the repeating subset. The relationship between score and age was only significant for the repeaters whose programs were currently active. It was not a significant factor for repeaters whose programs were inactive. The same was true for the relationships between score and RN experience; that is, for the repeaters who were graduates of inactive programs there was no significance.

Finally, program level (masters versus certificate) was controlled, with some interesting results. For the first-time takers, there was essentially no relationship between age and score for the masters program graduates. The same was true for RN experience and highest education, for the masters prepared first-time takers. The previous correlations, therefore, were attributable to the certificate graduates.

Table 41

Correlations between Examination Scores & Sociodemographic Variables by Selected Program Variables & NQE Status

NQE Status:	First-Time Takers		Repeats	
Variable Name:				
PNP preparation:	Formal	Informal	Formal	Informal
(N=)	(2,762-2,903)	(283-303)	(160-161)	(19-20)
Sex	.01	.01	.11	.00
Age	-.19*	-.32*	-.37*	-.70*
RN experience	-.15*	-.15*	-.22*	-.31
PNP experience	-.02	-.16*	-.12	-.09
Highest education	.29	.26*	.09	-.13
Current function	-.01	.04	.05	.08
Employment setting	.00	-.01	.05	-.16
Program status:	Active	Inactive	Active	Inactive
(N=)	(1,912-1,996)	(850-906)	(82-88)	(69-72)
Sex	.02	-.01	.17	.00
Age	-.16*	-.18*	-.47*	-.21
RN experience	-.10*	-.19*	-.27*	-.16
PNP experience	-.01	.03	-.02	-.17
Highest education	.28*	.25*	.12	.02
Current function	.00	.01	.12	-.01
Employment setting	.00	.01	.09	.02
Program level:	Masters	Certificate	Masters	Certificate
(N=)	(328-344)	(2,434-2,558)	(4-5)	(147-155)
Sex	.06	.00	.00	.11
Age	.01	-.18*	-.94*	-.34*
RN experience	-.04	-.14*	-.46	-.20*
PNP experience	.07	-.01	.23	-.12
Highest education	.03	.27*	-.14	.06
Current function	.06	.00	.92*	.04
Employment setting	.04	-.01	.64	.03

*p .05

Because of the small number of masters program graduates in the repeating subset, no meaningful interpretations can be made.

Average score and program variables. A final correlational analysis was done between the average score for each educational program (N=114) and the 10 program variables. The only significant correlation was between average program score and the educational level of the program (r=.23), indicating that masters programs had higher average scores than certificate programs.

Summary of Descriptive Results

This chapter presented the descriptive analysis for the variables under consideration in this research. That analysis included examination of the sociodemographic and program variables by cross-tabulation, analysis of variance, and correlational analysis. It also involved evaluation of relationships between these variables and the dependent variable, examination scores.

In general, there were significant differences in sociodemographic and educational program characteristics according to examinees' NQE status and year of examination. There were also significant differences in examination performance related to examinees' sociodemographic characteristics and, to a lesser extent, program characteristics. Chapter VII provides the detailed multivariate analyses for these data.

Chapter VII. Multivariate Analysis

This chapter describes the detailed multivariate analyses for this research. Regression analyses were performed to determine the amount of variance in examination performance (scores) that could be explained by the sociodemographic and program variables under study.

Separate regression equations were constructed to determine: (a) the ability of sociodemographic variables to predict examination performance; (b) the ability of program variables to predict examination performance; and (c) the ability of the combined model to predict examination performance. When appropriate, these analyses were conducted by the NQE status of examinees (first-time takers and repeaters) and by cohorts (year of examination). For those variables not measured at the interval level, dummy variables were created as described in Chapter VI.

Exploratory Analyses

Using the SAS stepwise regression procedure (PROC STEPWISE), an initial exploratory analysis was done. This analysis was performed separately for sociodemographic and program variables by examinee subsets. From these analyses, the investigator made decisions about variables to exclude from further analysis.

In the stepwise procedure, a variable must be significant at $p = .1500$ (default) to enter the equation (SAS Institute, 1982, p. 104).

As a result of this procedure for sociodemographic variables, 5 variables were excluded from further analysis (RN experience, PNP experience, current function, employment setting, sex). Therefore, subsequent analyses considered the 5 remaining variables: age, highest education, NQE status, exam year, and type of PNP preparation.

As a result of the stepwise procedure for program variables, 5 variables were also excluded from further analysis (program status, year established, class size, length in hours, length in weeks). The remaining 5 variables (educational level, accreditation status, setting, administration, directors) were considered in later analyses.

Regression Models

After exploratory analysis using the SAS stepwise procedure, further analyses were performed using the SAS regression procedure (PROC REG). This procedure is a general-purpose one for regression that fits least-squares estimates to linear regression models.

For these analyses, regression equations were constructed separately for sociodemographic variables, for program variables, and for the combined model. These analyses were done by subsets of examinees, by cohorts of examinees, and by type of PNP preparation of examinees. In addition, an analysis was conducted in which the average examination score of each program was regressed on the program variables.

Subset analyses. Equations were constructed according to the NQE status of examinees (first-time taker versus repeater) that regressed examination performance on 3 sociodemographic variables (see Table 42). This 3-variable model explained only 8% of the variance in examination scores for first-time takers. Examinees' highest

education made the largest contribution to differences in their scores (for example, with each increase in educational category, examinees averaged a 24 point increase in score).

For the repeating subset, however, this 3-variable model explained 22% of the variance in scores. For this group, highest education was not significant in the overall model. The year of examination made the largest contribtion to differences in scoring for this subset; for example, examination scores averaged 10 points lower with each increase in exam year (see Table 42).

Table 42
Regression of Examination Scores on Selected Sociodemographic Variables by NQE Status of Examinees

Variable Name	b value	SE	t	p
First-Time Takers:				
Intercept	151.43	104.32	1.45	.1468
Highest education	23.81	2.12	11.21	.0001
Age	-2.20	0.24	-9.24	.0001
Exam year	4.37	1.25	3.48	.0005
R^2=.0802 R=.2832	F=75.91	p=.0001	df=3, 2,612	
Repeaters:				
Intercept	1377.32	414.93	3.32	.0012
Highest education	5.90	8.21	0.72	.4737
Age	-4.57	0.80	-5.71	.0001
Exam year	-10.48	5.14	-2.04	.0434
R^2=.2218 R=.4710	F=12.26	p=.0001	df=3, 129	

To assess incremental changes in the explanation of variance in scores, marginal analyses were performed in which sociodemographic variables were entered into equations one at a time.

For first-time takers (see Table 43), highest education of examinees accounted for 4% of the variance in scores (Equation 1). When age was controlled (Equation 2), there was no change in the unstandardized regression coefficient (b value) for highest education, but

an additional 3% of variance in scores was explained. By controlling for year of examination (Equation 3), there was a slight increase in the b value for highest education, no change in the b value for age, and very little change in the R^2 value.

Table 43
Regression of Examination Scores on Selected Sociodemographic Variables: Incremental Changes, First-Time Takers

Variable Name	b value	SE	t	p
Equation 1:				
Intercept	433.99	6.82	63.66	.0001
Highest education	20.05	1.89	10.59	.0001
R^2=.0411 R=.2027	F=112.17	p=.0001	df=1, 2,614	
Equation 2:				
Intercept	513.06	10.41	49.28	.0001
Highest education	20.20	1.86	10.87	.0001
Age	-2.34	0.24	-9.92	.0001
R^2=.0759 R=.2755	F=107.34	p=.0001	df=2, 2,613	R^2 change=.0348
Equation 3:				
Intercept	151.43	104.32	1.45	.1468
Highest education	23.81	2.12	11.21	.0001
Age	-2.20	0.24	-9.24	.0001
Exam year	4.37	1.25	3.48	.0005
R^2=.0802 R=.2832	F=75.91	p=.0001	df=3, 2,612	R^2 change=.0043

For the repeating subset (see Table 44), highest education of examinees accounted for less than 1% of the variance in examination scores (Equation 1). As Equation 2 demonstrates, 19% of the variance in scores was explained by examinees' age. Controlling for exam year (Equation 3) added another 2% to the explanation of variance, and reduced the b value of highest education.

To summarize, when examination performance was regressed on selected sociodemographic variables by NQE status, the small amount of variance that could be explained for first-time takers was primarily due to their level of education and secondarily to their age. On the other hand, education made almost no contribution to the explanation

of variance for the repeating subset. The major factor for that group was their age.

Table 44

Regression of Examination Scores on Selected Sociodemographic Variables: Incremental Changes, Repeating Subset

Variable Name	b value	SE	t	p
Equation 1:				
Intercept	355.99	24.59	14.48	.0001
Highest education	7.08	9.21	0.77	.4432
R^2=.0045 R=.0671	F=0.59	p=.4432	df=1, 131	
Equation 2:				
Intercept	534.46	38.93	13.73	.0001
Highest education	6.20	8.31	0.75	.4571
Age	-4.51	0.81	-5.58	.0001
R^2=.1967 R=.4435	F=15.91	p=.0001	df=2, 130	R^2 change=.1922
Equation 3:				
Intercept	1377.32	414.93	3.32	.0012
Highest education	5.90	8.21	0.72	.4737
Age	-4.57	0.80	-5.71	.0001
Exam year	-10.48	5.14	-2.04	.0434
R^2=.2218 R=.4710	F=12.25	p=.0001	df=3, 129	R^2 change=.0251

After this examination, equations were constructed that regressed examination performance on 5 program variables, by NQE status of examinees (see Table 45). This 5-variable model explained only 3% of the variance in examination scores for first-time takers. The educational level of examinees' programs and their accreditation status made the largest contributions to explanation of variance in scores. For example, graduates of masters programs averaged 43 points higher than graduates of certificate level programs; also, graduates of accredited programs averaged 32 points higher than graduates of unaccredited programs. For the other 3 variables, there were inverse relationships between examination performance and graduates of programs in university/college settings, of nurse-administered programs, and of programs with joint co-directors.

For the repeating subset, twice as much variance in examination performance was explained by this 5-variable model (see Table 45). Program setting, however, was the only variable that was significant in the overall model. It also made the largest contribution to differences in scores.

Table 45
Regression of Examination Scores on Selected Program Variables by NQE Status of Examinees

Variable Name	b value	SE	t	p
First-Time Takers:				
Intercept	512.00	5.77	88.73	.0001
Educational level	43.00	5.98	7.19	.0001
Accreditation status	32.41	6.41	5.06	.0001
Program setting	-28.76	7.15	-4.02	.0001
Directors	-11.97	4.61	-2.60	.0094
Administration	-15.12	5.48	-2.76	.0058
R^2=.0334 R=.1828	F=18.05	p=.0001	df=5, 2,615	
Repeaters:				
Intercept	408.22	23.21	17.59	.0001
Educational level	51.98	40.68	1.28	.2036
Accreditation status	28.00	21.51	1.30	.1953
Program setting	-53.03	25.12	-2.11	.0367
Directors	-18.18	18.92	-0.96	.3384
Administration	6.82	21.35	0.32	.7500
R^2=.0612 R=.2474	F=1.65	p=.1493	df=5, 127	

Again, to assess incremental changes in the explanation of variance in scores, marginal analyses were performed in which program variables were entered into equations one at a time.

