A PSYCHOMETRIC ANALYSIS OF AN INSTRUMENT THAT TESTS GENETIC KNOWLEDGE OF ADVANCED PRACTICE NURSES IN GENETICS

A Dissertation

Presented to the Faculty of the

School of Nursing

Widener University

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Nursing Science

by

Jeanine T. Seguin

School of Nursing

April 9, 2007

Widener University School of Nursing

Title of Dissertation: A Psychometric Analysis of an Instrument that Tests Genetic Knowledge of Advanced Practice Nurses in Genetics

Author: Jeanine T. Seguin

Approved by:

[redacted]
Dr. Elizabeth W. Bayley, chair

[redacted]
Dr. Margaret A. Miller

[redacted]
Dr. Rita Black Monsen

[redacted]
Dr. Lois R. Allen

[redacted]
Dr. Judith Kilpatrick

Date: February 23, 2007

Submitted in partial fulfillment of the requirements for the degree of Doctor of Nursing Science

DEDICATION

To the memory of my grandparents

Carolyn Josephine Foster Benwitz

Chester Maximillus Benwitz

Doris Florence Wickens Buckman

Norman Leo Buckman

ACKNOWLEDGMENTS

I would like to thank my committee chair, Dr. Bette Bayley, and Drs. Peggy Miller and Rita Monsen for their patience, guidance, support and encouragement.

I would like to thank my friends and family. Mom, I could never have achieved this, or anything else, without your love and support. Joe, I know that you suffered along getting spoiled by Grandma while I spent my summers at school, but thanks for hanging in there for me. Michael, you signed on to this adventure with very little notice and stuck with us through to the end that, some days, we weren't sure was coming. I love you all.

I would also like to thank the Widener nursing faculty; Keuka College; the Institute for Credentialing Innovation of the American Nurses Credentialing Center; Omicron Chapter, Sigma Theta Tau, International; the International Society of Nurses in Genetics; the National Association of Clinical Nurse Specialists; the National Society of Genetic Counselors; the University of Cincinnati; the Emergency Nurses Association; the Academy of Medical-Surgical Nurses; the University of Maryland; and the Genetic Nurses Credentialing Commission for their support of me and my research.

Abstract

The purposes of this methodological psychometric study were to: 1) develop a criterion-referenced instrument using the steps outlined by Waltz, Strickland and Lenz (2005) to provide effective documentation of knowledge of advanced practice nursing in genetics; 2) determine the validity and reliability of the instrument; 3) compare pass/fail outcomes of nurses within and outside of the genetic specialty using the instrument; 4) compare pass/fail outcomes of Masters' prepared and non-Masters' prepared nurses using the instrument; and, 5) compare pass/fail outcomes of credentialed and non-credentialed advanced practice nurses in genetics using the instrument. An additional proposed aim, to evaluate congruence between instrument pass/fail outcomes and Genetic Nursing Credentialing Commission (GNCC) portfolio pass/fail outcomes, was not able to be fulfilled.

The Waltz and colleagues (2005) framework was used to develop a 100-item multiple-choice examination, based on published genetic nursing care standards, for the purpose of measuring knowledge of genetics of advanced practice nurses. The examination was based on criteria derived from the International Society of Nurses in Genetics, Inc. (ISONG) Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998). A panel of experts was used to establish content and construct validity and determine the cut-score for the instrument. The instrument was administered, via the Internet, to Masters' in nursing prepared nurses in genetics (n = 57), APNG certified advanced practice nurses in genetics (n = 5), non-Masters' in nursing prepared nurses in genetics (n = 33), non-genetic advanced practice nurses (n = 169) and non-genetic, non-Masters' in nursing nurses (n = 92).

The Criterion-Referenced Measurement Tool for Genetics (CRMTG) was analyzed for validity and reliability using a sample of 356 registered nurses. Reliability was evaluated using item-to-total and test-retest correlational analyses. Content and criterion-related validity were evaluated using content experts and pilot study data. Construct validity was evaluated using item analysis including level of difficulty and discrimination index calculations for each item. Divergent validity was evaluated using Chi square and ANOVA analyses to compare scores among groups of nurses with different educational preparations and different clinical practice areas.

The CRMTG was found to be valid with a content validity index of .88 and reliable with a Cronbach's alpha of .99 based on the findings of this study. The CRMTG pass rates indicated a significant difference between nurses who practice in a genetic setting and nurses who do not practice in a genetic setting ($p < .001$). The CRMTG was not found to be useful for differentiating between Masters' in nursing and non-Masters' in nursing prepared nurses when used for the sample as a whole. However, the pass rates for the subsample of participants who answered at least 85% of the items on the CRMTG did indicate a distinct difference between Masters' in nursing and non-Masters' in nursing prepared nurses. This study was unable to address portfolio pass/fail rates in relation to CRMTG pass/fail rates due to a very limited sample of nurses credentialed by GNCC through portfolio.

This study provides a model for the application and operationalization of nationally approved nursing practice scope and standards into a measurement tool, the application of both the Competency Outcomes and Performance Assessment Model (COPA) and Waltz' and colleagues' frameworks, and the use of the Internet for data

collection. This study also provides psychometric data for a tool that may be useful to the Genetic Nurses Credentialing Commission (GNCC) as a validation and/or supplement to their portfolio credentialing process.

This study was supported by the Institute for Credentialing Innovation of the American Nurses Credentialing Center; Omicron Chapter, Sigma Theta Tau International; and the Genetic Nurses Credentialing Commission.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

Page

CHAPTER ONE: INTRODUCTION

Chapter One begins with an overview of the study, including the problem statement, purpose, and research questions. Key terms are defined. This chapter also briefly introduces The Competency Outcomes and Performance Assessment (COPA) Model (Lenburg, 1979) and a framework for criterion-referenced measurement (Waltz, Strickland, & Lenz, 2005), which, together, provided a theoretical foundation for this methodologic study. Finally, Chapter One presents the assumptions and significance of this study.

Statement of the Problem

Credentials are a mechanism to protect and inform the public (Cary, 2001; Kelly & Joel, 2003; Mason, 2001). The credential has become a popular method of identifying beginning professional competence and advanced knowledge and skills in a variety of professions, including medicine and nursing, and at a variety of professional levels (Jensen & Saylor, 1994; Moloney & Schwirian, 1998; Pitts et al., 2002; Taylor, Thomas, & Sage, 1999; Weddle, Himburg, Collins, & Lewis, 2002; Wilkinson et al., 2002). The public acknowledgement of professional achievement provides a sense of professional pride and accomplishment. Credentialing of professional nurses is a method of professional self-regulation. However, the credentialing process is under scrutiny regarding its consistency, objectivity, validity, and reliability (Hood & Leddy, 2005; Knight & Knight, 1992; Whittaker, Carson, & Smolenski, 2000). This scrutiny applies to traditional means of awarding credentials through an examination process (Goeden, 1999; Hambleton, Swaminathan, & Rogers, 1991; Nichols, 1991) as well as to newer

approaches, such as portfolio review. Innovative methods of awarding and/or renewing credentials are emerging as ways of assuring professional competence of practitioners to the public but must be demonstrated to be valid (Alsop, 2001; Ben-David, 2000; Chambers, 2002; Greco & Mahon, 2003; Jasper, 2001).

As nursing care evolves through the 21st century, new knowledge and skills that have yet to be imagined will be required of practicing nurses and nurse educators. Instead of responding after-the-fact, nurses will be expected to "anticipate potential for alteration and address treatment while it is still a potentiality rather than an actuality" (Porter-O'Grady, 2001, p. 183). In order to ensure that nurses have current knowledge and skills in the increasingly complex health care arena, mechanisms are continuing to be developed to measure professional competence. Demonstration of professional competence is "the application of knowledge and the interpersonal, decision-making and psychomotor skills expected for the practice role, within the context of public health, safety and welfare" (National Council of State Boards of Nursing (NCSBN), 1998, p. 3) at the beginning level of entry to the profession and at advanced or expert levels.

Beginning professional competence is established through licensing examinations. Continuing professional competence can be established in job specific ways such as job performance evaluations and periodic peer and supervisor reviews. Professional competence can be established locally through relicensure requirements and registries. Professional competence can also be determined through recognized national and international organizations that specialize in the awarding of credentials to qualified

applicants (NCSBN, 1998). Many of these credentials are referred to as certificates and are seen as a sign of expertise and as a prestigious addition to one's professional standing.

Specialization as an advanced practice nurse is achieved through specific education requirements and a successful outcome of the credentialing process. Many professional organizations and most nursing organizations that award some form of credential do so through evaluating an applicant's performance on a standardized examination based on identified performance indicators. These examinations, typically multiple-choice, have been demonstrated to have some level of validity and reliability. Examinations are easy to administer to large numbers of candidates, simple to score, and the results are easily quantified. They are also expensive and labor intensive when development, maintenance, and revision are considered (Monsen, Cook, Middleton, & Kase, 2005).

Healthcare trends and escalating technology have led to an increase in medicine's subspecialization. Concurrent with the push for credentialing of nurses and other health professionals has been the development of increased specialization in nursing. For some nursing specialties small numbers preclude examination development for advanced practice credentialing; lower cost, alternative processes for the awarding of credentials are becoming more attractive.

One innovative approach to the credentialing process is the use of portfolios. Many different professions use portfolios for part or all of various competency and/or credentialing processes (Donen, 1998; Dyne, Strauss, & Rinnert, 2002; Routledge & Willson, 1997; Taylor et al., 1999). The portfolio is seen as a mechanism to provide real

practice information regarding performance decisions, educational endeavors, and critical self-reflection. The portfolio process is also seen as a more manageable strategy for the awarding of credentials than the creation and maintenance of a credentialing examination. Although also labor intensive, it is seen as a more cost effective approach to credentialing than the examination process, particularly for small organizations. However, little validity and reliability data are available on the portfolio as a means of evaluation.

This research sought to document the knowledge component of professional competence in nurses seeking the advanced practice nursing in genetics credential. An instrument was developed by the investigator to test knowledge of advanced practice nursing in genetics. Another intent of this study was to compare credential outcomes of this instrument with portfolio review outcomes; however, that purpose was not fulfilled due to small sample size and lack of the ability to identify nurses who failed the portfolio review.

Purposes of the Study

The purposes of this methodological psychometric study were to: 1) develop a criterion-referenced instrument using the steps outlined by Waltz, Strickland and Lenz (2005) to provide effective documentation of knowledge of advanced practice nursing in genetics; 2) determine the validity and reliability of the instrument; 3) compare pass/fail outcomes of nurses within and outside of the genetic specialty using the instrument; 4) compare pass/fail outcomes of Masters' prepared and non-Masters' prepared nurses using the instrument; and, 5) compare pass/fail outcomes of credentialed and non-credentialed advanced practice nurses in genetics using the instrument. An additional proposed aim, to

evaluate congruence between instrument pass/fail outcomes and Genetic Nursing Credentialing Commission (GNCC) portfolio pass/fail outcomes, was not able to be fulfilled. The investigator conducted this study to facilitate the evaluation of the current portfolio methodology used by the GNCC for awarding the advanced practice nurse in genetics credential.

Research Questions

The primary research question of interest was:

1) To what extent is the Criterion-Referenced Measurement Tool for Genetics (CRMTG) a valid and reliable indicator of knowledge of advanced practice nursing in genetics?

Other exploratory questions were:

2) Do nurses who practice in a genetic setting have a higher pass rate on the CRMTG than nurses who do not practice in a genetic setting?
3) Do advanced practice nurses have a higher pass rate on the CRMTG than nurses who do not have their Masters' degree in nursing?
4) Do nurses credentialed by the GNCC have a higher pass rate on the CRMTG than non-credentialed nurses?
5) Are GNCC portfolio pass/fail outcomes congruent with CRMTG pass/fail outcomes?

Definitions of Terms

Advanced Practice Nurse: This group includes Masters' prepared certified registered nurse anesthetists (CRNAs), certified nurse midwives (CNMs), clinical nurse specialists (CNSs), nurse practitioners (NPs) and case managers (Brown, 1996).

Advanced Practice Nurse in Genetics: A licensed registered nurse who has completed a graduate program in nursing, graduate coursework in genetics, and/or genetic clinical training and maintains current knowledge through continuing education provides advanced practice nursing in genetics. "Critical elements that distinguish advanced from basic level genetics nursing practice are the complexity of decision making, leadership, the ability to negotiate complex organizations, and expanded practice skills and knowledge in nursing and genetics" (ISONG, 1998, p. 6).

Credentialed Advanced Practice Nurse in Genetics (APNG): A registered professional nurse who has a minimum of a Masters' degree in nursing with additional genetic training who has been awarded the APNG credential by the GNCC.

Criterion-Referenced Measurement Tool for Genetics (CRMTG): This instrument is a 100-item multiple-choice test constructed by the investigator to reflect knowledge of advanced practice nursing in genetics based on the ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998).

Genetic Nursing: Genetic nursing is a clinical specialty that focuses on providing nursing care to "clients who have or are at risk for developing [inherited] conditions and/or birth defects, or who have children with [inherited] conditions and/or birth defects" (ISONG, 1998, p. 36).

Internet-based study: "Research that uses the Internet as the medium for data collection" (Siah, 2005, p. 117).

ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998): The International Society of Nurses in Genetics, Inc. (ISONG) in conjunction with the American Nurses Association (ANA) determined the scope and standards of genetics clinical nursing practice. This document delineates basic level and advanced level clinical genetic nursing practice standards. The published standards of genetics clinical nursing practice encompass the practice areas of: assessment, diagnosis, outcome identification, planning, implementation, and evaluation (ISONG, 1998). Implementation for all genetics nurses includes: problem identification, health teaching, case coordination, health promotion and health maintenance, psychosocial counseling, and genetic therapeutics. Standards for advanced practice nurses in genetics also encompass genetic counseling, case management, and consultation. The Scope and Standards document is currently being revised, but a final version is not available at the time of this writing. In this study the 1998 standards for advanced practice nurses in genetics were used as the criteria for the development of the measurement instrument content map.

Portfolio: A purposeful collection of artifacts that can serve as a portrait of individual and professional accomplishments (Weber, 2006).

Professional Competence: Demonstration of professional competence is "the application of knowledge and the interpersonal [and] decision-making...skills expected for the practice role, within the context of public health, safety and welfare" (NCSBN, 1998, p. 3). For this study the knowledge aspects of professional competence in genetics were

measured with the Criterion-Referenced Measurement Tool for Genetics (CRMTG) that was based on the content of the ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998).

Professional Credential: A professional credential is a designation used to protect the public from unsafe and incompetent providers, to provide the public with information to make informed choices about health care providers, to differentiate levels of care provider in a competitive environment, and to acknowledge excellence in practice (Cary, 2001; Hood & Leddy, 2005; Kelly & Joel, 1998; Mason, 2001; Moloney & Schwirian, 1998; Smolenski, 2005). This study focused on the Advanced Practice Nurse in Genetics (APNG) credential awarded by the Genetic Nurses Credentialing Commission.

Professional Standards: A professional standard is a threshold of best performance that can be used as a benchmark for comparison of practice behaviors and/or outcomes (Donabedian, 1981; Palmer, 1991). In this study the ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998) was used as the set of standards for performance in genetic nursing. These standards were the criteria upon which the CRMTG is based.

Theoretical Framework

The Competency Outcomes and Performance Assessment (COPA) Model (Lenburg, 1979, 1999a, 1999b, 2000, 2002; Luttrell, Lenburg, Scherubel, Jacob, & Koch, 1999; Nichols, Lenburg, & Soehnlen, 2000) was selected as the theoretical framework for this study because it provides a nursing context for performance evaluation. The model addresses competence through the discussion of the following four questions:

1. What are the essential competencies and outcomes for contemporary practice?
2. What are the indicators that define those competencies?
3. What are the most effective ways to learn those competencies?
4. What are the most effective ways to document that learners and/or practitioners have achieved the required competencies? (Lenburg, 1999a, 1999b; Redman et al., 1999).

The first three questions in the model address beginning or minimal levels of competence for entry into practice. The fourth question was the focus for this study from the perspective of minimal competence for advanced practice. An examination, the Criterion-Referenced Measurement Tool for Genetics (CRMTG), was developed to document one component of professional competence, the demonstration of knowledge of advanced practice nursing in genetics.

The criterion-referenced measurement work of Waltz, Strickland, and Lenz (2005) was selected for this study because it provides a framework for the determination of "what a person can or cannot do or knows or does not know, not on how the person compares to others" (p. 195). Criterion-referenced instruments are used to determine that the focus content is being measured. "Criterion-referenced measures are particularly useful when the purpose of testing is to ascertain whether an individual has attained minimum requirements" (p. 196). The CRMTG developed in this study evaluated the knowledge component of professional competence based on the criteria established in the ISONG Statement on the Scope and Standards of Genetics Clinical Nursing Practice

(1998). Both the COPA Model and the Waltz, Strickland and Lenz framework are discussed in greater detail in Chapter Two.

Assumptions of the Study

In this study the investigator made the following assumptions:

1. A criterion-referenced exam is able to measure the knowledge component of professional competence.
2. The ISONG Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998) accurately and appropriately reflects the advanced nursing practice specialty of genetics.
3. Valid and reliable criterion-referenced instruments can be developed based on professional standards and the currently accepted references and best practice in an area of specialization.
4. Testing conditions will vary as individuals would respond to the Internet survey in their home or work environments with varying environmental conditions and interruptions in the testing process.

Significance

The significance of the development of a valid and reliable instrument for the measurement of genetic nursing knowledge is that the availability of such an instrument is a necessary step in the process of evaluating and validating the current genetic nursing credentialing process. Certifying agencies have been challenged by the public to provide a valid and reliable mechanism for awarding credentials (Cary, 2000). The GNCC has an innovative method of competence assessment in the portfolio process; however,

evaluation of that process is needed. In order to effectively evaluate an innovative process, it should be compared to the established standard. In this case the standard for credentialing is a multiple-choice examination. Therefore, the CRMTG was developed and tested for validity and reliability in anticipation of future use as a psychometric measure of the validity of the GNCC portfolio process.

Nursing Practice

The development of a valid and reliable instrument for the evaluation of genetic nursing knowledge as a core element of professional competence supports the awarding of an APNG credential. In July 2003 the National Association of Clinical Nurse Specialists (NACNS) held a summit in Indianapolis, IN to discuss the recent recommendation from the National Council State Boards of Nursing (NCSBN) for licensure of advanced practice nurses via examination by 2005. It was noted that there currently are only 9 advanced practice nursing (APN) examinations although there are over 40 identified specialty practice areas for Clinical Nurse Specialists (CNS). Dr. Elizabeth Bayley attended the July 2003 NACNS Summit and reported that, based on the discussions held, "There is an urgent need to develop certification examinations where economically possible and demonstrate that alternative methods such as portfolio review are valid ways of measuring specialty competence" (E. Bayley, personal communication, July 31, 2003). The GNCC approach to credentialing through portfolio review was discussed. "Results of the psychometric data to be provided by this dissertation are eagerly anticipated by the American Nurses Credentialing Center (ANCC), the National Association of Clinical Nurse Specialists (NACNS), and many nursing specialty

organizations" (E. Bayley, personal communication, July 31, 2003) based on statements made by official representatives of these organizations.