As Table 46 indicates, the educational level of examinees' programs accounted for 2% of the variance in scores (Equation 1). Although controlling for other program variables (Equations 2-5) contributed only 1% to the explanation of variance for first-time takers, there were substantial changes in the b values (particularly accreditation status) as additional variables entered the equations.

For the repeating subset (see Table 47), about 3% of the variance in scores was explained by the setting of examinees' programs. Like the first-time takers, those graduates in university or college sponsored programs scored lower than those whose programs were in other settings. Controlling for variables incrementally (Equations 1-5) made little difference in the R^2 value, and the overall model was not signficant at $p=.05$ or less.

To summarize the subset analyses for program variables, program variables were not good predictors of examination performance for either first-time takers or repeaters. Those variables, however, that made the largest contributions to differences in scores were not the same between groups. For the first-time takers, educational level and accreditation status were the most important predictors. The only significant predictor for the repeaters was the program setting.

A final subset analysis was performed on the combined, 8-variable model for first-time takers and repeaters. The combined model explained 9% of the variance in scores for first-time takers (see Table 48), with exam year, administration, and directors not significant. Examinees' highest education and the educational level of their programs made the largest contributions to explanation of variance in their scores. The R^2 value for the combined model was not substantially different than the R^2 for the sociodemographic variables alone ($R^2=.0822$, 3-variable sociodemographic model; $R^2=.0334$, 5-variable program model).

Table 46

Regression of Examination Scores on Selected Program Variables: Incremental Changes, First-Time Takers

Variable Name	b value	SE	t	p
Equation 1:				
Intercept	497.57	2.06	242.03	.0001
Educational level	44.37	5.70	7.78	.0001
R^2=.0226 R=.1503	F=60.54	p=.0001	df=1, 2,614	
Equation 2:				
Intercept	491.07	3.56	137.77	.0001
Educational level	41.14	5.88	7.00	.0001
Accreditation status	9.73	4.36	2.23	.0258
R^2=.0245 R=.1565	F=32.81	p=.0001	df=2, 2,613	R^2 change=.0019
Equation 3:				
Intercept	504.16	4.92	102.38	.0001
Educational level	41.14	5.86	7.01	.0001
Accreditation status	23.98	5.72	4.19	.0001
Program setting	-27.34	7.12	-3.84	.0001
R^2=.0300 R=.1732	F=26.90	p=.0001	df=3, 2,612	R^2 change=.0055
Equation 4:				
Intercept	507.49	5.54	91.60	.0001
Educational level	40.25	5.90	6.82	.0001
Accreditation status	24.43	5.73	4.27	.0001
Program setting	-28.35	7.16	-3.96	.0001
Directors	-5.08	3.88	-1.31	.1905
R^2=.0306 R=.1749	F=20.61	p=.0001	df=4, 2,611	R^2 change=.0006
Equation 5:				
Intercept	512.00	5.77	88.73	.0001
Educational level	43.00	5.98	7.19	.0001
Accreditation status	32.41	6.41	5.06	.0001
Program setting	-28.76	7.15	-4.02	.0001
Directors	-11.97	4.61	-2.60	.0094
Administration	-15.12	5.48	-2.76	.0058
R^2=.0334 R=.1828	F=18.05	p=.0001	df=5, 2,615	R^2 change=.0028

Table 47

Regression of Examination Scores on Selected Program Variables: Incremental Changes, Repeating Subset

Variable Name	b value	SE	t	p
Equation 1:				
Intercept	371.68	7.81	74.60	.0001
Educational level	60.32	40.27	1.50	.1366
R^2=.0168 R=.1296	F=2.24	p=.1366	df=1, 131	
Equation 2:				
Intercept	376.30	12.53	29.31	.0001
Educational level	57.51	40.88	1.41	.1619
Accreditation status	7.19	16.05	0.45	.6551
R^2=.0183 R=.1353	F=1.21	p=.3001	df=2, 130	R^2 change=.0015
Equation 3:				
Intercept	395.00	18.68	21.14	.0001
Educational level	57.51	40.43	1.42	.1573
Accreditation status	28.95	19.31	1.50	.1362
Program setting	-49.46	24.97	-1.98	.0497
R^2=.0473 R=.2175	F=2.14	p=.0973	df=3, 129	R^2 change=.0290
Equation 4:				
Intercept	410.54	21.97	18.69	.0001
Educational level	51.81	40.53	1.28	.2035
Accreditation status	31.00	19.31	1.60	.1110
Program setting	-52.79	25.02	-2.11	.0368
Directors	-21.37	16.02	-1.33	.1845
R^2=.0604 R=.2458	F=2.06	p=.0903	df=4, 128	R^2 change=.0131
Equation 5:				
Intercept	408.22	23.21	17.59	.0001
Educational level	51.98	40.68	1.28	.2036
Accreditation status	28.00	21.51	1.30	.1953
Program setting	-53.03	25.12	-2.11	.0367
Directors	-18.18	18.92	-0.96	.3384
Administration	6.82	21.35	0.32	.7500
R^2=.0612 R=.2474	F=1.65	p=.1493	df=5, 127	R^2 change=.0008

On the other hand, the combined, 8-variable model for the repeating subset explained 26% of the variance in examination scores. In this model, 6 of 8 variables were not significant at the $p=.05$ level. Examinees' age and program setting were the only significant individual variables in the model. This combined model explained 4% more variance than the 3-variable sociodemographic model for repeaters ($R^2=.2250$, sociodemographic model; $R^2=.0612$, 5-variable program model). (See Table 49.)

To review these results by NQE status of examinees, the variables that were important predictors of examination performance were different for first-time takers and repeaters. When these differences were examined in view of the findings presented in Chapter VI, possible explanations emerged.

In relation to the differential importance of age, it was a larger factor for repeaters because they averaged 6 years older than first-time takers. In addition, there were proportionally more examinees in the repeater subset in the over 45 age groups (33.7% versus 13.8% for first-time takers). Those examinees in the over 45 age groups obtained significantly lower scores than younger groups, both for first-time takers and repeaters, on analysis of variance (average point spread 64 points).

Highest education was not a significant factor for the repeating subset. On analysis of variance, there were no differences in mean scores of repeaters according to education, so this finding was not surprising. (The doctorally prepared repeater was excluded here.) There were, however, differences in mean scores of the first-time takers according to education, which was reflected in the regression

Table 48

Regression of Examination Scores on Selected Sociodemographic & Program Variables: First-Time Takers

Variable Name	b value	SE	t	p
Intercept	333.28	113.27	2.94	.0033
Sociodemographic:				
Highest education	20.20	2.29	8.81	.0001
Age	-1.97	0.24	-8.11	.0001
Exam year	2.16	1.37	1.58	.1142
Program:				
Educational level	21.44	6.46	3.32	.0009
Accreditation status	18.92	6.39	2.96	.0031
Program setting	-16.26	7.05	-2.30	.0213
Directors	-4.55	4.52	-1.01	.3140
Administration	-7.14	5.39	-1.32	.1852
R^2=.0875 R=.2958	F=31.26	p=.0001	df=8, 2,607	

Table 49

Regression of Examination Scores on Selected Sociodemographic & Program Variables: Repeating Subset

Variable Name	b value	SE	t	p
Intercept	1257.16	423.42	2.97	.0036
Sociodemographic:				
Highest education	6.21	8.77	0.71	.4806
Age	-4.49	0.82	-5.47	.0001
Exam year	-8.77	5.24	-1.67	.0968
Program:				
Educational level	6.88	39.28	0.17	.8612
Accreditation status	26.56	19.77	1.34	.1815
Program setting	-44.82	22.70	-1.97	.0505
Directors	-8.08	17.14	-0.47	.6413
Administration	13.99	19.51	0.69	.4907
R^2=.2560 R=.5066	F=5.35	p=.0001	df=8, 124	

model. The average point spread between those with associate degrees and those with all other types of preparation (doctoral excluded) was 57 points.

Differences in the importance of the educational level of examinees' programs to explanation of variance were also related to previous results. On analysis of variance, there were no significant differences in average scores among repeaters according to program level. This was not true for first-time takers, where there were significant differences, with masters graduates obtaining higher scores (point spread 46 points).

The differential importance of program setting was not as easily explained. Analysis of variance revealed no significant differences in mean scores according to setting, for either first-time takers or repeaters. In those analyses, setting was coded as college/university/joint (1) or other (0)--military facility, hospital, community agency. When analysis of variance and Duncan's test was performed for each setting individually, differences were apparent.

For first-time takers, those examinees whose programs were sponsored by universities or colleges and military facilities scored alike (means=506 and 519, respectively) and those from programs sponsored by community agencies (mean=480), hospitals (mean=480), and joint sponsors (mean=484) scored alike. For the repeaters, those from military programs (mean=452) had significantly higher scores than those from all other programs (mean range=352-389). The investigator concluded, therefore, that regression results related to program setting were due to the effects of the scores of military program graduates.

Cohort analyses. For these analyses, regression equations were constructed separately for sociodemographic variables, programs variables, and the combined model, by year of examination. Table 50 shows the results for the sociodemographic and program variable equations.

When examination performance was regressed on the sociodemographic variables by year of examination, only highest education and NQE status were significant (at p=.05) across all years. Examinee age was significant in 3 of 6 years (1979, 1980, 1981). With the exception of highest education, where b values have steadily decreased over the years, there were no clear trends across years. There was consistency across years in the direction of relationships between scores and the sociodemographic variables.

Regression of examination performance on these variables by year of examination clearly explained more variance than was true in the analysis by NQE status. For 1977, 14% of the variance in scores was explained (R=.3720; F=68.77; p=.0001; df=2, 856). That amount increased to 15% for 1978 (R=.3903; F=19.52; p=.0001; df=3, 326), and to 26% for 1979 (R=.5079; F=60.39; p=.0001; df=3, 521). For the remaining 3 years, the amount of variance explained was: 1980, 19% (R=.4358; F=30.16; p=.0001; df=3, 386); 1981, 17% (R=.4119; F=24.31; p=.0001; df=3, 357); and 1982, 22% (R=.4699; F=26.44; p=.0001; df=3, 280).