Therefore, the significance for nursing practice of this study was to develop an instrument that can be used to validate the portfolio process for awarding of credentials. Once the process is demonstrated to be valid and reliable, other small organizations, both within and outside of nursing, may decide to establish sub-specialty credentials using a portfolio process when an examination is prohibitive due to the limited resources of the organization.

Nursing Science

The potential for the proposed study to advance nursing science is the application of the COPA Model to measure competency in the clinical practice of advanced practice nurses. The review of literature by the investigator yielded only undergraduate educational applications of the COPA Model, although the model was developed with both academia and practice settings in mind. The application of the model in advanced practice clinical arenas strengthens the function and scope of the model.

This study was also significant for the development of a new valid and reliable instrument to document advanced practice nursing knowledge of genetics. This instrument may be used to award credentials for genetic practice or validate existing genetic credentialing practice.

Therefore, the significance of this study to nursing science was that it modeled an application for the professional scope and standards documents as criteria for professional competence and developed an instrument to assess professional competence. This study

also added to nursing science by implementing the COPA Model in advanced nursing practice.

Nursing Education

The development of a criterion-referenced measurement instrument was significant for nursing education as a valid and reliable measure of knowledge of genetics. The CRMTG can be used to evaluate Masters'' level genetic nursing program outcomes. Therefore, the significance of this study to nursing education was to provide an instrument to measure genetic nursing knowledge in academic programs.

Nursing Research

The development of a valid and reliable instrument to measure knowledge of advanced practice nursing in genetics provides psychometric data needed to evaluate the use of portfolios for the awarding of credentials. This study constructed an instrument for testing the use of professional portfolios in nursing. Certification, outcomes evaluation and nursing education research are relatively new areas that need to be developed in nursing (Cary, 2000). Changes are sometimes made in nursing education and practice without planned, systematic evaluation of the effectiveness of the outcome (Anderson, 1997; Task Force on Accreditation of Health Professions Education, 1999). Portfolios have been used by nurses for a variety of purposes including evaluation of nursing education, clinical advancement, and academic tenure and promotion with anecdotal support. The Criterion-Referenced Measurement Tool for Genetics (CRMTG) as a measure of the nurse's knowledge base provides a mechanism for evaluating the existing portfolio review process used by the GNCC. Modeling a method of evaluation of

innovation in nursing education and practice was hoped to be a major contribution of this study. Although the sample size and distribution was not sufficient to provide psychometric data for the portfolio process, the procedure and instrument developed may be used as models for future research in portfolio evaluation.

The use of the Internet for data collection has an important role in nursing research. Nurse researchers have accepted the Internet as a means to gaining access to information; however, there is reluctance, in the health care community, to using the Internet for data collection (Daley, McDermott, Brown, & Kittleson, 2003; Strickland et al., 2003). This study used the Internet for the collection of data in the form of a multiple-choice instrument. Data were collected using an Internet data management program. The results were available and able to be downloaded as both raw data and in summarized form.

Therefore, the significance of this study to nursing research was as a model for evaluating nursing knowledge of genetics, development of an instrument for providing psychometric data on the portfolio process of credentialing genetic nurses and as a demonstration of a successful use of the Internet for the collection of data.

Summary

The awarding of professional credentials is a public acknowledgement of the achievement of a specific professional standard. The Genetic Nurses Credentialing Commission (GNCC) has developed a portfolio review process for the awarding of an advanced practice nurse in genetics (APNG) credential instead of the traditional method of examination for the determination of credentialing. This methodological psychometric study developed an instrument, the Criterion-Referenced Measurement Tool for Genetics

(CRMTG), that may be used to provide evidence of validity and reliability of the GNCC portfolio review process. Once the portfolio process for awarding of credentials is found to be valid and reliable other professional organizations may adopt a similar approach to credentialing. Key terms used in this study were defined and operationalized. The areas of significance of this study were identified. Discussion of the COPA Model was introduced and will be expanded in the second chapter as will discussion of the Waltz, Strickland, and Lenz (2005) framework for instrument construction.

CHAPTER TWO: REVIEW OF LITERATURE

Introduction

In this chapter the Competency Outcomes and Performance Assessment (COPA) model (Bargagliotti, Luttrell, & Lenburg, 1999; Cook, 1999; Lenburg, 1999a, 1999b; Redman, Lenburg, & Walker, 1999) are described and the four essential questions that are the framework of this model are explored in relation to this study. This chapter provides a synthesis of applicable research and literature on competencies including professional competence, the assessment of competence, credentialing and certification, a brief history of certification, and the American Nurses Credentialing Center. The concept of nursing competence as it relates to the specialty of genetic nursing is explored including existing and future roles of nurses in genetics and the portfolio assessment approach that the Genetic Nursing Credentialing Commission (GNCC) uses to award genetic nursing credentials. The documentation of competence and, specifically, competence in genetic nursing is included in this chapter.

Literature relative to the development of criterion-referenced instruments is synthesized including the linear model for test construction, the development and key components of a test plan, the development of quality test items and the analysis of test items for validity and reliability. The use of criterion-referenced instruments by nursing certifying organizations is discussed in this chapter. The stages for the development and validation of criterion-referenced measures as outlined by the Waltz, Strickland, and Lenz (2005) model are delineated. Also included in this chapter is a section on Internet data collection.

Academic Search Premier, Educational Resources Information Center (ERIC), Medline, Health source EbscoHost, and Cumulative Index to Nursing and Allied Health Literature (CINAHL) databases were accessed for the years 1975 – 2005. Manual and electronic searches were performed for the above mentioned areas as well as for the key words "the Human Genome Project (HGP)", "genetics", "genetic nursing", "credential", "credentialing", "credentialing" and "nursing", "certification" and "nursing", "competency", "evaluation", and "outcomes evaluation". A search was performed using Academic Search Premier, ERIC, Medline, and Health Source for the years 1975-2005 of "portfolio and competence", yielding 424 articles, nine of which were relevant to this study. An additional search was performed using Academic Search Premier, ERIC, Medline, and Health Source for the years 1975-2005 of "portfolio and credential", yielding 97 articles, one of which was relevant to this study; and "credentialing" which yielded 4,583 articles and was refined to "credentialing and nursing", yielding 866 articles, seven of which were relevant to this study. These searches were performed in order to evaluate current trends and standards in the literature. Results of these searches provided definitions of terms. The theoretical frameworks for this study were discovered and described through literature review. The literature also provided insight into potential problems which were evaluated, planned for, and minimized.

The Competency Outcomes and Performance Assessment Model (COPA)

Introduction

The Competency Outcomes and Performance Assessment (COPA) Model was developed in the 1970s by Dr. Carrie B. Lenburg. Changes in health care drove a need to

reform the evaluation of nursing competence in academic and clinical settings. The COPA model was designed to protect consumers, respond to technological innovations, accommodate sociopolitical and market forces, and adjust to the rising incidence of litigation related to health care (Lenburg, 1999a). Lenburg (1999a, 1999b) defined competency-based performance evaluation as a criterion-referenced, summative evaluation process that assesses a participant's actual ability to meet a predetermined set of performance standards under controlled conditions and protocols.

The COPA model has been applied to a variety of educational and clinical settings, including nursing and pharmacy programs (Horne & Medley, 2001; Lenburg, 2000). Luttrell, Lenburg, Scherubel, Jacob, and Koch (1999) described a Baccalaureate of Science in Nursing (BSN) curriculum that was successfully redesigned based on the COPA model. Nichols, Lenburg, and Soehnlen (2000) provided an example of the COPA model being applied to maximize prior learning and experience for registered nurses in a degree completion program.

COPA Model

Increased consumer concerns regarding the quality of health care and increased litigation of health care providers over the past several decades led Dr. Lenburg to develop The Competency Outcomes and Performance Assessment (COPA) Model (1979). She developed the COPA model based on her extensive work with the New York Regents College External Degree Nursing Program (NYREDP). The model was designed to protect consumers by assessing an individual's (student, faculty, nurse) ability to meet

a predetermined set of performance standards under controlled conditions and protocols (Lenburg, 2002).

The COPA Model was designed to evaluate initial and continuing competence in nursing practice. Continued competence is receiving increased emphasis for the entire United States workforce because employers, consumers, and third party payers expect a continually evolving and ever expanding skill set. Nursing faculty, students, and clinicians must be able to demonstrate their continued professional competence, growth, and currency in nursing and in their identified specialty(ies) (Lenburg, 1999a). The COPA Model provides a comprehensive, reflective, exploratory approach to assessing the full range of core competencies essential for nursing practice (Lenburg, 1999a; Redman et al., 1999).

Framework

Lenburg (1999a, 1999b) defines competence in terms of performance and the achievement of performance indicators. The basic organizing framework for the COPA model consists of four essential questions:

1. What are the essential competencies and outcomes for contemporary practice? ;
2. What are the indicators that define those competencies? ;
3. What are the most effective ways to learn those competencies?; and,
4. What are the most effective ways to document that learners and/or practitioners have achieved the required competencies? (Lenburg, 1999a; Lenburg, 1999b; Redman et al., 1999).

What are the essential competencies and outcomes for contemporary practice?

To answer this first guiding question one must identify the required competency outcomes. A standard needs to be established which identifies what authorities consider to be competence in any given situation. A professional consensus is required to establish realistic, appropriate, measurable outcome statements for the group being assessed. Lenburg (1999a, 1999b) has identified eight core practice competencies for graduate professional nurses and examples of sub skills for each (Table 1).

Table 1

Lenburg's Eight Core Practice Competences (Adapted from Lenburg, 1999a, 1999b)

1. Assessment and Intervention Skills (e.g. safety and protection, assessment and monitoring, therapeutic treatments and procedures)
2. Communication Skills (e.g. oral skills, writing skills, computing skills)
3. Critical Thinking Skills (e.g. evaluation, problem solving, decision making, scientific inquiry)
4. Human Caring and Relationship Skills (e.g. morality, ethics, legality, cultural respect, cooperative interpersonal relationships, client advocacy)
5. Management Skills (e.g. administration, organization, coordination, planning, human and material resource utilization, accountability and responsibility)
6. Leadership Skills (e.g. collaboration, creativity, anticipating, supporting with evidence, professional accountability)
7. Teaching Skills (e.g. individuals and groups, health promotion, health restoration)
8. Knowledge Integration Skills (e.g. liberal arts, natural and social sciences, nursing, healthcare and related disciplines)

What are the indicators that define those competencies?

This question requires the identification of behavioral indicators that are deemed mandatory for actual practice of each competence. Lenburg (1999b) defines these critical elements as "the set of single, discrete, observable behaviors that are mandatory for the designated skill, at the targeted level of practice" (p. 10). Again, the model uses the eight core practice competencies identified in Table 1 as reference points for these behavioral indicators.

What are the most effective ways to learn those competencies?

This question reflects the academic application of the COPA model. It allows for reflection on the most effective methods for learners, and practitioners, to accomplish the competencies (Lenburg, 1999b). When the COPA model is applied in educational settings, the most appropriate teaching methodology and the determination of effective learning must be evaluated for each situation.

What are the most effective ways to document that learners and/or practitioners have achieved the required competencies?

The fourth question in the model reflects how critical the evaluation of knowledge is to evidence-based practice. Integration of a broad range of knowledge is key to the achievement and maintenance of competence. Mechanisms of evaluation, supported documentation, and assurances to the public are necessary for a legitimate competence validation program.

Lenburg (1999b) identifies two essential psychometric concepts; examination and sampling. Lenburg has developed two competency assessment instruments to measure

these concepts. The Competency Performance Examinations (CPEs) evaluate the student's knowledge using a written examination. The Competency Performance Assessments (CPAs) evaluate a sample of the student's skills. The former is applicable in academic settings and the latter is used in clinical settings (Lenburg, 1999b; Bargagliotti et at., 1999).

The key elements of the COPA model are the learning, establishment, and documentation of professional competence. Learning of competence is an on-going achievement. The documentation of competence is also fluid. Behaviors based on standardized indicators are used to establish competence (Figure 1).

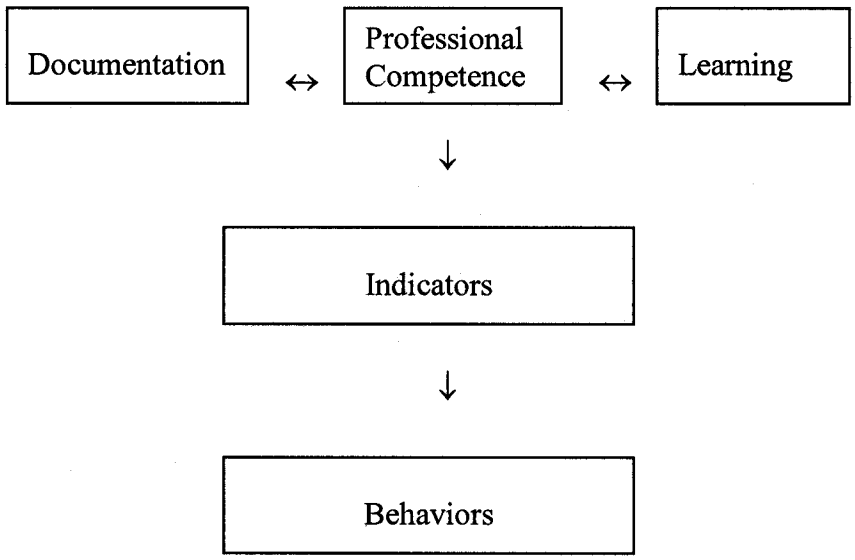


Figure 1. Relationships in The COPA Model

Applications

The COPA framework and methods are applicable for documenting either initial or continuing professional competence. The model was designed for academic and clinical practice environments. It can apply to "newly licensed nurses or to certified

nurses who need to confirm continuing abilities to engage in competent and contemporary practice" (Lenburg, 1999a, p. 11). The model has, primarily, been implemented in academic settings through the use of CPEs and CPAs to supplement traditional grading practices. These academic settings include the New York Regents College External Degree Nursing Program, as mentioned above, the University of Memphis, the University of Colorado Health Science Center, and King College in Bristol, TN. Review of the literature including Academic Search Premier, ERIC, Medline, EbscoHost and CINAHL databases by the author found no exemplars of the COPA model being studied outside of the education setting.

Summary of COPA Model

The Competency Outcomes and Performance Assessment (COPA) Model can be used to establish and evaluate a level of professional competence in nursing practice (Figure 1). This model has four essential questions as its framework. Each question reflects an understanding and application of the eight core practice competencies identified by Lenburg (1999a, 1999b, 2000) as essential to nursing clinical practice. Although the model is directed toward baccalaureate nursing education, it can be readily adapted to specialty and advanced nursing practice, by making the indicators of professional competence more specific and of a higher performance level, respectively.

Competencies and Outcomes

Introduction

An important concept in the COPA model is essential competencies and outcomes for contemporary practice (Lenburg, 1999a, 1999b, 2000; Redman et al., 1999). The

literature was reviewed to identify the meaning, context, and determination of professional competence and, subsequently, credentialing. The following section will discuss the concept of professional competence and how it is assessed. Credentialing and certification will be compared and contrasted including a discussion of the history of certification and the establishment of credentialing organizations including the American Nurses Credentialing Center.

Professional Competence

Competence was defined by the National Council of State Boards of Nursing (NCSBN) as "the application of knowledge and the interpersonal, decision-making and psychomotor skills expected for the practice role, within the context of public health, safety and welfare" (1998, p. 3). Alsop (2001) made the point that graduation from an accredited program signifies that a practitioner is qualified to practice the science of nursing. Competence, however, is a new dimension that encompasses the art of nursing (Alsop). Lenburg (2000) argued that having a degree, license and/or certificate, all recognized indicators of qualification for a position, are not enough to stay current, thus competent, in the field of nursing. Even fulfilling continuing education requirements does not ensure professional competence (Lenburg, 2000). Girot (2000) also discussed the difference between qualified and competent. Girot addressed the different levels of nursing preparation available in the United Kingdom, which are similar to those in the United States; therefore, the entry into practice issues are similar. The concerns presented relate to different levels of basic preparation, including diploma, associate, and baccalaureate degrees with similar competency expectations for graduate nurses (Girot).

The term "competency" is also found in the literature in relation to credit for prior learning documentation (Neely & Schuley, 1980), but that use will not be explored in this study.

Professional competence includes clinical reasoning and principled, intuitive practice (Diekelmann & Ironside, 2005). "Regulatory systems need to acknowledge the development of critical and transferable skills and not just an evolving knowledge base" (Alsop, 2001, p. 128). Bargagliotti et al. (1999) define competency as an evaluation of clinical abilities and state that "all practicing nurses are ... competent by reason of licensure, employment and perhaps professional certification" (p. 1). Competence, according to Alsop is socially situated, job referenced, and binary. A competent practitioner not only accepts, but also responds to and shapes change (Alsop, 201; The Center for the Health Professions, 2002; Hawk, 1999). Competence measured against defined outcomes "derives from public sentiment that how something was done is not as important as whether or not it was effective" (Bargagliotti et al., 1999, p. 2). Blumenreich (2002) stated, "society is dependent on healthcare professionals to evaluate the competence of their fellow practitioners" (p. 348).

Assessment of Professional Competence

Assessment of professional competence can be accomplished via a variety of methods including testing, skill demonstration, peer review, certification, critical thinking skills tests, or professional portfolio (Green & Ogden, 1999). Credentials are awarded based on the establishment of professional competence through a credentialing body. Certification is a form of professional credential (Figure 2). "Utilizing certification as a

measure of continued competence is one approach that can take into account the large variety of professional nursing roles and specialties" (ANA, 1996, p. 14).

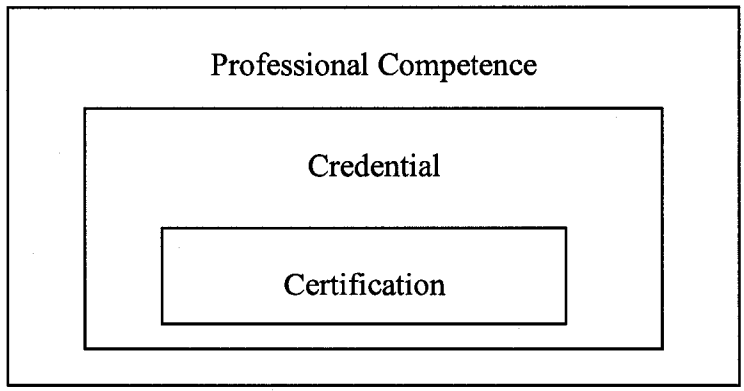


Figure 2. Interrelationship among professional competence, credential, and certification

The awarding of a credential or a certificate provides the public with an assurance that nurses are competent in their area of practice (Cary & Smolenski, 2005; Scherubel, 2002). The degree of confidence depends on the validity and reliability of the credentialing process, thus the prevalence of standardized tests for the awarding of credentials. Validity and reliability data through item-analysis is available for certification exams such as those administered by the ANCC. However, validity and reliability data are scarce for alternative methods for the awarding of credentials (Cary & Smolenski, 2005; Murrells, 2002) such as skills assessment and peer or portfolio review.

Credentialing

Credentialing is used in the literature as a general term meaning a mechanism to protect the public from unsafe and incompetent providers, provide the public with information to make informed choices in health care providers, differentiate levels of care provider in a competitive environment, and acknowledge excellence in practice (Cary,

2001; Hood & Leddy, 2005; Kelly & Joel, 2003; Mason, 2001; Moloney & Schwirian, 1998). Credentialing is described as public or private and "incorporates licensure, certification, accreditation, recognition, and registration" (Smolenski & Gagan, 2005, p.201). Public credentialing in the United States is referred to as licensure. A license is issued by a government organization and allows an individual to practice within the practice act of the specified profession.

Private credentialing is informational and voluntary. Competence of health care professionals is assured to the public through private credentialing in such disciplines as medicine (Ben-David, 2000; Chambers, 2002; Donen, 1998; Pitts et al., 2002; Routledge & Wilson, 1997; Wilkinson et al., 2002), occupational therapy (Alsop, 2001), physical therapy (Jensen & Saylor, 1994), social work (Taylor, Thomas, & Sage, 1999), dietetics (Weddle, Himburg, Collins, & Lewis, 2002) and nursing (Cary, 2001; Hood & Leddy, 2005; Kelly & Joel, 1999; Mason, 2001; Moloney & Schwirian, 1998). Credentialing in health professions can mean recognition by a professional organization (Shirey, 2005) or the granting of privileges to practice in a given setting based on an internal review at the institution in which privileges are being sought (Blumenreich, 2002).

In response to variability and incongruity across the credentialing spectrum the United States Department of Health, Education, and Welfare (DHEW) investigated the feasibility of a voluntary national certification system of licensed professionals in 1971 (Kelly & Joel, 2003). National certification was not adopted, but the National Commission for Certifying Agencies (NCCA) was established in 1978 in order to certify those organizations granting any professional certification (Kelly & Joel). This type of

organization is considered second-tier as it certifies the certifiers. The National Board of Nursing Specialties (NBNS) is another second-tier credentialing organization that was formed in 1990 (Kelly & Joel). Only nursing credentialing organizations are monitored by the NBNS (Kelly & Joel).

In nursing, credentialing is defined by the International Council of Nurses (ICN) as "processes used to designate that an individual, program, institution or product has met established standards set by an agent (governmental or nongovernmental) recognized as qualified to carry out this task" (Styles & Affara, 1998, p. 44). Terminology used to refer to or in conjunction with credentials includes licensure, registration, accreditation, approval, certification, recognition, or endorsement (Styles & Affara).

Certification

Certification is the term most commonly used when referring to specialized nursing practice credentials in the United States. Many organizations offer credentials for nursing specialties. Large nursing organizations such as the American Nurses Credentialing Commission and smaller nursing organizations such as the American Holistic Nurses' Association offer a variety of options for nurses seeking professional recognition. Kelly and Joel (1996) defined certification as:

> The process by which a nongovernmental agency or association grants recognition to an individual who has met certain predetermined qualifications specified by that agency or association. Such qualifications may include (1) graduation from an accredited or approved program; (2) acceptable performance on a qualifying

examination or series of examinations; and/or (3) completion of a given amount of work experience (p. 443).

In 1999, Kelly and Joel refined their definition of certification to: "a voluntary, nongovernmental credential awarded to individuals after they prove their ability to experts or peers in their field. It signified competence or specialization, and not just safe practice" (p. 453). Moloney and Schwirian (1998) defined certification as a means of recognizing professional achievement through excellence in clinical practice and/or advanced specialty practice. Certification motivated nurses to improve and maintain their skills, which improved patient care outcomes, and increased the autonomy of the nursing profession (Moloney & Schwirian; Alabama State Nurses' Association, 2000; Cary, 2000; Roberts, 2001).

History of Certification in Nursing

The certification of nurses began in 1946 with the certification of nurse anesthetists by the American Association of Nurse Anesthetists (Fairman & Lynaugh, 1998; Fickeissen, 1990). In 1966 the American Nurses Association (ANA) began to plan mechanisms for the advancement of professional clinical competence. As early as 1968 expert nurses wrote standards of practice for the ANA, and certification was based on the achievement of these established standards of practice (Dolan, 1968).

The American College of Nurse-Midwives was the first organization to credential nurses by written examination in 1971 (Fairman & Lynaugh, 1998). The American Association of Critical-Care Nurses (AACN) and the National Board of Pediatric Nurse

Practitioners and Associates (NBPNPA) began awarding credentials in 1976 and 1977, respectively (Fairman & Lynaugh).

The American Nurses Credentialing Center

Currently the largest, based on number of nurses credentialed and number of specialty credentials available, credential granting organization in nursing is the American Nurses Credentialing Center (ANCC) (Scherubel, 2002). Although the American Nurses Association (ANA) began certification of nurses in 1973, it was not until 1991 that the ANCC was established as a central organization for the certification of nurses (Scherubel). Also in 1991, the American Board of Nursing Specialties (ABNS) was established as a national peer review program, and it monitored the variations in educational, practice and experience requirements among specialty nursing certification groups (Parker, 1994). By 1995 there were 26 areas of certification (Kelly & Joel, 1996) and in 1999 there were 43 nationally identified certifications (Navigating, 1999). "Almost all graduates of nursing Masters' specialty programs (especially practitioner programs), want the status afforded by certification" (Barnum, 1997, p. 4). The benefits of certification to the individual nurse include improved marketability, greater professional mobility, monetary reward, and peer recognition (ASNA, 2000; Cary, 2000; Moloney & Schwirian, 1998; Roberts, 2001).

Currently, approximately 16% of all registered nurses have become certified (Mason, 2001). The ANCC, alone, offers 37 specialty and advanced practice certification exams. There are over 135,000 nurses certified by ANCC, of which, about 58,000 are advanced practice nurses (ANCC, 2006). The certification process is traditionally

performed through the successful completion of an appropriate educational program and the passing of a standardized examination (Barnum, 1997; Nichols, 1991; Smolenski & Gagan, 2005).

Summary of Competencies and Outcomes

While the public credential (licensure) is required to practice, private credentials and certificates awarded by professional organizations have become popular methods of identifying competence at a variety of professional levels. The public acknowledgement of professional achievement also provides the nurse with a sense of professional pride and accomplishment. Traditionally, credentials have been awarded through an examination process. Innovative methods of awarding and/or renewing credentials are emerging as ways of assuring professional competence of practitioners to the public; however, a review of the literature did not provide any validity and reliability studies of these emerging methods.

Nursing Competence in Genetics

All nurses need at least a basic knowledge of genetics (Greco, 2003; International Society of Nurses in Genetics, 1998; Jenkins, Dimond, & Steinberg, 2001; Lea, Anderson, & Monsen, 1998; Pew Health Professions Commission, 1995). Monsen (1999), founding president and past executive director of the Genetic Nursing Credentialing Commission (GNCC), stated that "there are over 2.5 million RNs in the US, and nearly all can expect to be confronted with questions about advances in genetics as growing coverage in the popular media focuses on gene-based health care" (p. 4). "Experts say it is as important for clinicians to be conversant with genetics as with

pharmacology and physiology" (Saunders, 1998, p. 39). Collins (2002), director of the National Human Genome Research Institute at National Institutes of Health, predicted that genetic advances would affect every nurse in every setting within the next ten years. Nurses have clinical access to people with genetic conditions. Nurses act as the primary interface between health care and the public. As primary contacts, nurses will be integral in identifying those at risk for genetic conditions and providing education regarding diagnostics and treatment options. Nurses may be asked to explain what the defects mean and how the mutation is transmitted (Carol, 2003; Conley & Gorin, 2003; Jenkins & Collins, 2003; Lea & Monsen, 2003; Loescher & Merkle, 2005).

Nurses provide basic counseling of individuals and families and refer them to genetic counselors as appropriate. Nurses are the educators, policy makers, and researchers for the genomic era (Collins, 2002; Feetham, Thomson, & Hinsaw, 2005; Olsen, et al., 2003; Prows, Glass, Nicol, Skirton, & Williams, 2005). Nurses will monitor ethics and health care access and delivery as advocates for the public (Biedrzycki, 2002; Hegyvary, 2005; Jenkins, Grady, & Collins, 2005; Jenkins & Lea, 2005). The mapping of the human genome has opened up genetics technology and information to nearly every person and group in the world with access to the Internet. All nurses, as providers of care, must have a basic knowledge of genetics (Genomics Policy Unit, & Medical Genetics Service for Wales, 2003; Jenkins, Calzone, Lea, & Prows, 2005).

Historically, genetic information was included in basic nursing preparation in courses related to maternity and pediatrics. Genetic conditions are now found to be significant across the lifespan. Medical advances have improved care outcomes for

children with genetic syndromes to the extent that these children are living into adulthood and, in some cases, even having children of their own (Carson & Hieber, 2001). Pharmacogenomics, the customization of medications based on cellular dysfunction, metabolism, or enzymes, and oncogenomics, the study of genetic alterations that allow the replication of tumor cells, are now changing the way that many adult illnesses are being diagnosed and treated (Henig, 2004; Lashley, 2005; Lewis, 2005; Munoz & Hilgenberg, 2005; Nicol, 2003; Persing & Cheek, 2000). Genetic nursing encompasses healthcare across the lifespan, from family planning and prenatal testing, through childhood and adulthood. Genetics plays a role in how people live and how they will die.

Genetics has also entered into the arena of nursing clinical specialties. Specialized advanced practice nurses in genetics help disseminate genetic information, monitor genetic legislation, and educate their peers and the public. Consumers expect to receive accurate information from competent, knowledgeable care providers such as nurses who meet or exceed standard care practices (Cary & Smolenski, 2005; Diekelmann & Ironside, 2005; Prows et al., 2003). The establishment of a visible credential demonstrating professional competence in the rapidly advancing field of genetics in health care is essential (Skirton, Patch, & Williams, 2005).

Portfolio Assessment

The investigator has been a part of the Genetic Nurses Credentialing Commission (GNCC) since it was chartered in 2001. Prior to that an ad hoc credentialing committee was formed in 1999 and charged by the International Society of Nurses in Genetics, Inc. (ISONG) with developing a process to award a credential that recognized knowledge and

performance of both basic and advanced practice nurses in genetics. The term credential was selected because a full-fledged certification process was not in place. The committee corresponded with the ANCC regarding the development of a standardized certification exam. Finding the exam process too expensive and the number of eligible genetic nurses too small, a portfolio process was established for the awarding of credentials in genetics (Moyer, 2002). Portfolios had begun to be used for the documentation of initial and continued professional competence in a variety of settings (Dennison, 2005; Friedman & Marr, 1995; Hayes, Chandler, Merriam, & King, 2002; Johnson, 2002; Lenburg, 2000; Seguin, 2005; Weber, 2006; Weddle, Himburg, Collins, & Lewis, 2002).

The GNCC score team currently uses the International Society of Nurses in Genetics (ISONG) and the American Nurses Association (ANA) Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998) to evaluate portfolios. Standards are defined as a threshold of best performance that can be used as a benchmark for comparison (Donabedian, 1981; Palmer, 1991). Clinical logs and case studies are included in the portfolios as indicators of performance. An overall agreement among reviewers of acceptability and evidence of demonstration of standards is required for award of the credential. Each portfolio is reviewed by at least 5 reviewers (GNCC, 2001). The score team is comprised of clinical and academic genetic nurses who have been trained in the portfolio review process being used by GNCC.

Summary of Nursing Competence in Genetics

As the need for all nurses to have basic genetic knowledge unfolds, a role is developing for the expert nurses in genetics. In order to evaluate genetic nursing

competency using the COPA model, performance indicators must be developed. The GNCC adapted the ISONG Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998) and used these standards as indicators of competent advanced practice nursing care in genetics. These standards have been adapted by the GNCC into a portfolio process for the awarding of an advanced practice nurse in genetics (APNG) credential. The portfolio process examines evidence of the candidate's professional practice and compares the evidence to the published standards as indicators of the individual's advanced practice nursing in genetics clinical competence.

Documentation of Competence

Introduction

This study focused primarily on the fourth essential concept in the COPA framework, documentation of achievement. The achievement of competence is both a professional requirement and a personal development responsibility (Lenburg, 2000). Documentation of this achievement can include documents such as job performance evaluations from employers, professional credentials from accreditors, peer evaluations from coworkers, documentation of continuing education, goal setting, focused self-reflection, and examination (Johnson, 2002; Lenburg; Weddle et al., 2002). Additionally, credentials have been awarded in health care, and particularly in nursing through an examination process (Nichols, 1991; Barnum, 1997). However, the development, maintenance and revision of a large pool of items are quite expensive and labor intensive. Also, in order to establish item validity and reliability, a large pool of candidates must take the examination. Standardized examinations are under fire for their inability to

measure critical thinking, creativity, and real-life behavior measurement. Standardized examinations may measure more test-taking skill than nursing skill (Goeden, 1999; Hambleton, Swaminathan, & Rogers, 1991; Nichols, 1991).

Documentation of Competence in Genetic Nursing

The GNCC, while investigating a certification exam for nurses specializing in genetics, discovered significant feasibility issues. In 1999, Dr. Mary Smolenski, Director of Certification Services for the American Nurses Credentialing Center (ANCC), estimated that a genetic certification exam through ANCC would cost about $40,000 to develop and would require a critical mass of at least 1,000 applicants (Moyer, 2002). ISONG, the primary source of potential applicants, has only 250 members and a small budget. Also of concern were the multiple sub-specialties within genetics nursing and the rapidly changing state of nursing practice related to genetics.

Founding members of GNCC then explored alternative methods of competency documentation. The portfolio process was selected by GNCC because it reflects individual practice, responds to changing professional environments, and provides an equivalent evaluation standard for the applicants regardless of the variations of responsibility and evaluation present in their own work environments. Dr. Rita Monsen, founding president and past executive director of GNCC, and Dr. Smolenski have identified that many external groups are carefully watching the progress of the portfolio process with the potential for a more widespread application in the certification/ recertification arena (R. Monsen, personal communications, June 11, 2002, June 25, 2002, July 1, 2002, July 4, 2002; M. Smolenski, personal communication, July 1, 2002).

Of greatest concern is the lack of psychometric data in the literature surrounding the portfolio process for evaluation. The aim of this study was to create a valid and reliable knowledge assessment instrument that may be used to serve as a foundation for validating the use of portfolios for awarding credentials. This instrument fills a void by providing a means to compare exam-based versus portfolio-based credentialing with GNCC as the example organization and APNG as the model group.

Development of Criterion-Referenced Instruments

Introduction

Measurement of nursing competence for the issuance of a degree or license or the awarding of a credential has traditionally been achieved through standardized testing. The purpose of the measurement is to quantify the characteristics of a competent professional nurse. Measurement can be obtained through direct or indirect approaches. Considerations when developing a measurement instrument include accuracy, precision, sensitivity, and error (Burns & Grove, 2005; Munro, 2004; Nunnally & Bernstein, 1994). Measurement instrument development includes multiple step-wise activities.

The Linear Model for Test Construction

The evaluation of humans, when an instrument, as a whole, measures an attribute, utilizes the linear model. The use of this model assumes that the sum of items scored has a linear relationship with the attribute being measured; when the attribute increases, the corresponding score increases. It is also expected that the sum of the variables being measured has a linear relationship with the sum of the items scored. The test score is a

sum of the items scored. Items may be weighted based on an established guideline or rubric that is congruent with the attributes to be measured (Nunnally, 1970).

Test Plan

The most critical method of ensuring the validity of the test is to use a test plan prior to developing the test, rather than measuring the validity after the test has been constructed (Nunnally, 1970). The test plan defines "the scope and emphasis of the test" (Tinkelman, 1971, p. 49) and directly relates to the purpose of the test. The test plan outlines the purpose and content for the test, describes the type and number of items by section and subsection, provide examples of the items, describe how long the test should take to administer, how it will be administered, and how it will be scored (Henrysson, 1971; Nunnally, 1970). The number of items to be included on a test should reflect the weighting of the sections and subsections as identified on the test plan (Tinkelman, 1971). Upon completion, the plan should be reviewed and approved by content experts (Nunnally, 1970; Tinkelman, 1971).

Number of Test Items

There should be a sufficient number of items to establish reliability. "If the cutoff score is to be used to award jobs or grant scholarships, reliable measurements may be critical indeed" (Tinkelman, 1971, p. 71) requiring a reliability of at least .90. Using the Spearman-Brown formula, if 100 items are included on a test with an expected average correlation between items and test score of .50 the reliability of the test is calculated to be .97. If 150 items are included, which is a standard length of certification examinations (AACN, 2004; ANCC, 2004), the reliability of the test is calculated to be .98. However,

for criterion-referenced tests reliability may be difficult to calculate, or even inappropriate, due to the low degree of variance of test scores (Glaser & Nitko, 1971).

The number of items should be reasonable for the length of time allocated for testing (Nunnally, 1970; Tinkelman, 1971). Multiple-choice items typically take between 45 and 60 seconds per item (NCSBN, 2003; Tinkelman, 1971). Therefore, an 100-item test should be scheduled for a 1 1/2 hour time slot and an 150-item test should be planned to last 2 hours. Pilot studies will also provide time allocation information based on the average length of time the pilot participants require to complete the test. Tests should be constructed to allow 75 to 90 % of the examinees to finish the test within the time allocated (Tinkelman).

Criterion-Referenced Measurement

Criterion-referenced tests should be directed toward measuring achievement expressed in terms of performance. This form of testing requires a clearly defined and delimited domain, clearly defined behavioral objectives, clearly specified standards of performance, adequate sampling within each area of performance, selection of test items reflecting specified behaviors, and a scoring and reporting system that adequately describes performance on clearly defined tasks (Gronlund, 1973; Schwarz, 1971; Waltz et al., 2005). In clinical genetic nursing, the standards of advanced practice performance are delineated in the ISONG Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998).