For the program variables, there was also an increase in the amount of variance explained when performance was examined by year of exam (see Table 50). The only program variable that was significant across all years was the educational level of examinees' programs. Accreditation status was significant in 4 of 6 years (not 1977 or 1978), and

Table 50

Regression of Examination Scores on Selected Sociodemographic & Program Variables by Year of Examination

Exam Year:	1977	1978	1979	1980	1981	1982
Sociodemographic Variables:	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)
Intercept	352.94	440.06	556.78	563.34	549.90	527.95
	(35.58)	(35.41)	(22.37)	(27.25)	(33.12)	(42.26)
Highest education	52.53	46.41	36.69	28.00	23.20	23.31
	(6.50)	(8.76)	(4.38)	(5.40)	(5.94)	(6.98)
Age	-2.33	0.29	-3.16	-3.17	-2.10	-1.16
	(0.40)	(0.64)	(0.46)	(0.61)	(0.73)	(0.88)
NQE status[a]	----	60.13	47.85	39.72	51.92	65.61
		(12.29)	(7.96)	(9.35)	(9.96)	(10.16)
R^2=	.1384	.1523	.2580	.1899	.1697	.2208
Program Variables:						
Intercept	505.66	520.14	523.59	515.35	478.80	473.65
	(9.39)	(16.47)	(13.61)	(16.62)	(18.17)	(19.19)
Educational level	65.92	54.26	57.68	81.00	31.46	37.65
	(15.31)	(22.44)	(15.32)	(14.54)	(14.00)	(15.67)
Accreditation status	1.54	-8.92	51.49	33.62	90.65	119.54
	(11.38)	(19.06)	(13.18)	(16.08)	(19.76)	(24.62)
Program setting	-17.71	-11.43	-51.55	-36.93	-54.34	-73.40
	(11.35)	(21.56)	(15.85)	(20.08)	(21.77)	(27.71)
Directors	7.99	-14.35	-30.13	-13.58	-7.97	-32.35
	(8.58)	(13.82)	(9.80)	(11.16)	(14.68)	(15.76)
Administration	6.84	18.75	-19.85	-41.08	-26.78	-30.12
	(10.62)	(16.22)	(11.42)	(12.78)	(17.25)	(18.63)
R^2=	.0253	.0443	.0754	.0951	.0860	.1322

[a]All first-time takers in 1977, therefore, NQE status not applicable

program setting was significant in 3 years (1979, 1981, 1982). Program directors made a significant contribution to the overall model in 1979 and 1982, but administration of the program was only a significant factor in 1980.

There were no clear trends across years for the program variables, and, with the exception of educational level and program setting, there was not even consistency in the direction of relationships. There was, however, a trend across years in the amount of variance explained by these variables, with an 11% increase between 1977 and 1982. In 1977, only 2% of the variance in scores was explained (R=.1591; F=4.43; p=.0006; df=5, 853); 4% was explained in 1978 (R=.2105; F=3.01; p=.0115; df=5, 324); and the amount increased to 7% in 1979 (R=.2756; F=8.46; p=.0001; df=5, 519). The trend continued in 1980, when 9% of the variance was explained (R=.3084; F=8.07; p=.0001; df=5, 384). Likewise, 9% of the variance was explained in 1981 (R=.2933; F=6.68; p=.0001; df=5, 355). Finally, the amount increased to 13% in 1982 (R=.3636; F=8.47; p=.0001; df=5, 278).

When examination performance was regressed on the combined 8-variable model, examinees' highest education and their NQE status were the only variables consistently significant across years. Age and accreditation status of examinees' programs were significant in 4 of 6 years, educational level of the program was significant in 3 of 6 years, and program setting was significant in 1 of 6 years. The program administration and directors were not significant in any year in the combined model.

As Table 51 indicates, the combined model by year explained more variance in examinees' scores than the analysis by NQE status (subset

Table 51
Regression of Examination Scores on Combined Sociodemographic & Program Variables by Year of Examination

Exam Year:	1977	1978	1979	1980	1981	1982
Variable Name:	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)	b (SE)
Intercept	356.97 (36.49)	449.52 (37.36)	576.91 (24.46)	577.33 (29.51)	526.33 (36.91)	526.26 (43.25)
Sociodemographic:						
Highest education	51.31 (6.65)	43.94 (9.31)	32.95 (4.67)	20.89 (6.23)	22.15 (6.78)	16.44 (8.15)
Age	-2.37 (0.40)	0.20 (0.66)	-2.70 (0.48)	-2.75 (0.64)	-1.71 (0.75)	-0.78 (0.89)
NQE status[a]	----	58.98 (12.26)	50.32 (7.97)	38.81 (9.40)	49.35 (9.93)	58.63 (10.22)
Program:						
Educational level	42.12 (14.50)	21.10 (22.18)	14.00 (14.48)	36.58 (16.03)	-6.31 (15.24)	9.51 (16.87)
Accreditation status	-23.06 (10.85)	-7.27 (18.26)	35.58 (12.05)	7.95 (15.55)	63.68 (19.12)	78.58 (23.71)
Program setting	2.88 (10.76)	-12.69 (20.28)	-40.80 (14.34)	-9.31 (19.40)	-33.38 (20.73)	-47.88 (26.16)
Directors	12.89 (8.03)	-20.25 (12.94)	-15.61 (8.88)	-2.26 (10.63)	-0.70 (13.97)	-21.99 (14.81)
Administration	16.59 (9.97)	12.97 (15.30)	-9.83 (10.26)	-21.83 (12.43)	-15.07 (16.44)	-20.59 (17.38)
R^2=	.1515	.1774	.2768	.2043	.2011	.2573

[a]All first-time takers in 1977, therefore, NQE status not applicable

comparisons). The only clear trend across years, as previously mentioned, was the decreasing influence of highest education on scores. There was the suggestion of a trend, however, for examinee age--decreasing influence across years.

For 1977, 15% of the variance in examinee scores was explained by this model (R=.3892; F=21.71; p=.0001; df=7, 851). That amount increased to 18% in 1978 (R=.4212; F=8.65; p=.0001; df=8, 321), and to 28% in 1979 (R=.5261; F=24.68; p=.0001; df=8, 516). In 1980, it decreased to 20% (R=.4520; F=12.23; p=.0001; df=8, 381), remained the same in 1981 (R=.4484; F=11.07; p=.0001; df=8, 352), and increased again in 1982 to 26% (R=.5072; F=11.91; p=.0001; df=8, 275).

In summary, regression of examination performance on sociodemographic and program variables by year of examination produced interesting results. As was true for the subset analysis, the most important preditors of performance across years were highest education, age, educational level of the program, and accreditation status. For this analysis, NQE status was added to the equations and, as could be expected, was a significant predictor across years.

The relative contributions (b values) of highest education, age, and educational level have decreased over this 6 year period as the examinee population's average age decreased, highest education increased, and the number of masters program graduates increased. While the direction of the relationships between these variables and scores has remained more or less consistent, there were 2 years in which this was not true for accreditation status. In 1977 and 1978 the relationship was negative, indicating that graduates of unaccredited programs obtained higher average scores than graduates of accredited programs. The reason for this is unclear, although a larger proportion of the population in those years were graduates of unaccredited programs (for example, 37% of examinees in 1977).

Analysis by PNP preparation. In these analyses, regression equations were constructed that regressed examination performance on the sociodemographic variables by examinees' type of PNP preparation (formal or informal). For the formally prepared examinees, this 4-variable model explained 15% of the variance in scores. NQE status and highest education of examinees made the largest contributions to differences in scores (see Table 52).

For the informally prepared examinees, this 4-variable model

explained 22% of the variance in scores; NQE status and exam year were not significant in the overall model. Highest education made the largest contribution to differences in scores for this group (see Table 52).

Table 52
Regression of Examination Scores on Selected Sociodemographic Variables by Type of PNP Preparation of Examinees

Variable Name	b value	SE	t	p
Formally prepared:				
Intercept	250.83	99.62	2.52	.0119
Highest education	23.31	2.03	11.46	.0001
Age	-2.42	0.23	-10.70	.0001
Exam year	3.83	1.20	3.19	.0014
NQE status	49.43	4.25	11.62	.0001
R^2=.1508 R=.3883	F=126.82	p=.0001	df=4, 2,856	
Informally prepared:				
Intercept	1257.47	425.39	2.96	.0034
Highest education	28.13	5.99	4.69	.0001
Age	-4.48	0.66	-6.80	.0001
Exam year	-8.89	5.30	-1.68	.0944
NQE status	15.43	11.98	1.29	.1987
R^2=.2252 R=.4745	F=21.65	p=.0001	df=4, 298	

In regard to the relative importance of these variables to the explanation of variance, there were differences between the groups in relation to age, exam year, and NQE status. The age variable was a larger contributor for the informally prepared examinees because they averaged 2 years older than the formally prepared examinees. This factor was also related to the relationship with year of examination.

In addition, the proportion of informally prepared candidates in the examinee population increased steadily between 1978 and 1981 (in 1981, 20.5% of the total population was informally prepared). In terms of NQE status, while first-time takers achieved higher average scores in both groups, the magnitude of the contribution was greater for the

formally prepared candidates. This was probably due to the fact that the average point spread in scores between formally and informally prepared first-time takers was 49 points, which is statistically significant (analysis of variance).

Program level analysis. A final analysis was performed in which examination performance was regressed on selected program variables, by educational program (N=114) and by examinee (all formally prepared). Table 53 presents the results of these equations.

Table 53
Regression of Examination Scores on Selected Program Variables by Examinee & by Educational Program

Variable Name	b value	SE	t	p
Examinees:				
Intercept	506.47	5.62	90.03	.0001
Educational level	48.24	6.04	7.98	.0001
Accreditation status	35.45	6.12	5.79	.0001
Program setting	-32.10	6.81	-4.71	.0001
Directors	-15.94	4.50	-3.55	.0004
Administration	-15.37	5.37	-2.86	.0043
R^2=.0391 R=.1977	F=23.90	p=.0001	df=5, 2,934	
Programs:				
Intercept	506.47	10.18	49.74	.0001
Educational level	48.24	10.94	4.41	.0001
Accreditation status	35.45	11.08	3.20	.0019
Program setting	-32.10	12.37	-2.60	.0108
Directors	-15.94	8.14	-1.96	.0530
Administration	-15.37	9.73	-1.58	.1174
R^2=.2775 R=.5268	F=7.30	p=.0001	df=5, 95	

In the individual-level analysis (by examinee), this 5-variable model explained 4% of the variance in examination scores. Examinees' educational level and the accreditation status of their programs made the largest contributions to differences.

The regression equation for the aggregate-level analysis (by program) was weighted according to the number of examinee-graduates from

each program. As Table 53 indicates, the 5-variable model explained 28% of the variance in average performance from program to program. The unstandardized regression coefficients (b values) for the two equations are mathematically identical, but the standard errors (SE) are smaller when individual scores are considered.

The results of the two regression equations reported in Table 53 can be combined to test the adequacy of the 5-variable linear model. The test involves a partitioning of the residual sum of squares from the individual level regression into two parts, one of which equals the residual sum of squares from the program level regression. The analysis is referred to as an "F-test for lack of fit" by Weisberg (1980, pp. 83-87), and is summarized in Table 54 (also see Iversen & Norpoth, 1976, p. 91).

The result (F=3.55) is statistically significant at the p=.01 level, indicating that a nonlinear function might provide a better fit or that the assumption of homoscedastic error variances is violated to some degree. It is consistent with the earlier finding that individual level variables make a difference when they are added to the regression equation.

Table 54
Analysis of Variance for Lack of Fit of Regression on 5 Program Variables

Source	Sum of Squares (SS)	df	Mean Square (MS)	F
Regression[a]	1,217,034	5	243,407	
Lack of fit	3,169,224	95	33,360	3.55*
Pure error	26,707,090	2,839	9,407	
Total	31,093,348	2,939		

[a]Model includes 5 variables listed in Table 53
*p<.01

Summary of Multivariate Results

This chapter presented the detailed multivariate analyses conducted for this research. Regression equations were constructed to determine the ability of sociodemographic, program, and combined models to predict examination performance. These analyses were performed by subsets (NQE status of examinees), by cohorts (year of examination of examinees), and by type of PNP preparation (formal or informal) of examinees. A final analysis regressed examination scores on selected program variables, comparing aggregate (by program) and individual (by examinee) level results.