Types of Test Items

Test items can be objective, such as multiple-choice items, true-false items, and matching items. Test items can also be of a constructed response type, such as completion items, essays, and problem-solving items (Center for Nursing Education and Testing, 1998). Objective-type items have the advantages of: (a) providing a wide sampling of content; (b) being able, if constructed properly, to measure multiple levels of cognitive ability; and (c) efficient, accurate, reliable and objective scoring. The disadvantages of objective-type items are that they are difficult to develop, particularly at higher cognitive levels, and that the test taker has a chance to guess the correct answer, without knowing the content (Center for Nursing Education and Testing, 1998; Jenkins & Michael, 1986; Stanton, 1983). Constructed response-type test items are advantageous because they are relatively quicker and easier to develop than the objective-type items; they can, as with objective-type, also test a wide sampling of content; guessing is minimized; and with essay and problem-solving items, one can efficiently measure higher-level cognitive objectives, such as analysis and synthesis. Disadvantages of constructed response items are that they are more difficult and time-consuming to score, are less reliable, and are more subjective than objective-type items (Center for Nursing Education and Testing, 1998; Stanton, 1983).

Multiple-choice test items

Multiple-choice items are the best objective measure and are the most widely used test type for commercially distributed tests (Nunnally, 1970; Wesman, 1971). They are the easiest style of test to administer and score, there is a large pool of expert multiple-

choice item writers, and most any content area can be evaluated using multiple-choice tests. Multiple-choice tests are more reliable than other styles of tests, such as essay, and are able to test a broader spectrum of a topic (Nunnally, 1970). Multiple-choice items are typically arranged in a spatial pattern that is uniform and easily differentiates stem from choices (Thorndike, 1971).

One drawback of multiple-choice items is that they are primarily able to measure only knowledge, comprehension, and application (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956; Demetrulias & McCubbin, 1982; Frisbie, 1983; Hayter, 1983; Waltz, et al., 2005). Another limitation is the risk to reliability of multiple-choice tests due to guessing (Henrysson, 1971). Increasing the number of alternative responses per item can reduce this threat. The correlation between reliability and number of alternative responses was found to be increasingly significant from two to five alternatives; above five, the significance decreased (Guilford, 1954; Wesman, 1971).

Bloom's Taxonomy

Bloom and associates (1956) identified six cognitive levels of evaluation: knowledge, comprehension, application, analysis, synthesis, and evaluation. The awareness of facts, terms, concepts and principles is knowledge. This level is defined as retrieving "relevant knowledge from long-term memory" (Anderson & Krathwohl, 2001, p. 31). It includes the processes of remembering, recognizing, recalling, identifying and retrieving. The difficulty level for knowledge questions is increased through the level of discrimination or obscurity of the information being recalled. This is the most commonly tested level of the taxonomy.

Comprehension requires a basic level of interpretation, comparison/contrasting of two concepts, explanation, estimation and/or translation. This level is defined as being able to "construct meaning from instructional messages" (Anderson & Krathwohl, 2001, p. 31). This second level includes the processes of exemplifying, classifying, summarizing, and inferring. Comprehension requires the demonstration of the use of an abstraction on demand. This level is emphasized the most in structured educational settings (Bloom, 1956).

The next level in Bloom's Taxonomy is the application of previous concepts to new situations. This third level includes problem solving and the construction of compiled information such as charts and graphs. Processes include executing and implementing (Anderson & Krathwohl, 2001). Hierarchy distinguishes application from comprehension. Application requires the application of an abstraction with the cue to do so (Bloom, 1956).

The last three levels of this taxonomy are analysis, synthesis, and evaluation. To analyze requires "break[ing] material into its constituent parts and determin[ing] how the parts relate to one another and to an overall structure or purpose" (Anderson & Krathwohl, 2001, p. 31). It includes the processes or differentiating, organizing, and deconstructing. Analysis requires distinguishing and interpretation of assumptions.

Synthesis is the integration of multiple ideas to formulate solutions or action plans. It is creating a functional whole by reorganizing "elements into a new pattern or structure" (Anderson & Krathwohl, 2001, p. 31). Processes include creating, generating, hypothesizing, planning, designing, constructing, and producing.

Evaluation is judgment of adequacy or value based on criteria and standards (Center for Nursing Education and Testing, 1998; Gaskins, Dunn, Forte, Wood, & Riley, 1996; Waltz et al., 2005). Processes include checking, coordinating, detecting, monitoring, testing, critiquing, and judging (Anderson & Krathwohl, 2001, p. 31). Evaluation is the highest level of the cognitive hierarchy because it involves aspects of all of the other levels (Bloom, 1956).

Most multiple-choice test items fall within the first three levels of Bloom's Taxonomy (Airasian, 1994; Center for Nursing Education and Testing, 1998). "Investigators encountered difficulties in developing measures, particularly objective ones, at the most complex end of the continuum" (Krathwohl & Payne, 1971, p. 30). Bloom (1956) provides illustrative test items for each of the cognitive domains. Analysis items require the identification of unstated assumptions in order to solve the posed query. Synthesis items include examples of essays, plans, and derivational hypotheses. Evaluation items include validation and support of the selected answer and rationale for the choices not selected. However, Krathwohl and Payne (1971) illustrate potential difficulties accurately classifying a higher cognitive level item.

> The complications of classifying accurately an item in a higher-level category are compounded by the fact that the student's prior experience with the material on which the item is based may have resulted in his learning by rote a problem that would be complex if new to the student. Such a problem would drop into a lower category, such as Knowledge, for the student who learned it by rote, whereas it

would be a measure of more complex behavior for the student who met it afresh (p. 30).

Test Item Development

The development of test items is a crucial factor in the accuracy of the test results. During the development and evaluation process questions arise, such as: Do the items test what I want them to test? Are the tests accurate? Are the test items clearly written; or are they ambiguous, or focusing on trivial or obscure content? (Nunnally, 1970; Van Ort, & Hazzard, 1985). The number of items to include is another consideration. "A test must be long enough to be considered valid and reliable, but short enough to be considered practical or usable" (Stanton, 1983, p. 338).

Quality Test Items

Common problems associated with test items are ambiguity of items, unclear linkages of items with stated objectives, inconsistent item construction, and low domain testing based on Bloom's (1956) Taxonomy (Center for Nursing Education and Testing, 1998; Gaskins, Dunn, Forte, Wood, & Riley, 1996; Waltz et al., 2005). Wesman (1971) suggests expressing each item clearly using concise words in a simple arrangement. Items should be relevant for the audience to be tested and should have expert editorial review prior to being included in the test. Wesman also suggests the following when writing multiple-choice items: the use of either a direct question or an incomplete statement as the item stem; the avoidance of a negatively stated item stem if possible; and, making all distractors plausible, attractive, appropriate, and unique.

Analysis of Test Items

Test item analysis data provide information on each individual item. Item analysis data include item difficulty level and item discrimination index. This information can answer questions related to the item performance, difficulty, and freedom from clues to the keyable answer. It also provides information on the plausibility of the distracters (Jenkins & Michael, 1986; Linn & Gronlund, 2005; Rizzolo, 1987; Waltz et al., 2005). In addition to item analysis, a panel of experts can be used to determine the level of difficulty of an instrument and to set the passing threshold, or cut score.

Validity

According to Nunnally (1970), a widely recognized authority on instrument development "A crucially important phase in the development of a measuring instrument concerns learning whether or not the instrument is useful for any purpose" (p. 132). Nunnally defines validity as the extent that an instrument measures what it is claimed to measure. Validity is the determination of the usefulness of an instrument. Validation is measured on a continuum. The measurement is based, not on the instrument, but rather on the intended use of the instrument (Nunnally). Four types of validity that are referred to in the literature are: content, criterion, construct, and divergent.

Content validity

Content validity measures face validity or the extent to which an instrument measures what it is supposed to measure, such as self-esteem versus self-concept (Cronbach, 1971; Psychometrics, 2004; Reliability and Validity, n.d.; Streiner & Norman, 2003; Williamson, 1999). Content validity measures the accuracy of the information

included in the exam (Center for Nursing Education and Testing, 1998; Flynn & Reese, 1988; Waltz et al., 2005) and if all relevant areas are included (Gronlund, 1973; Streiner & Norman, 2003; Waltz, et al., 2005). Content validity ensures that the keyable answer is correct and that the non-keyable choices are not correct. Content validity also supports that the information included in the instrument reflects all significant concepts within the domain to be measured.

Criterion-related validity

Criterion-related validity measures the individual test-taker's performance against set standards of performance such as Standardized Assessment Tests (SAT) (Center for Nursing Education and Testing, 1998; Gronlund, 1985; Linn & Gronlund, 2005; Waltz et al., 2005). This measurement of validity compares the ranking of test-takers for a particular instrument to their ranking on an established instrument. Consistency between scores or rankings would support the validity of the non-standardized instrument.

Criterion validity measures the degree to which an instrument measures a concept as compared to pre-determined criteria, such as nursing knowledge (Cronbach, 1971; Reliability and Validity, n.d.; Streiner & Norman, 2003; Williamson, 1999). Criterion validity can also be considered concurrent or predictive, but not an explanation of dependence. Concurrent validity is measured at the same time that the criterion is being measured, such as taking a swimming test in the deep end of the pool. Predictive validity is when the measurement predicts the performance or existence of a variable or trait, such as predicting performance on the Nursing Comprehensive Licensing Examination for Registered Nurse (NCLEX-RN) using the Mosby Assesstest ©.

Construct validity

Construct validity refers to the degree that an instrument measures an intended theoretical construct, such as clinical performance (Cronbach, 1971; Reliability and Validity, n.d.; Streiner & Norman, 2003; Williamson, 1999). Construct validity is a way to measure against a defined concept (Center for Nursing Education and Testing, 1998; Gronlund, 1985; Linn & Gronlund, 2005; Waltz et al., 2005). Examples of measurements of a conceptualized topic would include an anxiety index or a depression scale.

Divergent validity

Divergent validity is the degree to which a new measure poorly correlates with measures of different and/or unrelated constructs (Graziano & Raulin, 2000). Divergent validity should indicate that there is not a strong correlation between similar groups or measurements. If there were a strong correlation, then the instrument is not measuring anything unique (New Target, 2005). Divergent validity can be measured through the comparison of responses between groups (van Oppen, 1992). A comparison of convergent correlations with the correlations of unrelated assessments can be used to assess divergent validity (Bell, Greig, Kaplan, & Bryson, 1997; Hull et al., 1995). Also, the findings of different instruments can be compared to assess divergent validity (Loblaw, Bezjak, & Bunston, 1999).

Threats to validity

The validity of an instrument can be threatened by many factors. Difficult vocabulary and sentence structure, inappropriate level of item difficulty, poorly constructed items, ambiguity, inappropriate items for the outcomes being measured, a test

that is too short, improper or illogical arrangement of items, and identifiable patterns of answers can influence an instrument's validity (Gronlund, 1985; Linn & Gronlund, 2005). Unclear directions are also a threat to the validity of a test. Directions should be simple and clear. Directions should address the purpose of the test and informed consent, procedures for guessing, time limitations, procedures for marking answers, and specific instructions for each style of item (Clemans, 1971).

Reliability

Reliability refers to the repeatability or reproducibility of the measurement scores regardless of what is being measured (Nunnally, 1970). Reliability can be estimated through single- and multiple-administration methods to measure intra-individual and inter-individual variability (Stanley, 1971). Reliability can be defined mathematically using classical test theory (CTT) or functionally using item response theory (IRT). CTT is based on psychometrics. CTT includes test construction theory, reliability, internal consistency, and predictability (Langenbucher et al., 2004). IRT includes factor analysis in order to support that a single construct is being measured. The IRT instrument is used to measure the range of the identified construct or trait (Langenbucher et al.).

Tests can also be evaluated for reliability, which is the accuracy and consistency of measurement (Flynn & Reese, 1988; Streiner & Norman, 2003; Waltz, et al., 2005). Test-retest or parallel form comparisons are measures of stability. Internal consistency can be determined by a variety of methods. One method is odd and even split half reliability corrected with the Spearman-Brown formula. When instruments have a right or wrong (dichotomous) response format, the Kuder-Richardson (KR 20) reliability analysis is

appropriate. Point biserial correlation is an item to total score correlation. The most commonly used method for establishing a reliability coefficient is Cronbach's alpha, which measures the degree to which different items measure the same attribute (Center for Nursing Education and Testing, 1998; Gillis & Jackson, 2002; Gronlund, 1985; Linn & Gronlund, 2005; Polit, 1996). The test, the conditions of administration, the sample size, and the participant group's homogeneity or heterogeneity can affect reliability. The length of the test, spread of scores, and the item types and quality will affect the reliability of the test. Time limitations, physical and proctoring conditions, and clarity of the instructions will also affect the reliability of an instrument (Center for Nursing Education and Testing, 1998; Gaberson, 1996; Gronlund, 1985; Linn & Gronlund, 2005).

Testing conditions

Testing environments need to be conducive to optimizing the reliability of the test results. Potential distractions should be minimized during the test administration. Proctoring or supervision provides structure, security, and support to the test takers. Test directions need to be clear regarding the purpose of the test, the time limitations, the recording of answers, and the availability of help during the test. The time allocation needs to be appropriate for the style of the questions and the number of items. Typically for multiple-choice items, one minute per item is used for time planning (Gaberson, 1996; Hambleton & Eignor, 1978; NCSBN, 2003).

Threats to reliability

As mentioned above, the greatest risk to reliability is the validity of the test. If an instrument is not determined to be valid, it cannot be reliable. An example of inconsistency as a threat would be poor and/or not standardized instructions and variability among test takers such as when the sampling of content puts some test takers at an advantage and others at a disadvantage. Random events that may distract or influence the individual's focus on the test including poor testing conditions threaten reliability. Instability of test scores; errors in scoring or having more than one keyable answer for a multiple-choice item; guessing, particularly for true-false items; or subjectivity in scoring for essay items effects the reliability of the test (Nunnally, 1970).

An instrument that is reliable has been proven to be consistent and stable. If the instrument is administered in an identical situation at different times the results will be similar (Twycross & Shields, 2005). Reliability is required for validity, but does not ensure validity (Nunnally, 1970; Reliability and Validity, n.d.). An instrument may give the same results at different administrations; however, the results may not be measuring what the instrument was developed and intended to measure. For example, a scale may be developed with the intent to measure apprehension of patients undergoing invasive procedures. The instrument may be demonstrated to be reliable, giving consistent results when administered. Unfortunately, the instrument may be demonstrated to be invalid in that it is found, upon further study, to be measuring pain levels rather than degree of apprehension.

Summary of Development of Criterion-Referenced Instruments

Many instruments that measure human attributes are developed using a linear model. The most critical aspect of test development and subsequent validity is the test plan. The test plan includes the purpose of the test as well as information about the type, number, and allocation of items among the specified topic areas and the cognitive levels of Bloom's Taxonomy. Test items are developed to reflect the test plan using established guidelines for quality. Test items are analyzed for degrees of content, criterion-related (predictive) and construct validity and reliability.

Use of Criterion-Referenced Instruments by Nursing Certification Organizations

Certification organizations have struggled with the need for easily scored multiple-choice tests and the need to address higher domains, especially for the advanced practice certifications. The American Association of Critical Care Nurses (AACN) combines the cognitive levels: knowledge/comprehension, application/analysis, and synthesis/evaluation (Critical Care Nurse Specialist Exam (CCNS), 2004). The American Nurses Credentialing Center (ANCC) does not use Bloom's Taxonomy, but, rather, uses domains of practice to map their examination content (2004). Dr. Cynthia Miller-Murphy (personal communication, December 3, 2004), director of the Oncology Nurses Credentialing Center (ONCC), acknowledged the limitations of multiple-choice tests when higher levels of Bloom's Taxonomy are being evaluated. The ONCC recently revamped their content map to include only knowledge and application questions. Dr. Miller-Murphy stressed that certification only identifies the achievement of minimal competence thresholds with a pass/fail outcome. ONCC basic level credential

examinations are 40% knowledge and 60% application. ONCC advanced practice examinations are 30% knowledge and 70% application with the requirement of the applicant to mentally move through several steps to get to the answer.

Waltz, Strickland, and Lenz Stages for the Development and Validation of Criterion-Referenced Measures

Waltz and colleagues (2005) developed a collection of theories and principles for the understanding and utilization of nursing measures. Processes were described for selecting or designing measurement instruments. Testing of and applications for such instruments were also included. Their work was directed toward nurse educators, nurse researchers and consumers of nursing research. The authors sought "to meet the needs of a large and heterogeneous nursing audience ranging from neophyte to more advanced in their knowledge and experience in the measurement of nursing phenomena" (p. v). The ultimate goal of their work is to help nurses to make a positive difference in the state of people's health through the conduct of successful outcome studies and quality assessment (Strickland, 1997).

Waltz and colleagues (2005) identified eight stages for the development and validation of criterion-referenced measures. These stages are:

1. Specify the conceptual model of the measure.
2. Specify the purpose(s) of the measure.
3. Explicate objective(s) or the domain definition.
4. Prepare test specifications including:
    a. Method of administration

b. Number or proportion of items that will focus on each objective or subscale
c. Type of items and how they will be created
d. Test restrictions and givens
e. General scoring rules and procedures

5. Construct the measure including:
   a. Develop a pool of items or tasks matched to the objective(s) or subscales
   b. Review items of tasks to determine content validity and their appropriateness
   c. Select items after editing or deleting poorly developed items from the item pool
   d. Assemble the measure (including preparation of directions, scoring keys, answer sheets, etc.)
6. Set standards or cut score for interpreting results.
7. Field-test or administer the measure.
8. Assess reliability and validity of measure (including determining the statistical properties of items, and deleting and revising items further based on empirical data) (p. 122).

These stages clearly itemize the steps necessary for the development of an instrument that measures the established criteria assuming that the criteria accurately reflect the domain of interest.

Waltz and Strickland are internationally recognized specialists in nursing research (Strickland, Burgess, Oberst, & Kim, 1987), measurement, and evaluation. Lenz is known

for her work with mid-range nursing theories and nursing measurement as well as her work in promoting doctoral nursing education (Lenz & Hardin, 2000; Lenz & Ketefian, 1995). These authors are referenced in a multitude of literature as nursing measurement experts.