Table 55
Summary of Multivariate Analyses: Regression of Examination Scores on Selected Sociodemographic & Program Variables

Variables:	Sociodemographic	Program	Combined
R^2 by NQE Status:	(3 variables)	(5 variables)	(8 variables)
First-time takers	.0802	.0334	.0875
Repeaters	.2218	.0612[a]	.2560
R^2 by Exam Year:	(3 variables)	(5 variables)	(8 variables)
1977	.1384	.0253	.1515
1978	.1523	.0443	.1774
1979	.2580	.0754	.2768
1980	.1899	.0951	.2043
1981	.1697	.0860	.2011
1982	.2208	.1322	.2573
R^2 by PNP Preparation:	(4 variables)		
Formal	.1508		
Informal	.2252		

[a]5-variable model not significant at $p=.05$

Table 55 summarizes the results of the regression analyses that were done. As the table indicates, the program variables made a limited contribution to the explanation of variance in examination scores. In the analyses by NQE status and by exam year, the 3-variable sociodemographic models explained essentially the same amount of variance

as the 8-variable combined models. Those variables that made the largest contributions to differences in scores were consistently highest education, age, and NQE status of examinees, and the educational level and accreditation status of their programs.

The final chapter (VIII) presents a summary and conclusions of this research, with discussion of limitations, implications, and recommendations for further research.

Chapter VIII. Summary and Conclusions

This research involved an analysis of data for 3,387 candidates who took the National Qualifying Examination for pediatric nurse practitioners/associates between 1977 and 1982. Those data were available from the National Board of Pediatric Nurse Practitioners and Associates, which administers the examination, and its testing agency, the National Board of Medical Examiners. Included in the data were sociodemographic characteristics of examinees, characteristics of their nurse practitioner educational programs, and their examination scores. The investigator sought to determine whether those sociodemographic and educational program variables were related to examination performance.

Summary Comments

Sample profiles. The sociodemographic profile for the examinee population was similar to the profile of pediatric nurse practitioners surveyed by Sultz and others in 1980 (for the Longitudinal Study of Nurse Practitioners). This sample (N=3,206) and the Sultz sample (N=199 pediatric) were comparable in terms of age, education, experience, and current function (Sultz, Bullough et al., 1983; Sultz, Henry, Bullough et al., 1983; Sultz, Henry, Kinyon et al., 1983a, 1983b).

For the educational program characteristics, there were similarities and contrasts with the Sultz data. This sample included programs that

were no longer active, whereas the Sultz data included only active programs. Therefore, there were differences in the current status of programs, their institutional settings, and their length in hours and weeks. That is, the National Board data included discontinued programs that were typically shorter in hours and weeks and were sponsored by a wider variety of institutions. On the other hand, the two samples were similar in terms of typical educational level, accreditation status, administration, directors, and class size of programs.

There were differences in the sociodemographic and program profiles from year to year. Between 1977 and 1982, the examinee population became younger and consequently had less experience. Their highest level of education increased, and there were more masters level nurse practitioner programs and program graduates in general. The educational programs have moved into the mainstream of nursing education--they are typically located in NLN schools of nursing, with administrative control vested in nursing and a nurse director or nurse and physician co-directors. Over the past 6 years, programs have increased in length, both in number of weeks and in number of hours of classroom and clinical content.

Bivariate results. There were some intercorrelations between predictor variables. For examinees, those intercorrelations that were of moderate or high magnitude included: between age and RN experience (r=.61); between highest education and exam year (r=.45); between program accreditation status and current status, setting, and administration (r range=.42 to .81); between program administration and directors (r=-.51); and, between program hours and weeks (r=.52).

For educational programs, there were three other moderate magnitude

correlations (in addition to the five program variable intercorrelations mentioned for examinees). These were: between program accreditation status and educational level (r=.42); and between program administration and setting and educational level (r range=.43 to .46). There were more intercorrelations between program variables than between sociodemographic variables, which reflects the effects of individual variation among examinees.

In contrast, there were generally low correlations between the predictor variables and examination score. The highest magnitude correlations for first-time takers were with examinee age (r= -.21) and highest education (r=.28). For the repeats, the relationships with age (r=-.40) and RN experience (r=-.23) were the strongest. On further analysis, however, it was apparent that the low correlations for some variables were due to nonlinearity rather than lack of relationship between the variable and examination performance.

When the values for examinee age, RN experience, and highest education were categorized for analysis of variance, there was evidence of nonlinear relationships. For example, the youngest age group did not have the highest average score (20-24 years: mean=483); the 25-34 year age group did (mean=514). The 20-24 year age group had the same average score as the 35-44 year age group, while the 45-54 year age group had an average score (459) that was not significantly different than the over 55 age group (mean=392).

Similarly, those examinees in the 1-5 year or 5-10 year RN experience group obtained higher average scores (means=511 and 504, respectively) than those with less than 12 months of experience (mean=484). Those with less than 12 months of experience scored like those with

10-15 years of RN experience (mean=483), and those with more than 15 years of experience obtained the lowest average scores (mean=453).

Regression results. Regression analysis was used to evaluate the relationships between examination score and the various predictor variables. The amount of variance in examination scores explained by the sociodemographic variables (3-variable equation) was 8% (R=.28) for the first-time takers and 22% (R=.47) for the repeaters. When type of PNP preparation was added to the equation, these variables explained 15% (R=.39) of the variance for formally prepared examinees and 22% (R=.47) of the variance for informally prepared examinees. The amount of variance explained ranged from 14-26% (R=.37 to .51) when this equation was estimated separately for each examination year.

The educational program variables were even more limited in their predictive ability. For first-time takers, 3% (R=.18) of the variance in scores was explained by the 5-variable equation. That amount increased to 6% (R=.25) for repeats, but the equation was not significant at p=.05 or less. When this equation was estimated separately for each exam year, the percentage ranged from 2-13% (R=.16 to .36).

On the other hand, the 5-variable equation containing the program variables explained 28% (R=.53) of the variance in average performance from program to program. That is, at the aggregate level of analysis (by program) there is obviously less individual variation around the program means and, therefore, greater predictive ability.

There was not much improvement in ability to predict individual performance on the examination when the sociodemographic and program variables were combined. The 8-variable combined model explained 9% (R=.30) of the variance in scores for first-time takers and 26% (R=.51)

of the variance for repeats. When these equations were estimated by year of examination, the R^2 ranged from .15 (15%; R=.39) to .28 (28%; R=.53). The variables that made the largest contributions to differences in scoring were the examinee's age and highest education and the educational level and accreditation status of their nurse practitioner program. As noted previously, the small amount of variance explained by these equations is partially due to lack of linearity between some variables and examination scores.

Interpretation and Limitations

In general, the findings of this research confirm the hypotheses of the investigator (see Chapter IV, Research Model, pp. 67-79) and the results of previous research. Among the sociodemographic characteristics, the most important in determining performance were the examinee's age, highest education, type of preparation, and status as a first-time taker or repeater.

The negative relationship between age and cognitive performance is consistent with other research (AANA, 1983; Aldag & Rose, 1983; Conger & Fitz, 1963; Dawson-Saunders & Doolen, 1980; Dunn, 1981; Hopkins & Stanley, 1981; Johnson & Hutchins, 1966; Lavin, 1965; Mellsop, 1981; NAACOG Certification Corporation, 1980; Reed & Feldhusen, 1972; Reed, Feldhusen & Van Mondfrans, 1973; Tucker & McGaghie, 1982), but was not strictly linear. For this population, the most pronounced effect on performance was seen in those examinees over age 45 years.

The importance of highest education to performance was also confirmed by this research. However, the results do not support the notion that this is a linear relationship, which is consistent with some previous research (AANA, 1982; Farrand et al., 1982; Fleming, 1979). For

example, those examinees with associate degrees scored alike (mean range=432-443) and those with diplomas, bachelors degrees, and masters degrees scored alike (mean range=468-536). (The nine doctorally prepared examinees are excluded from discussion because of their small numbers.)

The investigator postulated that the bachelors degree was the threshold level for education. That is, that examinees with bachelors degrees would receive higher scores than those with less or more formal education. To examine this possibility, a supplementary analysis was performed that looked at the significance of differences in mean scores for examinees with BSNs and MSNs. There were no differences for the repeats, or for those first-time takers with masters degrees in nursing who attended certificate (mean=541) or masters (mean=536) level nurse practitioner programs. There were significant differences in scores between certificate program graduates who had MSNs (mean=541) versus BSNs (mean=514), and between MSN masters program graduates (mean=536) and BSN certificate program graduates (mean=514). Based on these results, the idea of the bachelors degree as the educational threshold could not be supported. These results also indicated that the educational level of the individual examinee (masters or bachelors degree) was more important that the educational level of the nurse practitioner program (masters or certificate).

It was expected that formally prepared examinees would achieve higher scores than informally prepared examinees. This was substantiated by these results and agrees with the findings of others (NAACOG Certification Corporation, 1980, 1982). In addition, it was no surprise that

the first-time takers performed at higher levels on the examination than the repeaters. The repeaters do appear to be a norm group different than the first-time takers, as previously described by Fleming (1979), Fullerton and Thompson (1983), and Mellsop (1981).

Other sociodemographic characteristics under investigation included experience, current function, employment setting, and sex. Because of the small number of men in the study population (1.8%), relationships between examinee sex and performance could not be confirmed or refuted by this research. As expected, there were negative relationships between length of experience as a RN and PNP and examination scores. The relationship between RN experience and performance, however, was not linear; the idea of an experience threshold has been documented by other research (Downing & Maatsch, 1979; Dunn, 1981; Farrand et al., 1982; Maatsch, 1981; Mellsop, 1981; NAACOG Certification Corporation, 1980; Pawluk et al., 1976).

There was no relationship between performance and whether the examinee was currently functioning as a nurse practitioner. However, the results did confirm the "faculty effect" described by Fullerton and Thompson (1983) for nurse midwives, with faculty members obtaining the highest mean scores (535). Interestingly, the unemployed examinees' scores were not significantly different (mean=527) than the faculty scores--perhaps related to greater examination preparation time. Those examinees in more specialized employment settings (school and inpatient) received lower scores than all other examinees (means=466, school; 450, inpatient). This effect was demonstrated previously, among obstetric-gynecologic nurse practitioners (NAACOG Certification Corporation, 1980).

Among the educational program characteristics, the most important

were the program's educational level and accreditation status. It was expected that graduates of masters level programs would achieve higher scores, on average, than graduates of certificate programs. This was confirmed, and agrees with previous research on nurses in expanded roles (AANA, 1983; Dunn, 1981; Farrand et al., 1982; Fleming, 1979; NAACOG Certification Corporation, 1980).

Although no other research was identified that examined the relationship between program accreditation status and cognitive performance of graduates, the results did substantiate the investigator's expectations. That is, that graduates of NLN accredited programs would perform at higher levels than graduates of unaccredited programs (or of ANA accredited programs). (Note that there were only five ANA accredited programs in this sample, two of which were also NLN accredited. The three programs whose sole accreditation was from ANA had only 145 graduates in this sample.)

It was anticipated that graduates of university-sponsored programs would perform better than others. While this was true (mean=506), military-sponsored program graduates performed about the same (mean=519). This result among military-sponsored program graduates was also found among nurse anesthetist certification candidates (AANA, 1983; Fleming, 1979).

Those examinees whose programs were administered by nursing were expected to receive higher scores than those whose programs were administered jointly or by medicine. This was based on theoretical assumptions, as no other research in this area was identified. In terms of the average performance of examinees, this assumption was verified (means=506, nursing; 496, medicine; 495, joint).

Graduates of programs with joint co-directors did not perform as well as those whose programs were directed by a nurse or by a physician, on the average. This was not the expected effect, and there was no other research to support these findings. In contrast to the investigator's expectations, there were no relationships between examination performance and the year the examinee's program was established or the class size. The finding regarding class size is consistent with previous research by Martin and others (1980).

Likewise, the length of examinees' programs in hours had no effect on their scores. There was, however, a significant inverse relationship between average score and program length in weeks. This negative relationship was earlier described by Mellsop (1981) among physicians, and does not support Fleming's (1979) finding of a positive relationship for nurse anesthetists. Finally, like other research (AANA, 1983; Fleming, 1979), graduates of programs that had been discontinued received lower scores than those whose programs were still active. This was the anticipated result.

Limitations. Previous research suggests that level of academic achievement is most consistently related to other cognitive measures. Unfortunately, such measures were not available for this research. One obvious limitation of this research, therefore, is lack of inclusion of the most relevant predictor variables.

With regard to the sample population in general, the examinees were a relatively homogeneous group--at least the formally prepared first-time takers--with resulting restriction of range on the criterion measure. Also, because the criterion measure was a composite examination score, there was some loss of dimensionality and, therefore,

information (Hogan, Gallagher, Sirotkin, Wolfe & Scalzi, 1975, pp. 315-318).

This sample represents about 55% of the total population of pediatric nurse practitioners in the country, and the findings may be generalizable to that population. However, the results are probably not generalizable to other types of nurse practitioners, because of differences in their educational experiences, in their sociodemographic profiles, and in their certification mechanisms.

According to the definition used by the National Commission for Health Certifying Agencies, it could be said that these results establish differential validity for the National Board examination. That is, differential performance that is related to demographic differences (Report of the NCHCA, 1981, p. 19). This research, on the other hand, has not established discriminant validity for this certification examination, since differences in scoring cannot be directly related to differences in competence (Carmines & Zeller, 1979, p. 54).

Implications and Further Research

If only the bivariate results of this research were considered, the implications would be very different than they are. Once the predictor variables were examined simultaneously, in the multivariate analyses, however, many of the relationships changed and in some cases became insignificant. Given that most of the differences in performance between examinees were not explained by this research, the conclusions must be limited.

It is clear that, on the individual level, sociodemographic characteristics make a larger contribution to differences in scoring than do characteristics of the examinee's educational program. In the

aggregate-level analysis of average program score by program, which took group membership into account, the predictive power of the program variables was greatly improved.

Individual level. For nurses considering application to nurse practitioner programs, this research would support attendance at a NLN accredited program located in a school of nursing, or at a military-sponsored program. Those applicants (BSN prepared) with a choice of entering masters or certificate level programs should carefully consider their objectives in attending the program, the program cost and length, and their future goals. For those individuals who intend to work as nurse practitioners, time and financial costs can be decreased by attending a certificate level program. However, this research indicates that masters prepared nurse practitioners are typically higher achievers.

For those nurse practitioner program graduates who have worked in narrow-focus employment settings since graduation, such as inpatient hospital units or school systems, it would be advisable to establish an organized self-assessment and review program prior to sitting for the certification examination. Additionally, those nurse practitioners who fail the examination on the first try should take specific steps that may improve their test performance, such as organized review courses, self-assessment, and coaching.

Program level. In relation to admissions requirements, most programs currently consider factors such as previous education and experience, age, and measures such as grade point average. Those applicants with associate degrees should be scrutinized carefully, as should those with doctorates and those over age 35.

The class size of the program does not appear to make much difference, nor does the length in hours. There appears to be no justification, in terms of graduate performance, for the trend toward longer programs (currently, the average is 9 months). If the results of this research are any indication, there is also no reason for programs located in NLN accredited schools of nursing to seek ANA accreditation.

These results also offer some support to those schools that have progressed from certificate to masters level programs, in terms of cognitive performance of graduates. However, information on the costs of masters programs in nursing, and particularly those with nurse practitioner options, needs to be collected on a national level and made available. Given that data, potential applicants as well as policy-makers can make informed choices about certificate or masters level programs.

State level. Information on the costs of masters level nurse practitioner programs is also needed for state-level funding agencies. If the goal is to produce primary care providers who will continue to work as nurse practitioners, in the shortest possible time, then support for certificate level programs should be continued.

In spite of the results of this research, state level regulatory bodies should be particularly careful in supporting a particular educational requirement for nurse practitioners. The ideological positions of national nursing groups (such as ANA and NLN) often influence those individuals responsible for regulatory policy at the state level. It is risky to base exclusionary regulations on ideological grounds or on the limited amount of research available. To do so makes private

credentialing by non-governmental boards a public prohibition against practice, which could certainly be attacked on antitrust grounds (Havighurst & King, 1983, p. 132).

It is prudent, therefore, in those jurisdictions that provide state certification of nursing specialties, to allow both educational and national certification options for practice. That is, if the regulatory agency believes it is justified in placing educational restrictions on applicants for state-level certification, those applicants without the designated education should have another option, such as successful examination by a national certification board.

In addition, this research does not support the idea that ANA accreditation of nurse practitioner programs is an indication of their quality, at least in terms of the performance of their graduates. State agencies that consider national accreditation in their approval of applicants' educational programs should rely on NLN accreditation or an examination of the curriculum of the particular program.

National level. Like state-level funding agencies, the federal government should continue to support certificate level nurse practitioner programs, if the goal is to produce primary care providers at the lowest cost.

This research supports some of the national guidelines for nurse practitioner programs and does not support others. It supports the funding guidelines established by the Department of Health, Education and Welfare in 1976 (now DHHS) related to location of the programs in university settings. On the other hand, there appears to be no justification--in terms of graduate performance--for specifying a minimum class size of 8 students or a minimum length of one academic year.

Federal agencies establishing regulations related to the practice or reimbursement of nurse practitioners should be cautious in requiring exclusionary educational preparation, as mentioned in regard to state-level implications. While this research provides evidence that masters prepared graduates of nurse practitioner programs perform better on certification exams, this by itself should not be translated into restrictive policies for certificate level or non-masters prepared nurse practitioners.

Similarly, private credentialing bodies, such as the National Board, should be circumspect in their eligibility requirements related to applicants' educational degrees and program level. As mentioned before, there is no indication that candidate performance on a cognitive examination is directly related to clinical competence. Previous research in this area, for both nursing and medicine, has produced conflicting results (Downing & Maatsch, 1979; Dunn, 1981; Gonnella, 1973; Hastings, Sasmor & Murray, 1975; Hoekelman, 1975; Lang, 1979; Maatsch, 1981; McGuire & Williamson, 1968; Pawluk et al., 1976).

Further research. As previously mentioned, there is a need for research on the costs of masters programs in nursing, and particularly those programs with nurse practitioner options. Also, there is a general need for more research on specialty certification in nursing, as well as a need for private credentialing boards to exchange information and research related to their mechanisms. Such research should include identification of predictors of performance and relationships between examination performance and clinical performance.

For pediatric nurse practitioners, it would be interesting to compare the results of this research to similar research on ANA certified

nurse practitioners. Additionally, it would be of interest to look at the dimensionality of certification examinations for expanded role nurses, to determine what cognitive skills they actually measure.

Finally, it is recommended that the National Board of Pediatric Nurse Practitioners and Associates continue to collect data on certification examinees, and to use this information for future decision making. For example, based on this and other research, there is no doubt that those who repeat examinations are a different norm group than first-time takers. One mechanism proposed for recertification, by the National Board and other credentialing bodies, is reexamination of applicants. It could be predicted (AANA, 1983; Fleming, 1979; Fullerton & Thompson, 1983) that recertification candidates would perform like repeaters.

If a different examination is designed for recertification candidates, will the standards be lower? How can lower standards be justified if the purpose of recertification is measurement of continuing competence? These and other credentialing dilemmas related to nurse practitioners, and to other health professionals, have yet to be resolved.

Reference List

Reference List

Achen, C.H. Interpreting and using regression. Beverly Hills, CA.: Sage Publications, Inc., 1982.

Ainslie, B.S., Anderson, L.E., Colby, B.K., Hoffman, M.A., Meserve, K.P., O'Connor, C., & Quimet, K.M. Predictive value of selected admission criteria for graduate nursing education. Nursing Research, 1976, 25, 296-299.

Aldag, J. & Rose, S. Relationship of age, American College Testing scores, grade point average, and state board examination scores. Research in Nursing and Health, 1983, 6, 69-73.

American Association of Nurse Anesthetists. Certification examination for nurse anesthetists (1979-1982). Park Ridge, IL.: AANA, 1983. (Unpublished data)

American Nurses Association. Directory of accredited organizations, approved programs/offerings, and accredited continuing education certificate programs preparing nurse practitioners. Kansas City, MO.: ANA, 1982.

Asher, H.B. Causal modeling. Beverly Hills, CA.: Sage Publications, Inc., 1976.

Association of Faculties of Pediatric Nurse Associate/Practitioner Programs. Philosophy, conceptual model, terminal competencies for the education of pediatric nurse practitioners. AFPNA/PP, Inc., 1982. (Brochure)

Backman, M. & Steindler, F. Let's examine-prediction of achievement in a collegiate nursing program and performance on state board exams. Nursing Outlook, 1971, 19, 487.

Bejar, I.I. Achievement testing--Recent advances. Beverly Hills, CA.: Sage Publications, Inc., 1983.

Bell, J.S. & Sanchez, K.S. The prediction of test performance on licensure examinations in nursing. Educational and Psychological Measurement, 1980, 40, 1147-1152.

Benenson, T.F., Stimmel, B., & Aufses, A. Concordance of surgical clerkship performance and National Board of Medical Examiners Part II subtest scores. Surgery, 1981, 89, 692-696.

Benor, D. & Hobfoll, S.E. Prediction of clinical performance: The role of prior experience. Journal of Medical Education, 1981, 65, 653-658.

Bohan, K.M. Performance relationship: Nursing student to professional nurse. Washington, D.C.: Catholic University, 1966. (Unpublished doctoral dissertation)

Brandt, E.M. & Metheny, B.H. Relationships between measures of student and graduate performance. Nursing Research, 1968, 17, 242-246.

Bucher, R. & Stelling, J.G. Becoming professional. Beverly Hills, CA.: Sage Publications, Inc., 1977.