Within the context of the variety of nursing and multidisciplinary instruments that are available, appropriateness of application and quality of an instrument are of paramount importance (Strickland, 1997). The focus of Waltz and colleagues' work is to increase the operationalization of sound, valid and reliable measures within theoretical frameworks. Qualitative, quantitative, and non-traditional measurement approaches are discussed. Measurement issues such as ethics, including collection of sensitive data; process/outcome measurement; physiologic approaches and measures; evaluation and use of existing instruments; and scaling techniques, including visual analogs, magnitude estimation procedures, and multidimensional scales are addressed in this text (Waltz et al., 2005).

Waltz and colleagues (2005) suggest that criterion-referenced measurement be applied to test skill achievement and/or "to determine an object's status in relation to some specific attribute or property" (p. 195). Criterion-referenced measurement uses a specific domain as a frame of reference for interpreting results (Glaser, 1963; Martuza, 1977; Popham, 1978; Waltz et al., 2005). The participants are measured against the established standard, rather than against each other. Criterion-referenced measures usually provided nominal or ordinal data.

The decision to use criterion-referenced measurement is based on the intended use of the results. "The approach is flexible enough for use in a wide range of different content areas [and it] also appears to enjoy high levels of 'social validity'" (Hart & Sciutto, 1996, p. 26). Criterion-referenced measurement is used to ensure accountability for government and credentialing organizations (Hart & Sciutto). The ANCC Certification Examinations, the National Council Licensure Examination, Denyes's (1980) Self-Care Agency Instrument, The Denver Developmental Screening Test (Frankenburg, Dodds, & Fandel, 1970), Gullickson's (1988) Simulated Clinical Performance Examination Measurement Tool, and Lenburg's (1999b) Competency Performance Assessments (CPAs) and Examinations (CPEs) are examples of criterion-referenced measures (Waltz et al., 2005).

In summary, the quality of each individual item on an instrument as well as the overall quality of the instrument is significant to the instrument's ultimate validity and reliability. Each item must be evaluated for clarity, significance, and accuracy. The instrument must then be reviewed in its entirety for overall validity and reliability. The use of content experts and statistical analyses demonstrate the degree to which a given instrument is valid and reliable.

Internet Data Collection

In order to gather data from a geographically diverse sample, more and more researchers are moving toward Internet data collection (Gosling, Vazire, Srivastava, & John, 2004; Im & Chee, 2003; Wilmoth, 1995). The rapid expansion of the Internet is moving surveys to "'paper-less' and 'people-less' data collection" (Faculty of Social

Sciences, 2002). The availability of online surveys in the 1990's was greeted with a sense of excitement (Huffman, 2006). Currently, the novelty has worn off and people are inundated with 'spam' and surveys. Internet data collection has documented disadvantages and advantages.

Disadvantages

Internet survey data reflect low response rates as compared to printed mailed surveys (Hayslett & Wildemuth, 2004; Yun & Trumbo, 2000). Kittleson (1995) found a survey response rate of 28.1% for Internet versus 76.5% for mailed hard copy surveys with the greatest Internet response within two days of receipt. Hayslett and Wildemuth (2004) discovered very little difference between the demographics of the sample populations who responded to Internet and mailed surveys.

Difficulty transitioning self-administered surveys from paper to Internet, while maintaining reliability and validity; computer program compatibility and the resulting costs if hardware and program interfaces need to be set up; technological errors in creation, implementation, and data gathering; and variations in screen layouts have been distracters from online survey development (Dillman, Tortora, & Bowker, 1999; Faculty of Social Sciences, 2002; Huffman, 2006; Schunlev, 2001; Strickland et al., 2003; White, Carey, & Dailey, 2001). Other challenges related to Internet data collection include non-coverage (i.e. not reaching people who do not have email or that screen or filter their email) for sampling and solicitation, user-friendly (i.e. ease of understanding, low demands of energy and time, and content at fifth grade reading level) interface of questionnaire layout, weighting and imputation post-survey, cultural and country-specific

effects on the quality of data, confidentiality and privacy issues, and the impact of new technologies (Evans & Mathur, 2005; Faculty of Social Sciences, 2002; Strickland et al., 2003). Participants were noted to become frustrated with Internet surveys, as they were unable to select multiple choices to personalize the survey. Many times 'other' was selected as a default (Hayslett & Wildemuth, 2004).

Validity of Internet surveys can be threatened by the ability for responses to Internet surveys to be purposefully or accidentally deleted or double counted, the inability to proctor respondents or even confirm that they exist, the possibility of multiple responses from the same participant, and incomplete responses due to rushed completion of the survey (Dillman et al., 1999; Farmer, 1998; Huffman, 2006; Wright, 2005). "The appeal of recruiting participants quickly, easily and cheaply ought to be weighed against the stumbling blocks one may encounter when making causal inferences from an experiment or when providing populations estimates" (Siah, 2005, p. 118). Another potential threat to the reliability and validity of Internet surveys is the ethics of conducting Internet research (Siah). Varnhagen and colleagues (2005) examined the ethical issue of informed consent. They determined that, although participants took longer to read informed consent documents when provided online, the retention of informed consent information was similar to that of paper versions of the same consents.

Advantages

Although it was noted above that Internet surveys can be more expensive to develop if interface and programming issues arise, pre-packaged survey programs can make Internet surveys more cost effective than the traditional printed mailed survey

(Huffman, 2006; Pitkow & Recker, 1995). Opportunities that the Internet provides for surveys includes better data quality with less risk of coding errors; the ability to recruit through listservs; access to specific, unique populations; easier preparation; reduced respondent burden; quicker survey timing; a more intimate venue for sharing information; an increased willingness to participate in future surveys; an effective mechanism for gathering opinions on sensitive issues; nearly identical response results compared to phone surveys; and availability of analytical instruments (Evans & Mathur, 2005; Miller & Hogg, 2000; Rhodes, Bowie, & Hegenrather, 2003; Smith, 1997; Strickland et al., 2003; Sweet & Russell, 1996; Swoboda, Muehlberger, Weitkunat, & Schneeweiss, 1997; Wright, 2005). Internet surveys also allow for higher-quality graphics and multimedia presentations (Kennedy, 2005; Sheehan & Hoy, 1999).

Gosling and colleagues (2004) explored documented concerns regarding the use of Internet methods for data collection and found that most were unfounded. Internet samples were found to be more demographically diverse than traditional samples; however, they are still not completely representative of the population. Internet users were not unmotivated, depressed or social deviants. Internet data may be generalized across presentation formats and findings are consistent with those based on traditional data collection methods. Internet data collection can compromise the anonymity of the participants; however, researchers can implement practices to monitor participation without compromising the protection of human subjects (Gosling et al.).

Recruitment of Nurses

Nurses have been found to be particularly difficult to recruit via the Internet (Ehrenberger & Murray, 1998; Im et al., 2006). Recruitment strategies need to be flexible as nurses are reached through different means (Barriball & While, 1999; Im et al., 2006). Authenticity and trust were difficult to establish without face-to-face contact between researcher and participant (Im et al., 2006). Changing email addresses, inconsistent listserv participation, and multiple recruitment contacts due to multiple organizational memberships proved frustrating to researchers (Im et al., 2006). Having the recruitment message cleared through 'spam' filters can also interfere with recruitment (Graham, 2002; Sorkin, 2005; Stone & Weil, 2003).

Service Providers

Data collection using the Internet has been available since 1999 and is offered in two methods. One method is gathering server data into log files, then using an analysis program to generate reports. It can be a time consuming process to create the instrument in an Internet format; however, data collection is accomplished automatically. Raw data can be transferred to a usable database with a minimal amount of effort (Daley et al., 2003; Web Trends, 2000). Some of the companies that provide online survey services are Survey Monkey © (www.surveymonkey.com), Zoomerang © (www.zoomerang.com), Question Pro © (www.questionpro.com), Web Surveyor © (www.websurveyor.com), Custom Insight © (www.custominsight.com), and Zip Survey © (www.zipsurvey.com) (Huffman, 2006; Read, Perry, & Duffy, 2005; Wright, 2005).

The second method involves 'packet sniffing' to gather data. 'Packet sniffing' is a process by which the input data from discussion boards, interview transcripts, or open-ended survey questions are searched for terms, phrases, characteristics or other identifiers. The data are then gathered into a database for further analysis (Perkins, 2004; Web Trends, 2000).

Internet Data Collection Summary

Thus, Internet data collection can be used to gather and manage data in both qualitative and quantitative research as well as for test administration. Prior to initiating data collection via the Internet, either qualitative or quantitative, specific steps must be identified regarding format and placement of the instrument and retrieval of data (Strickland et al., 2003). Protocols for screening criteria, orientation to the site, data linkage mechanisms, confidentiality management, and permanent removal from the website must be identified (Strickland et al.).

Recruitment of participants though the Internet has both advantages and disadvantages (Cooper, 2000; Duffy, 2002; Huffman, 2006; Wright & Neill, 1999). The advantages of listserv recruitment include potential diversity of the sample, speed of data access, and reduced cost compared with regular mailing. However, web-based surveys may not accurately represent the target population, since participation is limited to individuals with Internet access and computer literacy. Other limitations involve participant confidentiality, since an electronic record of the responses is generated. Participants may have distractions when they are on the Internet accessing the instrument. Internet testing conditions are not controlled like a face-to-face testing venue

provides. Internet connections and speed as well as user expertise are also factors that may limit study participation.

Summary

The review of literature on the COPA model yielded a few articles on this relatively new model. Most of the articles included the originator of the model as one of the authors. The model was successfully applied in a variety of settings and has been suggested as a competency assessment framework for professional and nonprofessional settings beyond nursing.

The literature on the documentation of professional achievement elaborates on the terms professional competence, qualification, credentialing, and certification. Many of the terms were used interchangeably; however, the literature indicated very specific, if overlapping, definitions for each term.

The review of literature yielded a large volume of anecdotal information about portfolios in various settings. Very little research-based information was available regarding portfolio review data and analysis. Few studies are available regarding outcome evaluation data and analysis.

Nurses in the clinical specialty of genetics are credentialed using a portfolio review process. Portfolios have been used by a variety of professions and in a variety of settings to document competence. The Genetic Nurses Credentialing Commission is the first national nursing organization to use the portfolio for the awarding of professional credentials.

In the complex context of nursing competence, certification and credentialing, the norm is for acknowledgement of knowledge in practice through the process of a standardized examination. The development and testing of a criterion-referenced instrument is a multi-step process as supported in the literature. Important factors to consider in the development of an instrument are number, type, cognitive level, difficulty, and quality of items. A newly developed instrument requires analysis for validity and reliability. Content, criterion-related, construct, and divergent validity can be used to evaluate the degree to which an instrument meets the intended purpose.

The instrument can be administered in person or over the Internet. Internet data collection is an approved method for conducting research with advantages and limitations that need to be considered. As with any data collection method the type and purpose of the study, access to the target population, style of instrument, level of researcher confidence and comfort, and availability of resources need to be considered when selecting a method of data collection.

CHAPTER THREE: METHODOLOGY

Introduction

This chapter includes the research design, description of instrument development and sample, Internet procedure for data collection, plan for psychometric data analysis, and the limitations of this study. The purposes of this study were to: 1) develop a criterion-referenced instrument using the steps outlined by Waltz, Strickland and Lenz (2005) to provide effective documentation of knowledge of advanced practice nursing in genetics; 2) determine the validity and reliability of the instrument; 3) compare pass/fail outcomes of nurses within and outside of the genetic specialty using the instrument; 4) compare pass/fail outcomes of Masters' prepared and non-Masters' prepared nurses using the instrument; and, 5) compare pass/fail outcomes of credentialed and non-credentialed advanced practice nurses in genetics using the instrument. An additional proposed aim, to evaluate congruence between instrument pass/fail outcomes and Genetic Nursing Credentialing Commission (GNCC) portfolio pass/fail outcomes, was not able to be fulfilled.

The investigator conducted this study to develop an instrument to document the knowledge of advanced practice nurses in genetics and to establish the validity and reliability of that instrument. Although the sample was not sufficient to meet the last purpose of the study stated above, the instrument may later be used to facilitate the evaluation of the current portfolio methodology used by the Genetic Nursing Credentialing Commission (GNCC) for awarding the advanced practice nurse in genetics credential. This study was based on Lenburg's (1999a, 1999b) Competency Outcomes

and Performance Assessment (COPA) Model and used Waltz, Strickland, and Lenz's (2005) framework for the development and validation of criterion-referenced measures.

The research questions of interest were:

1. To what extent is the Criterion-Referenced Measurement Tool for Genetics (CRMTG) a valid and reliable indicator of knowledge of advanced practice nursing in genetics?
2. Do nurses who practice in a genetic setting have a higher pass rate on the CRMTG than nurses who do not practice in a genetic setting?
3. Do advanced practice nurses have a higher pass rate on the CRMTG than nurses who do not have their Masters' degree in nursing?
4. Do nurses credentialed by the GNCC have a higher pass rate on the CRMTG than non-credentialed nurses?
5. Are GNCC portfolio pass/fail outcomes congruent with CRMTG pass/fail outcomes?

Research Design

This study was a methodological psychometric investigation. A methodological study follows an established series of steps for developing, validating, and evaluating a research instrument. The evaluation of a research instrument's quality in regards to data collection is considered psychometric analysis. This process is based on a review and estimation of the instrument's validity and reliability (Polit & Beck, 2005).

Using the work of Waltz et al. (2005), a criterion-referenced instrument was developed and subjected to psychometric analysis of its validity and reliability. The

purpose of the instrument is to provide effective documentation of knowledge of advanced practice nurses in genetics. The domain being tested is genetic nursing practice. The International Society of Nurses in Genetics (ISONG) Scope and Standards of Genetics Clinical Nursing Practice (1998) was selected by the investigator for the CRMTG criterion as this document provides the nationally approved practice standards within the domain of genetic nursing.

Sample

This study used a purposive convenience sample of 356 registered professional nurses in the United States who volunteered to access the CRMTG via the Internet. Specific information about the sample as well as protection of human subjects will be discussed under the "Administration of Test" section later in this chapter.

Procedure for Instrument Development

Waltz, Strickland, and Lenz's (2005) stages for the development and validation of criterion-referenced measures were used to develop the CRMTG. These stages are described below with relevant information pertaining to their applications to the study (Table 2).

Stage 1. Specify the conceptual model

The purpose of this instrument was to effectively document knowledge of advanced practice nursing in genetics. The COPA model was selected as the conceptual model for the study and for the instrument because it addresses the evaluation of professional competence in clinical nursing using established indicators. This model is

Table 2

Development of the Genetics Criterion-Referenced Measurement Tool Using the Waltz, Strickland and Lenz Model

Waltz, Strickland and Lenz Stages	This study
1. Specify the conceptual model.	COPA Model
2. Specify the purpose(s) of the measure.	To provide effective documentation of knowledge of advanced practice nurses in genetics
3. Explicate objective(s) or domain definition.	ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998)
4. Prepare test specification including:	
a. Method of administration	Internet
b. Number or proportion of items that will focus on each objective or subscale	Test Blueprint
c. Type of items and how they will be created	Multiple choice created by the investigator
d. Test restrictions and givens	Test takers have limited access to online instrument
e. Scoring rules and procedures	Scoring for the instrument was one keyable answer per item. Scoring was done by the investigator. Results were reported as Pass/Fail
5. Construct the measure including:	
a. Develop a pool of items or tasks matched to the objective(s)	Items were selected using the established pool in Lewis, R. (1999). Instructor's manual test item file to accompany human genetics (3rd ed.). Boston: WCB/McGraw-Hill.
b. Review items or tasks to determine content validity and their appropriateness	100 Items reviewed by construct and content experts

Waltz, Strickland and Lenz Stages	This study
c. Select items after editing or deleting poorly developed items from the item pool	Items of concern were rewritten based on the expert review
d. Assemble the measure	Genetics Criterion-Referenced Measurement Tool
6. Set standards or cut score	Content experts were used to verify content and Bloom mapping and to determine the cut score for the instrument using the Ebel (1979) method
7. Field-test or administer the measure	The 116-item Genetics Criterion-Referenced Measurement Tool was field-tested using paper and pencil at the ISONG Annual Convention on November 2, 2003 (N = 6) Instrument was administered to five Sample groups (N = 356)
8. Assess reliability and validity of measure	Psychometric data analyses included: item-to-total correlation, paired t-test, item analysis, and chi square

consistent with both the stages of instrument development being used in this study and the portfolio process being used by GNCC.

Criterion-referenced measures are used to determine status with respect to identified performance standards (Waltz et al., 2005). This study was designed to provide psychometric data for a pass/fail credentialing process. The goal was to determine if a minimum level of knowledge had been achieved, not to rank the participants; therefore, a criterion-referenced instrument was developed.

Stage 2. Specify the purpose of the measure

The purpose of the CRMTG is to reliably document knowledge of advanced practice nursing in genetics.

Stage 3. Explicate objectives or domain definition

Standards of practice define the expected objectives of nursing care. The instrument used to measure these objectives must be based on these standards (Bronstein, 2002; Clinton, Denyes, Goodwin, & Koto, 1977). The domain for the CRMTG is knowledge of advanced practice nursing in genetics. The standards of practice for this domain are described in The ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998). These standards were used to develop the test plan blueprint for the instrument as one of the accepted approaches for development of criterion-referenced instruments. The scope and standards document has been revised; however, Genetics/ Genomics Nursing: Scope and Standards of Practice (2007) was not available at the time that the CRMTG was developed.

The ISONG standards of advanced practice clinical genetic nursing, found in The ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998), contains three primary practice areas: evaluation, knowledge, and teaching and research. Professional competence in this specialty encompasses the six steps of the nursing process: assessment, diagnosis, outcome identification, planning, implementation, and evaluation. The step of implementation is further defined as including: identification of risk, case coordination, health promotion and disease prevention, genetic psychosocial counseling,

and therapeutic communication (ISONG, 1998). The test plan blueprint for the CRMTG reflects these practice areas and steps of the nursing process (Table 3).