Burg, F.D., Brownlee, R.C., Wright, F.H., Levine, H., Daeschner, C.W., Vaughan, V.C., & Anderson, J.A. A method for defining competency in pediatrics. Journal of Medical Education, 1976, 51, 824-828.

Carfang, C.J. Factors influencing the pediatric nurse practitioner's perceived effectiveness. Madison, WI.: University of Wisconsin, 1979. (Unpublished master's thesis)

Carline, J.D., Cullen, T.J., & Scott, C.S. Prediction of clerkship performance using the new MCAT examination: An attempted application of canonical redundancy analysis. In Research in Medical Education: 1982, Proceedings of the 21st Annual Conference. Washington, D.C.: Association of American Medical Colleges, 1982.

Carmines, E.G. & Zeller, R.A. Reliability and validity assessment. Beverly Hills, CA.: Sage Publications, Inc., 1979.

Cooley, W.W. & Lohnes, P.R. Evaluation research in education. New York: Irvington Publishers, Inc., 1976.

Conger, J.J. & Fitz, R.H. Prediction of success in medical school. Journal of Medical Education, 1963, 38, 943-948.

Cronbach, L.J. Designing evaluations of educational and social programs. San Francisco: Jossey-Bass Publishers, 1982.

Cullen, T.J., Dohner, C.W., Peckham, P.D., & Samson, W.E. Restriction of range and the predictive validity of the new Medical College Admissions Test. In Proceedings of the Nineteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1980.

Dagenais, F. & Rosinski, E.F. Social class level, performance, and change in medical school. In Proceedings of the Fourteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1975.

Daniel, W.W. Biostatistics: A foundation for analysis in the health sciences. New York: John Wiley & Sons, Inc., 1974.

Davis, C.K. The federal role in changing health care financing--Part I. Nursing Economics, 1983a, 1, 10-17.

Davis, C.K. The federal role in changing health care financing--Part II. Nursing Economics, 1983b, 1, 98-104; 146.

Dawson-Saunders, B. & Doolen, D.R. Canonical redundancy analysis: A new technique to predict performance. In Proceedings of the Nineteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1980.

Deighton, B.C., Smith, D.W., & Gallagher, R.E. The relationship between scores on the new Medical College Admissions Test and performance in the first year of medical school: A preliminary study. In Proceedings of the Eighteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1979.

Department of Health, Education & Welfare. Medicaid and Medicare programs: Rural health clinic services. Federal Register, 14 July 1978, 43(136), 30528.

Department of Health, Education & Welfare. Nurse practitioner training. Federal Register, 23 January 1976, 41(16), 3552-3553.

Department of Health, Education & Welfare, Secretary's Committee to Study Extended Roles for Nurses. Extending the scope of nursing practice. Washington, D.C.: U.S. Government Printing Office, 1971.

Department of Health & Human Services, Division of Nursing. A directory of expanded role programs for registered nurses: October 1982. Hyattsville, MD.: DHHS, 1982.

DeVaul, R.A., Jervey, F.L., O'Keefe, S., & Short, B. Admissions preference as a predictor of preclinical performance. In Research in Medical Education: 1982, Proceedings of the 21st Annual Conference. Washington, D.C.: Association of American Medical Colleges, 1982.

Dickenson-Hazard, N. Personal communication. National Board of Pediatric Nurse Practitioners and Associates, November 1982 & August 1983.

Donabedian, A. Some basic issues in evaluating the quality of health care. In Issues in Evaluation Research. Kansas City, MO.: American Nurses Association, 1976.

Donnelly, J.C. Coping and ego development in internship. In Proceedings of the Eighteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1979.

Downing, S.M. & Maatsch, J.L. The effects of clinically relevant multiple choice items on the statistical discrimination of physician clinical competence. In Proceedings of the Eighteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1979.

Duncan, D.B. Multiple range and multiple F tests. Biometrics, 1955, 11, 1-42.

Dunn, B.H. Nurse practitioners: Predictors of performance. Richmond, VA.: Virginia Commonwealth University, 1981. (Unpublished manuscript)

Dunn, B.H. & Brown, B.S. Certification: The credentialing dilemma. Virginia Nurse, 1982a, 50(3), 10-11.

Dunn, B.H. & Brown, B.S. Whither nurse practitioners...to wither? Virginia Nurse, 1982b, 50(1), 10-11.

Dunn, B.H. & Chard, M. Nurse practitioners: A review of the literature, 1965-1979. Kansas City, MO.: American Nurses Association, 1980.

Edari, R.S. & staff. A profile of the certified nurse. In The Study of Credentialing in Nursing: A New Approach (vol. 2, Staff Working Papers). Kansas City, MO.: American Nurses Association, 1979.

Egan, R.L. Evaluation and the recent medical graduate. Journal of the American Medical Association, 1982, 247, 2931-2933.

Erdmann, J.B., Jones, R.F., & Tonesk, X. AAMC longitudinal study of medical school graduates of 1960. Hyattsville, MD.: National Center for Health Services Research, DHEW, 1979. (DHEW Pub. No. (PHS) 79-3235)

Farrand, L.L., Holzemer, W.L., & Schleutermann, J.A. A study of construct validity: Simulations as a measure of nurse practitioners' problem-solving skills. Nursing Research, 1982, 31, 37-42.

Fleming, P.R. The national qualifying examination for certification of nurse anesthetists: A report covering six examinations--November, 1975-June, 1978. Journal of the American Association of Nurse Anesthetists, 1979, 47(1), 25-38.

Friedman, C.P., Cheatham, C.M., Porter, C.O., & Bakewell, W.E. A nonlinear discriminant analytic approach to predicting medical student performance. In Proceedings of the Eighteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1979.

Frisbie, D.A. Evaluating student achievement: Principles, trends, and problems. New York: National League for Nursing, 1979. (Pub. No. 23-1766)

Fuchs, V.R. Who shall live? Health, economics, and social choice. New York: Basic Books, Inc., 1974.

Fullerton, J.T. & Thompson, J.E. Recertification in nurse-midwifery: A critical analysis of use of a written examination. Manuscript submitted for publication, 1983.

Greenlick, M.R. Assessing clinical competence. Evaluation and the Health Professions, 1981, 4, 3-12.

Gonnella, J.S. Evaluation of patient care: An approach. Journal of the American Medical Association, 1973, 214, 2040-3043.

Gough, H.G. Some predictive implications of premedical school competence and performance. Journal of Medical Education, 1978, 53, 291-300.

Guidelines on short-term continuing education programs for pediatric nurse associates. Pediatrics, 1971, 47, 1075.

Guilford, J.P. & Fruchter, B. Fundamental statistics in psychology and education (5th ed.). New York: McGraw-Hill Book Co., 1973.

Hastings, G.E., Sasmor, L., & Murray, E. Performance levels of primary nurse practitioners--Implications for medical education. In Proceedings of the Fourteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1975.

Havighurst, C.C. & King, N.M.P. Private credentialing of health care personnel: An antitrust perspective--Part one. American Journal of Law & Medicine, 1983, 9, 131-201.

Hinsvark, I.G. & Dorsch, H. Credentialing as it effects schools of nursing. In The Study of Credentialing in Nursing: A New Approach (vol. 2, Staff Working Papers). Kansas City, MO.: American Nurses Association, 1979.

Hobfoll, S.E. & Benor, D.E. Prediction of student clinical performance. Medical Education, 1981, 15, 231-236.

Hoekelman, R.N. An evaluation of preparation of pediatric nurse practitioners. Rochester, N.Y.: University of Rochester, 1975. (Unpublished manuscript)

Hogan, M.J., Gallagher, R.E., Sirotkin, R.A., Wolfe, B.L., & Scalzi, P.M. Interpretation of the dimensionality of multiple choice tests. In Proceedings of the Fourteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1975.

Holzemer, W.L., Farrand, L.L., & Schleutermann, J.A. The validation of patient management problems for nurse practitioners. Nursing Research, 1981, 30, 139-144.

Hopkins, K.D. & Stanley, J.C. Educational and psychological measurement and evaluation (6th ed.). Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1981.

Hughes, F.P. Norm-referenced or criterion-referenced standards: A question of judgment. Philadelphia: National Board of Medical Examiners, 1982.

Hutcheson, J.D., Garland, L.M., & Prather, J.E. Toward reducing attrition in baccalaureate degree programs: An exploratory study. Nursing Research, 1973, 22, 530-533.

Hutton, E. & Rorabaugh, M.L. Survey of state certification requirements for nurse practitioners. Manuscript submitted for publication, 1983.

Iversen, G.R. & Norpoth, H. Analysis of variance. Beverly Hills, CA.: Sage Publications, Inc., 1976.

Johnson, D.G. & Hutchins, E.B. Doctor or dropout: A study of medical student attrition. Journal of Medical Education, 1966, 41, 1099-1269.

Kahn, L. The influence of funding on the future of pediatric nurse practitioner programs. Pediatrics, 1979, 64, 106-110.

Kane, M.T. Evaluating student achievement. In Measurement and Evaluation in Nursing Education. New York: National League for Nursing, 1980a. (Pub. No. 17-1807)

Kane, M.T. Norm-referenced tests and criterion-referenced tests. In Measurement and Evaluation in Nursing Education. New York: National League for Nursing, 1980b. (Pub. No. 17-1807)

Keck, J.W., Arnold, L., Willoughby, L., & Calkins, V. Efficacy of cogntive/noncognitive measures in predicting resident physician performance. Journal of Medical Education, 1979, 54, 759-765.

Kissinger, J.D. & Munjas, B.A. Predictors of student success. Nursing Outlook, 1982, 32, 53-54.

LaDuca, A. The structure of competence in health professions. Evaluation and the Health Professions, 1980, 3, 253-288.

Lang, N.M. The relationships between credentialing and outcomes of service: A review and analysis of existing reports. In The Study of Credentialing in Nursing: A New Approach (vol. 2, Staff Working Papers). Kansas City, MO.: American Nurses Association, 1979.

Lavin, D.E. The prediction of academic performance. New York: Russell Sage Foundation, 1965.

Lee, P.R., LeRoy, L., Stalcup, J., & Beck, J. Primary care in a specialized world. Cambridge, MA.: Ballinger Publishing Co., 1976.

LeRoy, L. & Solkowitz, S. The costs and effectiveness of nurse practitioners. In The Implications of Cost-Effectiveness Analysis of Medical Technology. Washington, D.C.: Office of Technology Assessment, U.S. Congress, 1981.

Litherland, R.L. Iowa tests of educational development as a predictor of academic success in Iowa schools of professional nursing. Iowa City, IA.: University of Iowa, 1966. (Unpublished doctoral dissertation)

Lloyd, J.S. Explicit definitions of competence developed by specialty boards. In Proceedings of the Nineteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1980.

Maatsch, J.L. Assessment of clinical competence on the emergency medicine specialty certification examination: The validity of examiner ratings of simulated clinical encounters. Annals of Emergency Medicine, 1981, 10, 504-507.

Margolis, C.Z. & Cook, C.D. Rating pediatric houseofficer performance. Pediatric Research, 1974, 8, 472.