Table 3

Criterion-Referenced Measurement Tool for Genetics (CRMTG) Test Blueprint

	Number of items per genetic category				
ISONG/ANA Standards of Genetic Clinical Nursing Practice (1998)	General Genetics (30%)	Reproductive Genetics (16%)	Adult Genetics (27%)	Pediatric Genetics (27%)	Row Total
Evaluation (40%)	10	8	14	20	52
Assessment (4%)	1	1	3	3	8
Diagnosis (4%)	1	0	0	4	5
Outcome Identification (4%)	1	1	2	3	7
Planning (4%)	2	0	2	2	6
Implementation (20%)	5	6	6	5	22
Identification of Risk (4%)	1	1	0	0	2
Case Coordination (4%)	2	0	1	1	4
Health Promotion/ Disease Prevention (4%)	2	1	2	1	6
Genetic Psychosocial Counseling (4%)	0	4	2	2	8
Therapeutic Communication (4%)	0	0	1	1	2
Evaluation (4%)	0	0	1	3	4
Knowledge (30%)	16	7	10	6	39
Teaching and Research (30%)	4	1	3	1	9
Total Items	30	16	27	27	100

Stage 4. Prepare test specifications

Number or proportion of items per subscale

Genetic practice areas of general, reproductive, adult, and pediatric have been selected as inclusive content areas for the CRMTG (Lashley, 2005; Lea, Jenkins, & Francomano, 1998) (Appendix A). Fallon (2004) identified prenatal, pediatric, adult, and cancer as key areas of genetic counseling practice. Advanced practice nurses in genetics offer genetic support in these specialty areas as well. For the CRMTG a general genetic knowledge section was included that would be applicable to all practice areas. The adult category includes cancer genetics as well as cardiac, diabetes, and other primarily adult genetic disorders. Prenatal genetics is labeled "reproductive genetics" on the CRMTG content map. Content was mapped to include these practice settings, general genetic knowledge, and the practice standards identified in The ISONG Scope and Standards of Genetics Clinical Nursing Practice (1998).

Type of items and how created

Test specifications were prepared for a pilot study of the first draft using the identified domain knowledge of advanced practice nursing in genetics. A multiple-choice format was selected because this format parallels the credentialing examinations given by other credentialing organizations. The final instrument contains 100 multiple-choice items. This number is consistent with most credentialing examinations, which include 100 to 300 items (ANCC, 2003a, 2003b) with the most common number being 150 (ANCC, 2004), and was selected based on the number of subjects needed to validate the

instrument. For this study 356 subjects were recruited which exceeded the minimum requirement of 300 subjects for the instrument length of 100 items.

Test restrictions and givens

The instrument could be accessed only through a specific URL. Participants required no password once they obtained the URL. The website opened directly to the first page of the instrument upon initial access. A cookie was left on users' computers, so that, when they resumed the survey instrument, they could begin where they left off, thus decreasing respondent fatigue (Waltz et al., 2005) and preventing respondents from changing answers.

Scoring rules and procedures

Consistent with other multiple-choice examinations, the CRMTG had one keyable response per item. The investigator manually scored each participant's CRMTG by comparing participant responses to the answer key. The content experts verified the answer key during the content validity process that will be discussed in the "Construct the Measure" stage. Consistent with other credentialing examinations, the CRMTG results were reported as pass or fail based on the cut score determined by the expert reviewers.

Participants were not able to change answers or to go back in the CRMTG and fill in missing items. There were no forced response items; therefore, the participants were able to skip demographic and CRMTG items if they so opted. The investigator controlled missing data by marking the item incorrect, as is standard for unanswered items on multiple-choice examinations. Skipped item data are available in Appendix B. Data from each completed item were used for item analyses.

Stage 5. Construct the measure

Develop a pool of items

The investigator constructed a draft instrument using nursing text and practice sources (Jenkins & Lea, 2005; Lashley, 2005; Lea, Jenkins, & Francomano, 1998; Lewis, 2005), national nursing standards, and case studies in addition to the Lewis (1999) Instructor's Manual Test Item File to Accompany Human Genetics (3rd ed.) test bank. Permission was granted for the use of this test bank in this study (Appendix C). CRMTG content was mapped to reflect the ISONG standards; general genetics; and the sub-specialty practice areas of reproductive, adult, and pediatric genetics. The GNCC used a team of expert nurse clinicians and educators to develop the weighting of the portfolio elements based on the ISONG standards. GNCC portfolios are evaluated using the following weightings: 40% of the content was based on the operationalized standards of genetic nursing care; 30% of the content was based on general genetics knowledge; and the remaining 30% was based on patient, family, community, and peer education (GNCC, 2001; Middleton, 2002) (Table 3).

The final pool of items included in the CRMTG did not exactly match the weighting of standards used by GNCC for their portfolio scoring. Once items were reviewed for content validity the number of valid test items were not in balance with the original test blueprint plan. The investigator decided to use the stronger items and deviate from the planned blueprint rather than maintain the blueprint and risk item and instrument validity. In the future, as additional items are created and tested, teaching and research items will be an area to be further developed. The investigator wrote items for this

particular area that were specific to genetics. Additional items could be developed which reflect general processes used by the advanced practice nurse in the roles of teacher and researcher.

Review items to determine content validity and their appropriateness

This investigator, who has had training in test item writing and review, created the draft instrument using standard test construction methods (Demetrulias & McCubbin, 1982; Frisbie, 1983; Gaberson, 1996; Rizzolo, 1987; Stanton, 1983). Construct and content validity were enhanced by expert reviews using Van Ort and Hazzard's (1985) test item evaluation criteria. The construct experts were two Widener faculty members with expertise in test item construction. Five nurses with expertise in genetic nursing were selected for content review.

Select items after editing or deleting poorly developed items

Construct and content experts independently reviewed the first draft of the instrument which contained 100 items (Appendix D). Feedback regarding content distribution and items of concern (i.e., two keyable answers, misleading wording, imbalance in response format or length, inaccurate information) was returned to the investigator through errata notations on the test form. Revisions to the instrument were made based on the reviewers' recommendations. At this time 2 questions were deleted, having been determined non-relevant; 15 questions were rewritten based on expert suggestions for clarity and accuracy; and 18 new questions were written to include a broader scope of genetic nursing practice, greater complexity based on Bloom's Taxonomy and to increase the test item pool (Table 4).

Table 4

CRMTG Item Development

Event	Items	Number of Items on CRMTG	Total Item Pool
Construct and Content Expert Review		100	100
Deleted	2	98	100
Rewritten	15	98	100
Created	18	116	118
Pilot		116	118
Deleted	12	104	118
Rewritten	4	104	118
Created	29	133	147
Expert CVI		133	147
Deleted	33	100	147
Rewritten	44	100	147
Created	0	100	147
CRMTG		100	147

As discussed in Chapter Two, Bloom and associates (1956) identified levels of evaluation. These cognitive levels are: knowledge, comprehension, application, analysis, synthesis, and evaluation. The literature supports the challenge that the investigator encountered in writing multiple-choice test items that measured the more complex cognitive levels of analysis, synthesis, and evaluation (Airasian, 1994; Center for Nursing Education and Testing, 1998; Krathwohl & Payne, 1971; Waltz et al., 2005). Case study items were developed and added to the CRMTG in an attempt to increase the complexity of test items (Nkanginieme, 1997) (Table 5).

Writing multiple-choice items at appropriate cognitive levels was a very difficult undertaking. The bulk of the items on the CRMTG were at the lowest cognitive level,

knowledge. Because this instrument is designed for advanced practice nurses, higher cognitive levels would be preferred. Although the expert reviewers identified 11 items as analysis and 4 items as evaluation the literature suggests that multiple choice questions cannot be used for measurement of analysis, synthesis, or evaluation because selecting the one correct answer does not deconstruct, establish a new pattern, or test the validity of a concept (Anderson & Krathwohl, 2001).

Table 5

CRMTG Items by Bloom Cognitive Level

Bloom Cognitive Level	Percent of items
Knowledge	35
Comprehension	27
Application	23
Analysis	11
Synthesis	0
Evaluation	4

Assemble the measure

A 116-item instrument was developed using the approved items, as revised, from the initial draft submitted to the experts. The Criterion-Referenced Measurement Tool for Genetics (CRMTG) was developed and mapped based on the ISONG standards.

Pilot test

After receiving Institutional Review Board approval from Widener University and Board of Directors approval from ISONG, a paper and pencil pilot test of the CRMTG was conducted (Appendix E). In anticipation of the use of the CRMTG for advanced practice nurses in genetics, members of ISONG were contacted for the pilot. The investigator contacted the ISONG Board of Directors and received approval for conducting the pilot at the ISONG 16th Annual International Conference in Los Angeles, California. The investigator worked with the conference planning committee to secure a time slot, room, and refreshments for the pilot. A call to participate was included in the conference brochure. The pilot participants all had a background in genetics, although they had a variety of nursing education and practice backgrounds (Appendix F).

The primary purpose of the pilot was to edit or eliminate items that were unclear or not related to advanced practice nursing in genetics (Henrysson, 1971). In addition, information was sought from pilot study participants regarding their ease in answering each item (Appendices E & G). Confirmation that two hours was an adequate time allowance for completion of the CRMTG was also sought and obtained, with most participants having completed the CRMTG within one hour.

After six nurses pilot tested the CRMTG, piloting was terminated because responses yielded essentially no new information. Fifteen items were re-written for clarification, accuracy, and to increase level on Bloom's Taxonomy and 2 items were deleted. Three multiple-choice and 16 case study-style questions were created and added

to the instrument for a total of 133 items (Table 4). Pilot study results were compared to content expert feedback for the final draft of the instrument.

Post pilot review of items to determine content validity and appropriateness

Content experts reviewed the 133-item instrument for content validity (Beck & Gable, 2001; Lackman, Nieto, & Gliem, 1997; Waltz et al., 2005; Weis & Schank, 2000) using Gronlund's (1973) Check List for Evaluating a Criterion-Referenced Test and a content validity index was calculated (Polit & Beck, 2005). The experts were selected from advanced practice nurses who are members of ISONG. These experts included: 1) a past-president of ISONG who is a professional registered nurse practicing in genetics and is certified by the American Board of Medical Genetics; 2) a past ISONG secretary who is a study coordinator for a genetic research company and a professional registered nurse; 3) a past executive director and founding president of GNCC who has more than 20 years experience in nursing education including 10 years as program chair and is a well-published nursing expert in genetics; 4) a nursing faculty member with her Masters' of Science in Nursing specializing in genetics; and 5) the co-chair of the ISONG education committee who is Vice-Dean of Nursing at a Carnegie (2000) classified Doctoral/Research University - Intensive in the greater New York area and has clinical background in maternal child nursing.

The five content experts were approached to review the 133-item draft of the instrument. They were provided with the test plan blueprint, the instrument, the content areas for each item as they relate to the blueprint, and the keyable answer for each item (Appendix H). They were asked to rate each item for degree of construct quality, degree

of content accuracy, degree that item reflects identified content areas, level of difficulty of the item, and the degree that the item reflects knowledge of advanced practice nursing in genetics. They were asked to identify the appropriate cognitive level for each item using Bloom's taxonomy. The expert reviewers were also asked to rate the overall CRMTG on the degree that items measure an adequate representation of advanced practice nursing in genetics (Gronlund, 1973; Halpern, Thompson, & Schaffer, 2001; Wu & Yu, 2003) (Appendix I).

The index of content validity (CVI = .88) was calculated from the five content experts' ratings of the content relevance of the items on the instrument (Wu & Yu, 2003). A 4-point Likert scale was used where 1 = "strongly disagree," 2 = "disagree," 3 = "agree" and 4 = "strongly agree." The number of items that scored a 3 or 4 was calculated. A low percentage of agreement between the experts indicated a problem with the item (Beck & Gable, 2001; Lynn, 1986). Items identified as relevant to advanced practice nursing in genetics knowledge by 80% or more of the experts and pilot study participants were retained. Items identified as problematic or irrelevant by 60% or more of the experts and pilot study participants were removed from the CRMTG (BenDebba, Heller, Ducker, & Eisinger, 2002; Waltz et al., 2005).

The final CRMTG contains the highest rated 100 of the 147 total expert-reviewed and pilot tested items that had been created or modified for the item pool (Appendix J & Table 4). This is supported by the recommendations that one-and-one-half times as many items should be reviewed as will be included in the final instrument to allow for elimination of the most problematic items (Nunnally, 1978; Pett et al., 2003). The

CRMTG (Appendix K) has a Flesch Reading Ease of 46.8 and is at Flesch-Kincaid Grade Level 8.2. The Flesch Reading Ease and Flesch-Kincaid Grade Level are two measures of readability. They are based on the number of words in the sentences and the number of syllables per word. The Flesch reading ease provides a score in a range from 0 to 100; the higher the score, the easier the reading. The score is based on a calculated percentage of potential readers who would be able to read the text. A score of 48.9 means that 48.9 % of potential readers could read the text and 51.1% would have difficulty. The Flesch-Kincaid Grade Level converts the Flesch score to United States grade-school levels (Bastable, 1997; Readability formulas, 2004). A grade level of 6 to 8 and/or a reading ease of 60 to 70 are ideal for most public documents (Child, 2004). Because the target population for this instrument is registered professional nurses with a Masters' degree, the reading level of the CRMTG is not prohibitive. Only a sample of questions is included in Appendices D, E, H, and K for CRMTG security purposes.

Stage 6. Set standards or cut score

Data were collected both as raw scores and converted to dichotomous scores. The total raw score for each participant was compared to the calculated cut score. The dichotomous scale of pass = 1 and fail = 0 was established once the cut score had been identified (Ferguson, 1976). The cut score was determined based on expert panel review (Violato, Marini, & Lee, 2003). The five content experts were enlisted to determine the cut score because intra-judge inconsistency is inversely proportional to the judges' related content knowledge (Chang, 1999; Goodwin, 1999). Hurtz and Hertz (1999) have

established that as few as five experts in the subject matter can successfully determine a cut score.

The Ebel (1979) method was used to determine the passing, or cut-score, for the CRMTG (Appendix I). In this approach, a group of experts independently reviewed the test items and were asked to rate each item for relevance and difficulty (Waltz et al., 2005). The experts were given the correct answers for each test item. The experts reflected on the minimum knowledge level that was being measured by the test based on the identified professional performance standards. Each expert judge then answered the question "If a borderline test taker had to answer a large number of questions like these [in this cell], what percentage would he or she answer correctly" (Zieky & Perie, 2006)? The percentages were then multiplied by the number of items assigned to that cell. The responses from the five experts were averaged to determine the cut score (Ebel & Frisbie, 1991; Waltz et al., 2005).

Typically, for criterion-referenced tests, mastery of knowledge is equated with a score of 80% of the items answered correctly (Wilde & Sockey, 1995). The cut score for the CRMTG was determined, using the Ebel method, to be a score of 65% of the items answered correctly. One factor that may have affected the CRMTG cut score was the knowledge level of the experts. Entry-level genetic nurses might have perceived items that were seen as common knowledge to the experts as quite difficult. Another factor that might have affected the cut score was the distribution of CRMTG items over three specialty areas in addition to the general genetics content area (Table 3). The lower than

typical cut score might reflect that any single genetic nurse would be expected to be competent in general genetics and his or her specialty area, but not in all specialty areas.

Stage 7. Administer the measure

This section, discussing the administration of the CRMTG, provides detailed information regarding study sample, protection of human subjects, and Internet survey procedure.

Target population and sample

In order to evaluate the validity and reliability of the CRMTG, advanced practice genetic nurses were the target population for this study. However, because there are a very small number of advanced practice genetic nurses, and, in order to provide the diversity needed for validity assessment, other registered nurses were included as well. Clinical genetic nurses of a variety of educational preparations were contacted in order to evaluate the ability of the CRMTG to differentiate advanced practice knowledge from basic clinical genetic knowledge. Non-genetic nurses were also contacted in order to evaluate the ability of the CRMTG to differentiate genetic nursing knowledge from basic nursing knowledge. To gain appropriate study participants, genetic nursing programs and nursing organizations were selected based on the profile of the graduates/membership and ease and cost of access to the graduates/members.

Access to subjects

Several steps were required to obtain access to the five distinct groups of registered nurse participants who comprised the study sample. First, an Internet search was performed looking for nursing programs that have a Masters' level genetic nursing

program. Program deans were then contacted by telephone to determine if contact with the students for study purposes was allowed. Information regarding enrollment numbers and the preferred procedure for submission of the call for participants was also obtained during these phone conversations. Next, several national genetic and advanced practice nursing organizations were contacted by telephone to determine membership numbers, the research contact person, and the procedure one needed to use in order to submit the call for participants to members. Finally, the program deans and research directors of the schools and professional organizations were re-contacted via a follow-up email (Appendices L & M) once approval for the study was obtained (Appendix N). See Table 6 for a listing of nursing organizations, mechanism for contact, number of members, and number of actual study participants obtained from that organization. Participants from organizations not directly contacted may have been contacted through personal email or through other organizations and listed an affiliation with the identified organization.

Five distinct groups of participants were required for this study. The groups and their primary inclusion criteria were: (A) Nurses with, at least, an MSN who practiced in a genetic setting; (B) Nurses with, at least, an MSN who practiced in a genetic setting and who were credentialed as APNGs; (C) Nurses who did not have, at least, an MSN who practiced in a genetic setting; (D) Nurses with, at least, an MSN who did not practice in a genetic setting; and (E) Nurses who did not have, at least, an MSN who did not practice in a genetic setting. Subjects for Group A were accessed through schools of nursing with Masters' in Nursing programs with specialization in genetics and through professional nursing organizations in the genetic specialty. These schools included the University of

Table 6

Table of Nursing Organizations and Participant Contact Format (N = 356)

Organization	Mechanism for Contact	Number of Potential Contacts	Number of Participants
National Association of Clinical Nurse Specialists (NACNS)	Email	2,550	117
National Society of Genetic Counselors (NSGC)	NSGC E-Blast	1,757	112
Personal contacts	Email	250	61
International Society of Nurses in Genetics (ISONG)	ISONG listserv	277	42
American Association of Critical Care Nurses (AACN)	Not directly contacted	65,000	10
Oncology Nursing Society (ONS)	Not directly contacted	33,000	7
University of Cincinnati	Email to University	5	2
Emergency Nurses Association (ENA)	ENA listserv	23,000	1
American Academy of Nurse Practitioners (AANP)	Not directly contacted	15,000	1
Academy of Medical-Surgical Nurses (AMSN)	AMSN E-Newsletter	12,529	1
American Organization of Nurse Executives (AONE)	Not directly contacted	5,000	1
American Psychiatric Nurses Association (APNA)	Not directly contacted	4,900	1

Organization	Mechanism for Contact	Number of Potential Contacts	Number of Participants
University of Maryland	Email to University	6	1
Affiliation not identified			20

Cincinnati and the University of Maryland. Program directors from these schools were contacted and asked to distribute the request for participation to their advanced practice nursing in genetics students at the end of their nursing program and to recent graduates of these programs (Appendices L & O). Additional subjects for Group A were accessed electronically through the ISONG listserv and a National Society of Genetic Counselors (NSGC) E-Blast (Appendices M & O). Subjects were given a request for participation and instructions as noted above.

Subjects for Group B were APNGs accessed electronically through the GNCC newsletter (Appendix O). Subjects were given instructions as noted above. Subjects for Group C were accessed electronically through the ISONG listserv and a National Society of Genetic Counselors (NSGC) E-Blast (Appendices M & O). Subjects were requested to participate and given instructions as noted above. In some cases, subjects accessed for Group C met Group A criteria and were, therefore, assigned to Group A accordingly.

Subjects for Groups D and E were accessed electronically after obtaining approval for membership access from the Directors of Practice and Research from selected national nursing organizations (Appendices M & P). The organizations accessed were as

follows: Emergency Nurses Association (listserv), Academy of Medical Surgical Nurses (E-News), and National Association of Clinical Nurse Specialists (individual emails to members). Additional subjects were also accessed through personal emails from the investigator. All subjects were given instructions as described above. Table 7 shows the distribution of study respondents by contact organization and group. Some respondents indicated affiliation to more than one group.

Method of administration

The CRMTG was administered via the Internet. The instrument was managed by an internet survey company (Surveymonkey © at www.surveymonkey.com) that collects and stores the data, which were then available to the investigator as raw data in a Microsoft Excel ® file. This company was selected based on colleague referrals, literature review, cost, ease of use, format of raw data, data analysis capabilities, and service stability. Surveymonkey © staff did not have access to the identities of the users as users are given a computer-generated code upon initial access to the site. Detailed information regarding use of the Internet for data collection appears later in this chapter.

Characteristics of study participants

Appropriate categorization of the five study groups was verified based on subjects' responses to the demographic portion of the internet survey instrument. Specifically, group assignment was based on demographic responses to questions 1- What is your highest completed nursing degree? and questions 9, 10, and 11 relating to genetic practice (Appendix Q). In addition, question number 2 on the opening page of the site was used to identify members of The Genetic Nurses Credentialing Commission

Table 7

Distribution of Respondents by Contact Organization and Group

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Group D (n = 169)		Group E (n = 92)		Total (N = 356)	
	n	%	n	%	n	%	n	%	n	%	n	%
NACNS	1	1.8	-	-	-	-	115	68.0	-	-	116	34.4
NSGC	32	56.0	-	-	28	84.8	-	-	48	53.9	108	32.0
ISONG	20	35.1	-	-	4	12.1	10	5.9	4	4.5	38	11.3
GNCC	1	1.8	5	100.0	-	-	-	-	-	-	6	1.8
AACN	-	-	-	-	-	-	3	1.8	2	2.2	5	1.5
ONS	-	-	-	-	-	-	3	1.8	-	-	3	0.9
U of C	-	-	-	-	1	3.1	1	0.6	-	-	2	0.6
APNA	-	-	-	-	-	-	1	0.6	-	-	1	0.3
AMSN	-	-	-	-	-	-	1	0.6	-	-	1	0.3
U of M	-	-	-	-	-	-	1	0.6	-	-	1	0.3
ENA	-	-	-	-	-	-	-	-	-	-	-	-
Personal contacts	3	5.3	-	-	-	-	28	16.6	25	28.1	56	16.6
No affiliation identified	-	-	-	-	-	-	6	3.6	13	14.1	19	5.3

(GNCC) (Appendix R). Subjects assigned to Group A indicated a Masters' or Doctorate degree in nursing and employment in a genetic setting. Subjects assigned to Group B indicated membership in GNCC in addition to having a Masters' or Doctorate in nursing

and employment in a genetic setting. Subjects assigned to Group C indicated nursing degrees that were not a Masters' or Doctorate in nursing but were employed in a genetic setting. It is of interest that 12 of the subjects in Group C indicated that they had a Masters' degree in genetic counseling. Subjects assigned to Group D indicated a Masters' or Doctorate degree in nursing but were not employed in a genetic setting. Subjects assigned to Group E indicated nursing degrees that were not a Masters' or Doctorate and were not employed in a genetic setting (Table 8). These specific groups were needed in order to carry out divergent validity testing.

Table 8

Comparison of Sample Group Size by Setting and Education (N = 356)

Genetic Setting	Potential Sample Size	Actual Sample Size
Group A: MSN	> 200	57
Group B: MSN with APNG	32	5
Group C: Non-MSN	> 600	33
Non-Genetic Setting	**Potential Sample Size**	**Actual Sample Size**
Group D: MSN	>377,000	169
Group E: Non-MSN	> 2,900,000	92

Sample size

When constructing a test, in order to have a sufficient sample for conducting an item analysis, it is desirable to obtain at least three subjects for each item (Erickson, Duffy, Gibbons, Fitzmaurice, Ditomassi, & Jones, 2004; Linn & Gronlund, 2005; Tinsley

& Tinsley, 1987). Because the instrument developed in this study contained 100 items, the sample size of 356 participants was appropriate for item analysis. Given that genetic nursing is a relatively new specialty with only 32 credentialed advanced practice nurses in genetics (APNGs) and less than 1000 genetic nurses, of any educational degree, nationally (Jenkins & Lea, 2005), the pool of potential research subjects was relatively small.

Demographics of the sample

A limited number of demographic questions were asked due to the length of the actual CRMTG. Not all of the respondents answered all of the demographic questions. The sample's average age was 44.2 years which closely represents the average age of nurses in the United States (46.8 years) (HRSA, n.d.). Group C was the youngest group and was 10 years younger than the average study participant. Group B was the oldest with an average age of 49.2 years.

Not surprisingly, members of Group B had the most years of overall nursing experience (26.80) and Group C had the least (5.68). Accordingly, Group B had the most years of genetic nursing experience (13.60) while Group C, again, had the least (3.29). The hours per week of genetic practice reported by the study respondents followed the same pattern. Group B nurses worked over six hours per week more than Group C nurses (Table 9).

The study's sample reflects the predominantly female gender of nurses in the United States, where only 5.7% of licensed Registered Nurses are male (HRSA, n.d.).

Table 9

Demographic Profile of Study Sample: *Range, Means, and Standard Deviation

	Group A (n = 57)	Group B (n = 5)	Group C (n = 33)	Group D (n = 169)	Group E (n = 92)	Total (N = 356)
Age Range (n = 295)	24-60	44-62	23-68	26-74	20-81	20-81
Age Means	38.77 (SD = 13.51)	49.20 (SD = 7.43)	34.06 (SD = 12.42)	47.69 (SD = 8.06)	45.57 (SD = 13.14)	44.24 (SD = 11.05)
Nursing Experience Range (yrs) (n = 291)	0-37	21-41	0-40	3-49	0-41	0-49
Nursing Experience Means (yrs)	12.03 (SD = 13.51)	26.80 (SD = 8.32)	5.68 (SD = 12.04)	23.62 (SD = 8.42)	20.00 (SD = 11.95)	19.28 (SD = 12.10)
Genetic Experience Range (yrs) (n = 208)	0-30	7-28	0-25			0-30
Genetic Experience Means (yrs)	5.53 (SD = 7.29)	13.60 (SD = 8.33)	3.29 (SD = 6.97)			2.44 (SD = 5.68)
Genetic Practice Range (hrs/wk) (n = 213)	0-80	20-50	0-45			0-80
Genetic Practice Means (hrs/wk)	31.59 (SD = 15.22)	34.00 (SD = 13.42)	27.69 (SD = 15.80)			13.36 (SD = 18.12)

*unless otherwise indicated

Over 98% of the respondents to this study were female. The sample included only 5 males, representing 1.7% of the total sample.

The largest number of study respondents were members of NACNS and NSGC. Group A members were predominantly from NSGC and ISONG. All of the participants in Group B were from GNCC. The majority of Group C and E respondents were contacted through NSGC. Group D respondents were primarily from NACNS (Table 7).

Nationally, diploma nurses represent 17.5 % of the nursing population. Associate degrees are held by 33.7 % and baccalaureate degrees by 34.2 %. Only 13 % hold a Masters' or Doctorate degree (HRSA, n.d.). In contrast, 56.7 % of study respondents were Masters' prepared nurses.

The bulk of Groups A, B and D were nurses with a Masters' degree in nursing. In Groups C and E most of the respondents were at the Baccalaureate degree in nursing level of education; however, many participants in these two groups did not respond to this item on the demographic section of the survey instrument. When no education level was identified, the participant was placed in the less educated group by default (Table 10).

The largest numbers of respondents in all groups worked in the clinical role. Even among the Masters' and Doctorally prepared respondents, the primary practice role was identified as clinical. A large percentage of Group E members did not identify a primary practice area (Table 11).

The respondents represented a wide variety of clinical practice areas. Many of the respondents indicated "other" for their practice area. The open-ended "other" responses included the areas identified in Table 12. Additional responses included:

Table 10

Distribution of Respondents by Nursing Degree and Group

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Group D (n = 169)		Group E (n = 92)		Total (N = 356)	
	n	%	n	%	n	%	n	%	n	%	n	%
Diploma	-	-	-	-	-	-	-	-	4	4.3	4	1.1
Associate	-	-	-	-	-	-	-	-	6	6.5	6	1.7
BS	-	-	-	-	8	24.2	-	-	24	26.1	32	9.0
MS	50	87.7	5	100.0	-	-	147	87.0	-	-	202	56.7
Doctorate	7	12.3	-	-	-	-	22	13.0	-	-	29	8.1
Missing	-	-	-	-	25	75.8	-	-	58	63.0	83	23.3

Table 11

Distribution of Respondents by Primary Practice Role and Group

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Group D (n = 169)		Group E (n = 92)		Total (N = 356)	
	n	%	n	%	n	%	n	%	n	%	n	%
Admin.	2	3.5	-	-	3	9.1	13	7.7	1	1.1	18	6.1
Research	5	8.8	-	-	2	6.1	12	7.1	1	1.1	19	6.4
Education	7	12.3	-	-	1	3.0	42	24.9	6	6.5	66	22.3
Clinical	43	75.4	5	100	24	72.7	102	60.4	27	29.3	193	65.2
Missing	-	-	-	-	3	9.1	-	-	57	62.0	60	16.9

Table 12

Distribution of Respondents by Clinical Practice Area and Group

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Group D (n = 169)		Group E (n = 92)		Total (N = 356)	
	n	%	n	%	n	%	n	%	n	%	n	%
Med/Surg	-	-	-	-	1	3.0	42	24.9	5	5.4	48	16.2
OB	12	21.1	-	-	5	15.2	10	5.9	3	3.3	30	10.5
Peds	10	17.5	2	40.0	5	15.2	11	6.5	1	1.1	29	10.8
ICU	-	-	-	-	-	-	22	13.0	4	4.3	26	8.8
ED	-	-	-	-	-	-	9	5.3	3	3.3	12	4.1
Psych	-	-	-	-	-	-	8	4.7	1	1.1	9	3.0
Family	1	1.8	-	-	1	3.0	5	3.0	1	1.1	8	2.7
OR	-	-	-	-	-	-	4	2.4	2	2.2	6	2.0
Other:	34	59.6	3	60.0	19	57.6	56	33.1	15	16.3	127	41.9
Hematology/ Oncology	15	26.3	1	20.0	2	6.1	16	9.5	1	1.1	35	9.8
Community and public health	2	3.5	-	-	3	9.1	7	4.1	2	2.2	14	3.9
Genetic counseling	4	7.0	-	-	10	30.3	-	-	-	-	14	3.9
Genetics	7	12.3	2	40.0	3	9.1	-	-	-	-	12	3.4
Cardiac	2	3.5	-	-	-	-	5	3.0	1	1.1	8	2.2
Gerontology	-	-	-	-	-	-	6	3.6	1	1.1	7	2.0
Missing	-	-	-	-	2	6.1	2	1.2	57	62.0	61	17.1

neurology/neuroscience (n = 5); prenatal (n = 5); neonatal (n = 4); gastrointestinal (n = 4); and wound/ostomy/infusion (n = 3). There were two individuals each in categories of organ transplant, ambulatory, genetics education, clinical informatics, women's health, hospice/palliative care, and endocrine/diabetes; and one each in internal medicine, orthopedics, clinical lab, critical care, immediate care center, disabled adults, pharmaceutical industry, midwifery, maternal-fetal medicine, hepatology, and surgical and post-anesthesia care. Some respondents indicated more than one practice area.

Respondents represented a variety of geographic regions in the United States with most of the respondents coming from the northeast (Table 13). Regarding race, most respondents were Caucasian (Table 14). Over 16% of the respondents did not indicate their race.

Table 13

Distribution of Respondents by Geographic Region and Group

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Group D (n = 169)		Group E (n = 92)		Total (N = 356)	
	n	%	n	%	n	%	n	%	n	%	n	%
NE	19	33.3	2	40	15	45.5	48	28.4	22	23.9	106	35.6
Central	12	21.1	3	60	4	12.1	50	29.6	7	7.6	76	25.5
SW	8	14.0	-	-	8	24.2	24	14.2	4	4.3	44	15.1
SE	10	17.5	-	-	3	9.1	28	16.6	-	-	41	13.4
NW	7	12.3	-	-	2	6.1	19	11.2	3	3.3	31	10.4
Missing	1	1.8	-	-	1	3.0	-	-	56	60.9	58	16.3

Table 14

Distribution of Respondents by Race and Group

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Group D (n = 169)		Group E (n = 92)		Total (N = 356)	
	n	%	n	%	n	%	n	%	n	%	n	%
Caucasian	50	87.7	4	80.0	29	87.9	161	95.3	35	38.0	279	78.4
Asian	4	7.0	-	-	2	6.1	2	1.2	-	-	8	2.2
African-American	1	1.8	-	-	-	-	3	1.8	-	-	4	1.1
Hispanic	2	3.5	1	20.0	-	-	-	-	-	-	3	0.8
Other (Ashkenazi Jewish)	-	-	-	-	2	6.1	-	-	-	-	2	0.6
Native American	-	-	-	-	-	-	1	0.6	-	-	1	0.3
Missing	-	-	-	-	-	-	2	1.2	57	62.0	59	16.6

The respondents who worked in a genetic setting indicated a variety of genetic clinical practice areas consistent with the clinical practice areas represented on the CRMTG test blueprint (Tables 3 & 15). Respondents that selected "other" provided open-ended results as shown in Table 15. Fifteen of the open-ended responses indicated that the respondent did not work in genetics. For the CRMTG, reproductive genetics included prenatal and newborn screening, and adult genetics included oncology and community/ public health. Genetic counseling was included on the CRMTG as a practice component,

as indicated in the ISONG Scope and Standards (1998), rather than a clinical practice area.

Table 15

Distribution of Respondents by Primary Genetic Clinical Practice Area and Group (N=356)

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Total Genetic Sample (n = 95)	
	n	%	n	%	n	%	n	%
Pediatric	13	22.8	3	60.0	10	30.3	26	27.4
Reproductive	12	21.1	-	-	7	21.2	19	20.0
Adult	10	17.5	-	-	3	9.1	13	13.7
Other	22	38.6	2	40.0	13	39.4	37	38.9
Oncology	13	22.8	1	-	5	15.2	19	20.0
Prenatal	4	7.0	1	-	3	9.1	8	8.4
Community/ public health	2	3.5	-	-	1	3.0	3	3.2
General genetics	2	3.5	-	-	1	3.0	3	3.2
Genetic counseling	-	-	-	-	2	6.1	2	2.1
Newborn screening	-	-	-	-	1	3.0	1	1.1
Clinical laboratory	1	1.8	-	-	-	-	1	1.1

The respondents who worked in a genetic setting indicated a variety of genetic educational preparations (Table 16). Respondents indicating "other" provided open-ended responses of: National Institute of Nursing Research Summer Genetics Institute (n = 6); one-day classes (n = 4); and, Oncology Nurses Society Congress Workshop (n = 1). It is interesting to note that none of the members of Group B had formal genetic coursework or degrees. They all were prepared for their roles in genetics through on-the-job training. It is also of interest that a majority of the members of Group C, having been categorized as such based on their, Bachelors' or less, nursing degree, had a supplemental degree in genetics. Twelve members of Group C had a Masters' degree in genetics.

Table 16

Distribution of Respondents by Genetic Education and Group (N=356)

	Group A (n = 57)		Group B (n = 5)		Group C (n = 33)		Total Genetic Sample (n = 95)	
	n	%	n	%	n	%	n	%
Degree in Genetics	28	49.1	-	-	22	66.7	50	52.6
On-the-job Training	17	29.8	5	100	2	6.1	24	25.3
Continuing Education	3	5.3	-	-	1	3.0	4	4.2
Certification Program	2	3.5	-	-	-	-	2	2.1
College Courses	1	1.8	-	-	-	-	1	1.1
Other	6	10.5	-	-	5	15.2	11	11.6
Missing	-	-	-	-	3	9.1	3	3.2

Protection of Human Subjects

The proposal for this study was reviewed and approved for protection of human subjects by the Widener University Institutional Review Board (Appendix N). The deans/research directors of the schools/professional organizations whose students/ members were invited to participate in the study also approved the study. There were no foreseeable risks to the subjects. Demographic data were used to describe the sample and for divergent validity purposes. Participation in the study was voluntary. All participants had the right to refuse or withdraw once logged onto the website or anytime during the test-taking process prior to submitting the completed instrument, without penalty. While participants received no compensation for their participation in the study, those who completed the survey twice to provide test-retest data were invited to enter a drawing for a gift certificate. Participants' scores were matched using the email addresses that they supplied on the opening page of the website. Because nurses have been found to be difficult to recruit for Internet studies, as an incentive, upon completion of the retest, the test-retest participant email addresses were entered into a drawing. Four $100 gift certificates redeemable at amazon.com were sent electronically to the four, randomly selected, winners. The drawing was held independently from the scoring of the instrument. The Institute for Credentialing Innovation of the American Nurses Credentialing Center funded the drawing.

Subjects were given the option to receive a summary of the study results by contacting the investigator by email. This information was included in the invitation to participate and the letter of explanation (Appendices L, M, O, P, & R).

Once the data were downloaded into the database by Surveymonkey © it was not possible to link participant identifiers with their data. Raw data has been stored electronically on a personal computer in the investigator's home. The data will be kept for seven years, and then, any hard copies of the study data will be destroyed.

The opening pages of the website included: 1) Study information and withdrawal option (Appendix R); 2) an optional request for email addresses to be used for re-contact, test-retest tracking, and prize drawing participation; 3) the demographic form (Appendix Q); and 4) the Criterion-Referenced Measurement Tool for Genetics (CRMTG) (Appendix K). The demographic form and the CRMTG responses were coded for group assignment. The investigator and the dissertation chair were the only people to view the email addresses in order to protect participant identities.

Participant enrollment and Internet survey procedure

Once approval was obtained (Appendix N), an explanation of the study with an invitation to participate was circulated via organizational procedure (Table 6) to all members of all five groups. The text of the initial email requests (Appendices L, M, O, & P) has a Flesch Reading Ease of 48.9 and is at Flesch-Kincaid Grade Level 9.6 (Bastable, 1997) as calculated using Microsoft Word ® 2000 software. The email request for this study had a higher reading level and lower reading ease value than recommended for public documents; however, the audience for these emails was registered professional nurses. The primary audience was advanced practice nurses, who would have earned, at least, a Masters' degree.