Martin, G.M., Gullickson, G., & Gerken, C. Graduate medical education and certification in physical medicine and rehabilitation. Archives of Physical Medicine and Rehabilitation, 1980, 61, 291-297.

Martuza, V.R. Applying norm-referenced and criterion-referenced measurement in education. Boston: Allyn & Bacon, Inc., 1977.

McCloskey, J.C. Nursing education and job effectiveness. Nursing Research, 1983, 32, 53-58.

McGaghie, W.C. The evaluation of competence. Evaluation and the Health Professions, 1980, 3, 289-320.

McGuire, C. & Williamson, J.W. Consecutive case conference: An educational evaluation. Journal of the American Medical Association, 1968, 43, 1068-1074.

Mehrens, W.A. & Lehmann, I.J. Measurement and evaluation in education and psychology (2nd ed.). New York: Holt, Rinehart & Winston, 1978.

Mellsop, G. Performance correlates in a postgraduate psychiatric examination. Medical Education, 1981, 15, 167-170.

Mezey, M.D. Securing a financial base. American Journal of Nursing, 1983, 83, 1297-1298.

Michael, W.B., Haney, R., & Jones, R.A. The predictive validities of selected aptitude and achievement measures and of three personality inventories in relation to nursing training criteria. Educational and Psychological Measurement, 1966, 26, 1035-1040.

Mueller, E.J. & Lyman, H.B. The prediction of scores on state board test pool examinations. Nursing Research, 1973, 19, 263-267.

Munday, L. & Hoyt, D.P. Predicting academic success for nursing students. Nursing Research, 1965, 14, 341-344.

Murden, R., Galloway, G.M., Reed, J.C., & Colwill, J.M. Academic and personal characteristics as predictors of clinical success in medical school. In Proceedings of the Sixteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1977.

NAACOG Certification Corporation. Newsletter. December 1980 & January/February 1982.

National League for Nursing. Master's education in nursing: Route to opportunities in contemporary nursing 1982-83. New York: NLN, 1982. (Pub. No. 15-1312)

Nie, N.H., Hull, C.H., Jenkins, J.G., Steinbrenner, K., & Brent, D.H. Statistical package for the social sciences. New York: McGraw-Hill Book Co., 1975.

O'Donnell, M.J. NBME part I examination: Possible explanations for performance based on personality type. Journal of Medical Education, 1982, 57, 868-870.

Paiva, R.E.A. Reliability in follow-up evaluations of graduates as residents. In Proceedings of the Eighteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1979.

Passarelli, A. Credentialing in nursing: Background issues. In The Study of Credentialing in Nursing: A New Approach (vol. 2, Staff Working Papers). Kansas City, MO.: American Nurses Association, 1979.

Pawluk, J., Roberts, R., Tech, M., & Neufeld, M.D. Concurrent validity of the Canadian certification examination in family medicine. In Proceedings of the Fifteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges 1976.

Plapp, J.M., Psathas, G., & Caputo, C.V. Intellective predictors of success in nursing school. Educational and Psychological Measurement, 1965, 25, 565-577.

Pollard, M.R. & Leibenluft, R.F. Antitrust and the health professions. Washington, D.C.: Federal Trade Commission, Office of Policy Planning, July 1981.

Pollard, M.R. & Schultheiss, P. FTC and the professions: Continuing controversy. Nursing Economics, 1983, 1, 158-163.

Raffetto, A.M. & Zabarenko, L.M. Noncognitive evaluation of medical training--A paracognitive approach. Evaluation and the Health Professions, 1979, 2, 161-180.

Reed, C.L. & Feldhusen, J.F. State board examination score predictors for associate degree nursing program graduates. Nursing Research, 1972, 21, 149-153.

Reed, C.L., Feldhusen, J.F., & Van Mondfrans, A.P. Prediction of grade point average using cognitive and noncognitive predictor variables. Psychological Reports, 1973, 32, 143-148.

Reed, S.B. & Riley, W. Comprehensive evaluation model for nursing education. Evaluation and the Health Professions, 1979, 2, 438-454.

Report of the Committee. The study of credentialing in nursing: A new approach (vol. 1). Kansas City, MO.: American Nurses Association, 1979.

Report of the Committee on Goals and Priorities of the National Board of Medical Examiners. Evaluation in the continuum of medical education. Philadelphia: National Board of Medical Examiners, 1973.

Report of the Graduate Medical Education National Advisory Committee. GMENAC summary report (vol. 1). Washington, D.C.: U.S. Government Printing Office, 1981a. (DHHS Pub. No. (HRA) 81-651)

Report of the Graduate Medical Education National Advisory Committee. Nonphysician health care provider technical panel (vol. 6). Washington, D.C.: U.S. Government Printing Office, 1981b. (DHHS Pub. No. (HRA) 81-656)

Report of the National Commission for Health Certifying Agencies. Perspectives on health occupational credentialing. Hyattsville, MD.: Division of Associated Health Professions, DHHS, 1981. (DHHS Pub. No. (HRA) 81-4)

Roemer, M.I. Comparative national policies on health care. New York: Marcel Dekker, Inc., 1977.

Rose, S.D., Corman, L.C., & Robbins, J. Failure of examination answers to evaluate actual practice patterns by medical housestaff in an outpatient clinic. In Proceedings of the Eighteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1979.

Saffer, J.B. & Saffer, L.D. Academic record as a predictor of future job performance of nurses. Nursing Research, 1972, 21, 457-462.

Salman, R.D. Continuing competence and the credentialing of professionals. In Report of the National Commission for Health Certifying Agencies: To Assure Continuing Competence. Washington, D.C.: DHHS, 1981. (DHHS Pub. No. (HRA) 81-5)

SAS Institute. SAS user's guide: Statistics, 1982 edition. Cary, N.C.: SAS Institute, Inc., 1982.

Shimberg, B. Occupational licensing. Princeton, N.J.: Educational Testing Service, 1982.

Sime, A.M., Corcoran, S.A., & Libera, M.B. Prediting success in graduate education. Journal of Nursing Education, 1983, 22, 7-11.

Steele, S. Educational evaluation in nursing. Thorofare, N.J.: Charles B. Slack, Inc., 1978.

Stein, R.F. The graduate record examination: Does it predict performance in nursing programs? Nurse Educator, 1978, 3(4), 16-19.

Subcommittee on Health Manpower Credentialing. A proposal for credentialing health manpower. Washington, D.C.: HRA, PHS, DHEW, 1976.

Sultz, H.A., Bullough, B., Kinyon, L.J., Buck, G.M., & Sherwin, M.A. National study of nurse practitioners. Buffalo, N.Y.: State University of New York at Buffalo, 1983. (Unpublished data)

Sultz, H.A., Henry, O.M., Bullough, B., Buck, G.M., & Kinyon, L.J. Nurse practitioners: A decade of change---Part III. Nursing Outlook, 1983, 31, 266-269.

Sultz, H.A., Henry, O.M., Kinyon, L.J., Buck, G.M., & Bullough, B. A decade of change for nurse practitioners. Nursing Outlook, 1983a, 31, 137-141; 188.

Sultz, H.A., Henry, O.M., Kinyon, L.J., Buck, G.M., & Bullough, B. Nurse practitioners: A decade of change--Part II. Nursing Outlook, 1983b, 31, 216-219.

Sultz, H.A. & Kinyon, L.J. Longitudinal study of nurse practitioners: Phase I. Bethesda, MD.: DHEW, 1976. (DHEW Pub. No. (HRA) 76-43)

Sultz, H.A., Zielezny, M., Gentry, J.M., & Kinyon, L.J. Longitudinal study of nurse practitioners: Phase II. Hyattsville, MD.: DHEW, 1978. (DHEW Pub. No. (HRA) 78-92)

Sultz, H.A., Zielezny, M., Gentry, J.M., & Kinyon, L.J. Longitudinal study of nurse practitioners: Phase III. Hyattsville, MD.: DHHS, 1980. (DHHS Pub. No. (HRA) 80-2)

Thomas, B. Prediction of success in a graduate nursing service administration program. Nursing Research, 1974, 23, 156-159.

Thorndike, R.L. & Hagen, E.P. Measurement and evaluation in psychology and education (4th ed.). New York: John Wiley & Sons, Inc., 1977.

Tillinghast, B.S. & Norris, B. Let's examine: The relationship of selected admissions variables to student achievement. Nursing Outlook, 1968, 16, 58.

Tripp, A. & Duffey, M. Discriminant analysis to predict graduation-nongraduation in a master's degree program in nursing. Research in Nursing and Health, 1981, 4, 345-353.

Trussell, R.P. & Pappas, J.P. Some variables for student nurse selection. Report No. 47, Rocky Mountain Psychological Association, University of Utah, 1974.

Tucker, J.W. & McGaghie, W.C. Does student age predict preclinical medical school achievement? Evaluation and the Health Professions, 1982, 5, 489-498.

Veloski, J. & Gonnella, J.S. Performance on part III of the national boards--The effect of residency training. In Proceedings of the Nineteenth Annual Conference on Research in Medical Education. Washington, D.C.: Association of American Medical Colleges, 1980.

Weisberg, S. Applied linear regression. New York: John Wiley & Sons, Inc., 1980.

Werner, E.R., Adler, R., Robinson, R., & Korsch, B. Attitudes and interpersonal skills during pediatric internship. Pediatrics, 1979, 63, 491-499.

Wolfe, L.M. & Bryant, L.W. A causal model of nursing education and state board examination scores. Nursing Research, 1978, 27, 311-315.

Zagar, R., Arbit, J., & Wengel, W. Personality factors as predictors of grade point average and graduation from nursing school. Educational and Psychological Measurement, 1982, 42, 1169-1175.

Appendix

Appendix A. Data Analysis

NBME Tape Data

The computer tape obtained from the National Board of Medical Examiners (NBME) included the sociodemographic characteristics and examination scores of examinees. In preparing these data for analysis, some clean-up and recoding was necessary.

Age. Of the missing values listed for this variable (13), information was actually missing in only 2 cases. In the other 11 cases, the information was determined to be invalid and was deleted. The invalid cases involved examinees listing the exam year instead of their year of birth when asked to provide their birth date (month, day, year) on the application form. When these birthdates were converted to years, therefore, 11 examinees were "0" years.

RN experience. On the application form, examinees were asked to provide the number of months of RN experience, exclusive of PNP experience. On the computer tape, entries with missing information were coded "000." There was no apparent distinction between those with no RN experience and those with missing information (that is, both would be coded "000" according to the NBME coding scheme). Entries coded this way were treated as missing values for these analyses, since it is unlikely that candidates had no RN experience. Of the 171 cases involved, 111 occurred in 1977 or 1978, which raises a question of

coding differences in those years (particularly in view of other coding problems in those years, discussed below).

That accounts for 171 of 183 entries listed as missing for this variable; the remaining 12 were actually invalid data and were deleted. These entries were invalid because the months of RN experience given exceeded the maximum number possible, given the oldest examinee in each examination year. For example, the oldest examinee tested in 1978 was 63. Assuming that this examinee had practiced full-time since graduation from nursing school (about age 21), the maximum length of experience possible would be 42 years or 504 months. For each examination year, this determination was made and those entries over the limit were deleted.

PNP experience. The NBME coding for this variable was the same as for RN experience. That is, missing entries were coded "000," and there was no apparent distinction for candidates with no PNP experience. In this case, however, it was likely that a large number of examinees had no experience, since many take the examination as soon as they complete their programs. Therefore, where PNP experience was coded "000" and RN experience was not "000," it was assumed that the examinee had no experience. This procedure resulted in 423 of 453 cases originally coded as missing being reclassified as "no PNP experience" (leaving 30 cases as missing).

The remaining 8 cases considered missing (total missing=38) were actually invalid. As was true for RN experience, some candidates' length of experience as PNPs exceeded the limits possible. To determine outliers, 1966 was used as the earliest possible date to begin PNP practice (since this was the year that the first formally prepared

PNPs completed their program). In each case where it appeared that the maximum number of months for a particular examination year was exceeded, the investigator checked to see whether the examinee was formally or informally prepared. As a result of this process, 8 entries for formally prepared candidates were deleted.

Highest education. In 1977 and 1978, the application form was structured differently than in later years. For example, in 1977 the information requested for highest education asked examinees to list their highest non-nursing degree. In 1978, it asked for highest degree and diploma graduates were excluded (considered missing). In spite of these discrepancies, the investigator was able to convert and recode these variables, for 1977 and 1978, so that they were in the same format as later years. No missing values were generated as a result of this procedure.

Basic education. Although this information was requested on the application form, and was provided on the NBME tape, it could not be used. In this case, the question and coding for 1977 was different than for 1978-1982. In 1977 candidates were asked to give their basic nursing education (masters, bachelors, diploma, associate, other), while in other years they were given three choices (bachelors, diploma, associate). The investigator was not able to convert and reclassify this information, which resulted in loss of this variable in 1977. Because 1977 had the largest examinee population (943), the decision was made to omit the variable entirely rather than generating missing values for 29% of the total examinee population.

Employment setting. In 1977 there was no category on the application form for those examinees who were employed as RNs (but were not

functioning as PNPs) or for those who were not employed (these candidates were listed as missing information on the NBME tape). These entries (101) for 1977 were reclassified as "employed as RNs, not functioning as PNPs" by the investigator.

For all other years, 10 employment classifications were provided on the application form. The investigator chose to use only 9 classifications, combining 2 original categories for private practice settings (with pediatrician; with other physician). This was done because of the relatively small number of examinees practicing with "other" physicians (84).

Type of PNP preparation. This variable was created by the investigator, based on program information provided on the application form and NBME tape. The information provided was the National Board's 3-digit code for the educational program; informally prepared candidates were coded "000." Since there were no missing data for program codes, all those coded "000" were considered informally prepared, and all those with any other code were considered formally prepared.

As mentioned in the text, one informally prepared candidate was allowed to sit for the National Qualifying Examination in 1977, and four in 1982. The reason for this departure from National Board policy for 1977 is not clear. The four candidates in 1982 were originally tested, at one site, in 1981. Because of problems with testing conditions at the particular site, these candidates were retested in 1982 and their 1981 test results were deleted from all official records.

Program codes. As mentioned above, the NBME tape contained the National Board's code number for each formal educational program. Between 1977 and 1982, a number of programs went from active to inactive

(not operational) status. When programs become inactive, the National Board changes their code from a 100 to a 300 series (for example, program 111 becomes program 311). However, the tape data was based on the status of the program at the time the examinee took the examination. Therefore, there were duplicate listings (100 and 300 series) for some programs.

The investigator wanted to remove these duplicates by converting those inactive programs listed as 100 series to 300 series. To determine whether this would have any effect on the analysis (especially examination scores), t-tests between examination scores for each of these programs were performed. There were no significant differences in scores for any of the 25 programs involved, so the duplicates were deleted from the computer file.

NQE status. The NBME tape provided information about the NQE status of each examinee, that is, whether they were being tested for the first time or were repeating the examination. However, because of concerns about confidentiality, the NBME did not indicate the numbers of repeats of those being retested. That is, there was no identification of multiple-repeaters.

Because the investigator wanted to identify the multiple-repeaters for purposes of analysis, repeaters were sorted by their birthdates and the results printed. After sorting birthdates, the investigator matched repeaters on other variables (by inspection of the printout). In this way it was possible to determine entries for each multiple-repeater and to track their performance on successive examinations.

Collected Data

As described in the methodology chapter, the investigator collected

data from the National Board's program files to create a computer file on characteristics of examinees' educational programs.

Educational level. Each program's educational level was determined by file information and a directory of expanded role programs for nurse practitioners (DHHS, 1982). Some program sponsors operated both certificate and masters level programs, and in most cases, the National Board had two separate program codes. However, three institutions that had both types of programs had only one program code. Another four programs had progressed from certificate to masters level over the years, but retained one program code.

Since all programs were classified as either certificate or masters level for analysis, decisions needed to be made about these seven programs. To determine classification, the investigator looked at the percentage of the program's graduates with masters degrees. If less than 50% of the graduates had masters degrees, the program was classified as certificate level. On the other hand, if more than 50% of graduates had masters degrees, the program was classified as masters level. (It was recognized that those certificate program graduates who entered their programs with masters degrees might distort these classifications. However, in six of seven cases, the proportions far exceeded the 50% criterion, so there was no question of distortion. In the seventh case, the program had been masters level since 1976, and the investigator felt safe in assuming that the masters prepared graduates were actually masters level nurse practitioner program graduates.)

As a result of this process, the three sponsors with both types of programs were classified as certificate level. Likewise, three of four programs that had progressed from certificate to masters were

classified as certificate level. The remaining program was classified as masters level.

Accreditation status. The supplementary sources used to determine accreditation status of programs were both published in 1982 (ANA, 1982; NLN, 1982). The variable coding, therefore, reflects the accreditation status of a particular program in 1982, and not necessarily when the examinee attended the program or took the examination.

Length in hours. There were more missing values for this variable than for any other program characteristic (missing data on 52 programs and 1,230 individuals). Some of this is accounted for by the discontinued programs, on which little or no information was available. However, there were other problems with the data available for this variable.

Some program directors provided information on the length of their program in terms of number of semesters or quarters, without any breakdown by hours. Others listed semester or quarter hours without converting to clock hours or providing the information necessary for conversion. Because the investigator could not accurately determine the clock hours, these cases were treated as missing values.

Appendix B. Administrative Data Collection

Since any research effort is only as good as the data available, it is appropriate to make suggestions for future collection of data by the National Board. The investigator does believe that it is important to collect certain baseline data about certification candidates, for administrative and research purposes. Besides the sociodemographic information that the National Board currently collects, it would be useful to have more information about the applicant's nurse practitioner educational program. For example: name of the program, educational level (certificate or masters), and dates of attendance.

In addition, changes in coding for the sociodemographic information need to be made. The investigator suggests that: birthdates be changed to age in years; provision by made for coding those examinees with no experience as RNs or PNPs; months of experience be changed to years of experience; and, current function be eliminated. Current function should be eliminated because there is redundancy between this information and employment setting. It could be assumed that those individuals employed in one of the eight PNP employment settings were functioning as nurse practitioners, and those in the other settings were not functioning as nurse practitioners.

Finally, it is recommended that institutional sponsors with both

certificate and masters level programs be identified by two separate program codes. Further, that those programs that have progressed from certificate to masters level be identified by different code numbers, so that there is no confusion about the educational level of an examinee's program.

Appendix C. National Board of Pediatric Nurse Practitioners and Associates: Letter of Agreement and Policies on Research, Publication, and Confidentiality

The National Board of
Pediatric Nurse Practitioners & Associates
414 Hungerford Drive, Suite 310, Rockville, Maryland 20850 [redacted]

American Academy of Pediatrics
Association of Faculties of Pediatric Nurse Associate/Practitioner Programs
National Association of Pediatric Nurse Associates and Practitioners

November 2, 1982

Barbara Hall Dunn, RN, MSN, CPNP
1243-A Gaskins Road
Richmond, Virginia

Dear Barbie,

I am writing to confirm the National Board of PNP/As' decision to assist you with obtaining NBPNP/A program data relevant to your dissertation research proposal. Since the Board shares your opinion of the need for research relevant to PNP certification examination performance and program preparation, they and I will attempt to facilitate your research needs within the purview of the NBPNP/A policy on research and publication (enclosed).

Of utmost concern to the Board is maintenance of security and confidentiality. For this reason, no data which identifies individuals to scores and/or to programs can be released. In addition, statistics for individuals who have specifically requested non-release of score and names will not appear in the data.

It is also the Board's understanding that the data released to you will be utilized for proposed research purposes only and will not be shared with other individuals and/or agencies/organizations. The Board also requests that upon the completion of your research, all data materials be returned to the Board.

To facilitate securing data relevant to your research efforts, I am requesting that you compile a list of specific data which will be required. I will then discuss with NBME and yourself the most efficient manner to accommodate your request.

Finally, the Board requests that you sign an informed consent agreement which outlines your willingness to comply with the Board's policy on research and publication and their specific requests as apply to your research project. This letter can serve as such an agreement.

I look forward to hearing from you in the near future and in assisting you with this much needed research project.

Sincerely yours,

[redacted]

Nancy Dickenson-Hazard, RN, CPNP, MSN
Executive Director
National Board of Pediatric Nurse
Practitioners and Associates

NDH/nc

I concur with the terms of the National Board of PNP/As in conducting research in accordance with NBPNP/A policy.

[redacted] 11/8/82

Signature

POLICY ON RESEARCH AND PUBLICATION

It is the policy of the National Board of Pediatric Nurse Practitioners and Associates to support research and publications which are consistent with the goals of the Board, i.e. quality child health care and demonstrated continuous competency to practice by pediatric nurse practitioners and associates providing said care.

To ensure these goals, individuals with research requests will follow the procedure as outlined below.

1. The researcher will request in writing the National Board data/information/participation which is to be utilized for the research project.
2. The researcher will provide a summary and/or overview of the research project.
3. The request and overview will be reviewed by no less than three (3) members of the Board and the Executive Director.
4. Expenses incurred by the Board for data/information/participation will be sustained by the researcher.
5. Acknowledgement to the Board for data/information/participation will be requested of the researcher.
6. Confidential materials and information of the National Board will not be released. The materials include the NQE or any portion of the test or items; data identifying candidate to score; data identifying candidate to program.
7. The National Board reserves the right to deny data/information/participation to any researcher who does not comply with Board procedure and goals.
8. The National Board reserves the right to review and/or co-author any potential publication which is the result of information/data secured from the Board.

POLICY ON CONFIDENTIALITY OF MATERIALS AND INFORMATION

It is the policy of the National Board of Pediatric Nurse Practitioners and Associates to provide a national certification examination and a series of self assessment exercises. To ensure confidentiality of materials and information, the following are considered "secure" by the National Board, and will not be released by the Board:

1. all copies of the National Qualifying Examination
2. individual test scores of all National Qualifying Examination candidates
3. individual self assessment exercise scores of all participating candidates
4. test scores of all candidates participating in the recertification examination.

Vita