These email requests included information for access to the final version of the Criterion-Referenced Measurement Tool for Genetics (CRMTG) via the Internet. The instrument was available on-line via Surveymonkey © from June through December 2005 (Appendix S). A total of 17,372 potential participants were contacted through a variety of electronic methods. Subjects were given instructions for secure access to the test through Surveymonkey © netware. The first page (Appendix R) of the website included the recruitment and consent information. This page of the website provided the participants an explanation of the study, retest process, criteria for the gift certificate drawing, process for withdrawal from the study, and selection of "submit" upon completion of the instrument indicating consent to participate. A total of 356 participants logged onto the CRMTG website.

Individuals who chose to participate clicked "next" to begin the instrument. Subsequent pages were set up to include the demographic questions (Appendix Q) and the CRMTG. The CRMTG pages were formatted for ten stand-alone items per page to minimize an extremely long scroll. The case study items were formatted to begin with the scenario and then the associated items followed on the same page to allow participants to scroll back to the top of the page for review of the scenario.

One month after the completion of the CRMTG, participants were contacted via email and asked to retest (Appendix T).

Stage 8. Assess reliability and validity of measure

Validity and reliability of the CRMTG was assessed using a variety of methodologies (Table 17). Each type of reliability and validity is addressed in Chapter Four.

Table 17

Psychometric Data Analysis

Methodology	Type of Reliability	Type of Validity
Item-to-total correlation	Internal consistency	
Paired t-test	Over time	
Pilot Study		Content
Expert Review		Content & Criterion-referenced
Item Analysis		Construct
Chi square		Divergent

Summary

The Waltz and colleagues (2005) framework was used to develop a 100-item multiple-choice examination, based on published genetic nursing care standards, for the purpose of measuring the genetic knowledge of advanced practice nurses in genetics. The examination was based on criteria derived from the International Society of Nurses in Genetics, Inc. (ISONG) Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998). A panel of experts was used to establish content and construct

validity and determine the cut-score for the instrument. The study was reviewed and approved for human subjects by the Widener University Institutional Review Board (Appendix N) and by the research directors of the schools and professional organizations whose students/members were contacted to participate in the study. The instrument was administered, via the Internet, to Masters' prepared nurses in genetics (n = 57), APNG certified advanced practice nurses in genetics (n = 5), non-Masters' prepared genetic nurses (n = 33), non-genetic advanced practice nurses (n = 169) and non-genetic, non-Masters' nurses (n = 92).

CHAPTER FOUR: RESULTS

Introduction

In this chapter the study results are discussed. Following a summary of the descriptive statistical analyses for CRMTG scores, the results are organized based on the research questions for this study. Psychometric analysis of the Criterion-Referenced Measurement Tool for Genetics (CRMTG) included: item-to-total analysis, test-retest reliability, expert review, pilot administration, item analysis, and Chi square.

Descriptive Analysis of CRMTG Scores

Three hundred fifty-six participants logged onto the Surveymonkey © site for the CRMTG. Scores on the CRMTG ranged from 0 to 90. One hundred seventeen participants scored a 0 on the CRMTG. Skipped items were scored as incorrect. The mean score on the CRMTG was 31.56 with a standard deviation of 28.92. The median score was 33.50 with a mode of 0. The interquartile range was 0 to 50.00. The distribution of scores was negatively skewed (-.34) with kurtosis of −1.18.

Research Question 1

The first research question was: To what extent is the Criterion-Referenced Measurement Tool for Genetics (CRMTG) a valid and reliable indicator of knowledge of advanced practice nursing in genetics? Psychometric data were collected and analyzed in order to provide validity and reliability information on the CRMTG (Table 17). Content, criterion-referenced, construct, and divergent validity were also evaluated for the CRMTG. Reliability was calculated to evaluate internal consistency and test-retest reliability over a one-month time period.

Validity

Validity is measured as a continuum of usefulness (Nunnally, 1970). Validity results are not expressed in a yes or no response. The validity measure reflects the greater or lesser degree that an instrument meets the intended use. The intended use of the CRMTG is to indicate a minimum knowledge level of advanced practice nursing in genetics. Content, criterion, construct, and divergent validity of the CRMTG were evaluated.

Content validity

Content validity measures the extent to which an instrument measures what it is supposed to measure and the accuracy and relevance of the information included in the exam (Cronbach, 1971; Gronlund, 1973; Psychometrics, 2004; Waltz, et al., 2005). Expert reviewers and pilot participants provided feedback on each item included in the original test bank. The average congruency percentage was calculated between expert and pilot participants. Based on their feedback regarding item difficulty, accuracy and relevance to advanced practice nursing in genetics, content validity of each item included in the final version of the CRMTG was determined. Each item included in the final version of the CRMTG was considered relevant to advanced practice nursing in genetics knowledge by at least 80% of the content expert reviewers and pilot study participants. Content areas for each item were identified by the investigator and confirmed by the content experts (Table 3 & Appendices A & H). The content experts identified taxonomy areas for each item (Appendices A & H). The overall content validity index for the final version of the CRMTG was calculated to be .88.

Criterion-related validity

Criterion-related validity measures the ability of an instrument to meet the established function (Waltz et al., 2005). Expert reviewers were asked to evaluate each item against the test blueprint (Gessaroli & Poliquin, 1995) (Table 3). The criterion-related validity of the instrument as a whole was supported through calculation of interrater agreement among experts. A congruency of 80% was the criterion of inclusion for items in the CRMTG (Latvala, 2002). The expert reviewers were asked if they agreed with the placement of each item on the test blueprint as it related to the content areas of general, reproductive, adult, and pediatric genetics and Bloom's Taxonomy to provide validity evidence for the instrument (Hair, Anderson, Tatham, & Black, 1998; Munro, 2004) (Tables 3 & 5).

Construct validity

Construct validity measures the ability of the instrument to function as a measure of an intended theoretical construct (Cronbach, 1971; Psychometrics, 2004; Reliability and Validity, n.d.; Streiner & Norman, 2003; Williamson, 1999). Construct validity was evaluated using item analysis. Item analysis results provided information regarding the reliability, validity, and objectivity of this instrument in measuring knowledge of advanced practice nursing in genetics (Van Ort & Hazzard, 1985).

The level of difficulty for each item was calculated along with the discrimination index (Guilford, 1954; Linn & Gronlund, 2005) (Appendix U). Item difficulty is computed by dividing the number of participants who answered the item correctly by the total number of participants. Item difficulty is reported as the percentage of participants

who answered the item correctly (Linn & Gronlund, 2005). CRMTG item difficulty values ranged from 1.60 to 96.40% (M = 54.39%).

Discrimination index reflects the difference between the number of participants in the upper one third of the scores and the participants in the lower third who answered the item correctly. A quality item would have most of the high scoring subjects answering correctly and most of the low scoring subjects selecting an incorrect response. If all of the participants in the upper group answer correctly and all of the participants in the lower group answer incorrectly, the discrimination index would be a perfect 1.00 (Linn & Gronlund, 2005). CRMTG item discrimination index values ranged from -.01 to .38 (M = .22).

Divergent validity.

Divergent validity is used to demonstrate that there is not a strong correlation when responses of dissimilar groups are compared (New Target, 2005; van Oppen, 1992). If groups differ by important demographic criteria, they should have different responses to the instrument. The divergent validity for the CRMTG was analyzed using the nursing education and practice-setting demographic information of the study participants. The participants were divided into five groups for the purpose of calculating divergent validity (Table 18). Chi square analysis was used to evaluate the dichotomous pass/fail outcomes of the five sample groups (Ferguson, 1976; Lewin, 1987). A significant association was found ($\chi^2 = 225.29$, $df = 4$, $p < .001$).

A chi-square test of independence was calculated comparing the pass/fail rate for genetic (combined Groups A, B, & C) and non-genetic nurses (combined Groups D & E).

Table 18

CRMTG Pass Rates by Group* (N = 356)

	Group A	Group B	Group C	Group D	Group E	Total
Passed	30	3	21	7	2	63
Failed	27	2	12	162	90	293
Total	57	5	33	169	92	356
Pass Rate	52.6%	60%	63.6%	4.1%	2.2%	17.7%

* Group A – genetic nurses with a Masters' or Doctorate in nursing

Group B – Masters' prepared APNGs certified through the GNCC portfolio process

Group C – genetic nurses without a Masters' in nursing

Group D – non-genetic nurses with a Masters' or Doctorate in nursing

Group E – non-genetic nurses without a Masters' in nursing

Due to low values in some cells, the Yates' correction for continuity was used for calculation of chi-square ($\chi^2 = 132.69$, $df = 1$, $p < .001$). A significant association was found (Table 19). These analyses indicate that, in this study, the CRMTG clearly discriminated between nurses with a genetic background and those without a background in genetic nursing.

A one-way ANOVA was computed comparing the CRMTG raw scores of participants of all five groups. A significant difference was found among groups ($F(4,351) = 36.34$, $p < .001$). Tukey's honestly significant difference comparison with harmonic mean was used to determine the nature of the differences among groups

Table 19

Contingency Table Showing CRMTG Pass/Fail According to Genetic Background (N = 356)

	Genetic (n = 95)	Non-genetic (n = 261)	χ^2	df	p
Number Passed	54 (56.84%)	9 (3.45%)			
Number Failed	41 (43.16%)	252 (96.55%)	132.69	1	<.001

(Table 20). Pairwise comparisons demonstrated that each group of nurses working in a genetic setting (Groups A, B, & C) scored significantly higher than each group of nurses who did not work in a genetic setting (Groups D & E). However, there was no significant difference between the raw scores on the CRMTG for Masters' in nursing prepared genetic nurses (Groups A & B) as compared to the raw scores on the CRMTG for non-Masters' in nursing prepared genetic nurses (Group C).

Table 20

Tukey Analysis of Mean Difference (Standard Error) and Significance Between Groups

	Group A (n = 57)	Group B (n = 5)	Group C (n = 33)	Group D (n = 169)
Group B	-7.23 (11.41) (p = .969)			
Group C	-2.06 (5.35) (p = .995)	5.18 (11.74) (p = .992)		
Group D	26.37 (3.75) (p < .001)	33.60 (11.10) (p = .022)	28.42 (4.66) (p < .001)	
Group E	40.95 (4.12) (p < .001)	48.18 (11.23) (p < .001)	43.00 (4.96) (p < .001)	14.58 (3.17) (p < .001)

A one-way ANOVA comparing the raw scores of genetic (combined Groups A, B, & C) and non-genetic nurses (combined Groups D & E) was also computed. This analysis revealed that participants who had genetic experience scored significantly higher than participants who did not have genetic experience ($F(1,354) = 117.52$, $p < .001$).

Reliability

The CRMTG was evaluated for internal consistency and test-retest reliability over time. Internal consistency was evaluated using item-to-total analysis. Each test item was evaluated for consistency with all of the other items and consistency with the total score of each participant. Pearson item-to-total correlations ranged from .26 to .81 and was significant at the .001 level (2-tailed) for 99 of the 100 CRMTG items. Item 76 ($r = .01$) was the only exception (Appendix U). The Cronbach's alpha for the CRMTG was calculated to be .99. This value did not change when item 76 was removed from the CRMTG.

A correlational coefficient was used for test-retest comparison. Participants of the study were asked to repeat taking the CRMTG one month after they initially completed the CRMTG. Individual raw initial and repeat scores were compared using a paired t-test to measure stability over time (Lackman et al., 1997; Lewin, 1987; Miles, Penny, Power, & Mercey, 2003). A total of 36 participants participated in the retest. Scores were linked using optional email information supplied on a screen prior to the demographic screen on the instrument website. The mean score for the test was 43.47 (28.30) and the retest mean was 42.53 (30.31) for this subgroup of participants (Appendix V). Using the total score of the first administration correlated to the total score of the second administration for those

36 participants who responded to the retest request, the test-retest Pearson correlation was calculated to be .86 ($p < .001$). Paired t-test calculations revealed no significant difference ($t = -1.68$, df 35, $p = .101$) between initial and repeat test scores.

Research Question 2

The second research question was: Do nurses who practice in a genetic setting have a higher pass rate on the CRMTG than nurses who do not practice in a genetic setting? As described above, a chi-square test of independence was calculated comparing the pass/fail rate for genetic (combined Groups A, B, & C) and non-genetic nurses (combined Groups D & E) ($\chi^2 = 132.69$, $df = 1$, $p < .001$) (Table 19). Due to the low values in some cells, the Yates' correction for continuity was calculated. A one-way ANOVA comparing the raw scores of genetic and non-genetic nurses was also computed ($F(1,354) = 117.52$, $p < .001$). These analyses indicate that, in this study, the nurses who practiced in a genetic setting had a significantly higher pass rate on the CRMTG than nurses who did not practice in a genetic setting.

Research Question 3

The third research question was: Do advanced practice nurses have a higher pass rate on the CRMTG than nurses who do not have their Masters' degree in nursing? A chi-square test of independence was calculated comparing the pass/fail rate for participants with nursing Masters' (combined Groups A, B, & D) and those nurses without a Masters' in nursing (combined Groups C & E). Due to the low values in some cells, the Yates' correction for continuity was calculated ($\chi^2 = .01$, $df = 1$, $p = .912$). There was no significant difference in pass/fail outcomes when all nursing Masters' prepared nurses

(advanced practice nurses) were compared to those nurses who did not have a Masters' degree in nursing.

A one-way ANOVA comparing the raw scores of these nurses was also computed ($F(1,354) = 10.90$, $p = .001$). Tukey's honestly significant difference comparison with harmonic mean was used to compare Group A with Group C and Group D with Group E (Table 20). No significant difference was found when comparing the raw scores of genetic nurses with a Masters' degree in nursing and those without ($MD = -2.06$ (5.35), $p = .995$). A significant difference was found when comparing the raw scores of non-genetic nurses with a Masters' degree in nursing and those without ($MD = 14.58$ (3.17), $p < .001$). These analyses provide inconclusive results on the question of whether advanced practice nurses have a higher pass rate on the CRMTG than nurses who did not have a Masters' degree in nursing.

When groups were compared on advanced practice specific content areas on the CRMTG, some variability in the percent of correct responses was found (Table 21). Members of Group C performed better on most of the genetic counseling items than the members of Groups A and B. No group trends were seen on the case coordination items. The genetic credentialed nurses, Group B, performed better on most of the teaching and research items than the members of Groups A and C.

Research Question 4

The fourth research question was: Do nurses credentialed by the GNCC have a higher pass rate on the CRMTG than non-credentialed nurses? Nurses credentialed by GNCC (Group B) had a slightly higher pass rate (60%) on the CRMTG than non-

Table 21

Comparison of Advanced Practice Specific Item Percent of Correct Responses by Group

	Item number	Group A (n = 57)	Group B (n = 5)	Group C (n = 33)
Genetic counseling	10	5%	20%	42%
	11	42%	40%	49%
	36	60%	60%	73%
	45	51%	60%	61%
	60	25%	20%	33%
	62	70%	60%	67%
	78	28%	0%	39%
	99	32%	20%	27%
Case coordination	61	46%	40%	55%
	69	42%	60%	49%
	76	0%	0%	0%
	98	21%	0%	12%
Teaching and research	17	40%	20%	64%
	19	63%	100%	61%
	20	75%	100%	73%
	25	68%	100%	70%
	34	51%	20%	58%
	35	68%	100%	70%

	Item number	Group A	Group B	Group C
Teaching and research (con't)	41	54%	60%	58%
	67	46%	60%	49%
	94	32%	20%	39%

credentialed nurses with similar educational and clinical backgrounds (Group A) (52.6%). A chi-square test of independence was calculated comparing the pass/fail rate for GNCC credentialed nurses (Group B) and non-credentialed nurses with comparable educational and clinical backgrounds (Group A) ($\chi^2 = .00$, $\underline{df} = 1$, $\underline{p} = 1.000$). Due to the low values in each cell, the Yates' correction for continuity was calculated. A one-way ANOVA comparing the raw scores of the members of these groups was also computed ($\underline{F}(1,60) = .27$, $\underline{p} = .604$). No significant differences were found when comparing the pass/fail outcomes or raw scores of genetic nurses with a Masters' degree in nursing and those with the APNG credential.

Research Question 5

The fifth research question was: Are GNCC portfolio pass/fail outcomes congruent with CRMTG pass/fail outcomes? The five study participants in Group B were all successful portfolio applicants to GNCC. A 100 % pass rate would be anticipated for such an elite group of credentialed specialists; however, the pass rate on the CRMTG for the credentialed nurses was only 60 % (Table 18). Due to the anonymity of the data, there was no way to determine if any unsuccessful portfolio applicants responded to the

invitation to participate the study. Unsuccessful applicants may have been contacted through other genetic professional organizations but they would not have been members of GNCC. The demographic form did not include any indication of GNCC portfolio submission. Therefore, this question cannot be answered at this time.

Corrections for Missing Data

Due to the length and difficulty of the CRMTG, participant fatigue and other unknown factors resulted in a lot of missing data. Skipped items were scored as incorrect on the CRMTG to be consistent with standard multiple-choice item testing practice. When the raw data were reviewed for types for missing results, the pattern of missing data supported the fatigue factor. One hundred thirty-four respondents quit after answering the demographic questions and up to 10 actual test items. Two respondents answered up to 20 test items and then quit. Eight respondents answered up to 30 test items and then quit. Nine respondents answered up to 40 test items and then quit. One respondent answered up to 50 test items and then quit. Six respondents answered up to 60 test items and then quit. One respondent answered up to 70 test items and then quit. The remaining 195 participants answered at least 85 items. For comparison, the analyses for this study were rerun using the filtered data for the 195 participants who answered at least 85% of the items on the CRMTG (Table 22). These analyses were not initially reported because with more than 5% missing values, cases should not be deleted (Data imputation for missing values, n.d.).

Table 22

Distribution of Participants and CRMTG Scores for Subsample

Group	Subsample Size	Range of Scores	Mean Score (s.d.)	Pass Rate (%)
A	41	36-89	71.12 (12.15)	73.2
B	4	50-79	68.50 (13.48)	75.0
C	24	54-90	77.04 (9.62)	87.5
D	100	20-75	43.46 (11.65)	7.0
E	26	24-81	43.19 (13.06)	7.7

Descriptive Analysis of CRMTG Subsample Scores

For the 195 participants logged onto the Surveymonkey © site for the CRMTG who completed at least 85% of the items, the scores on the CRMTG ranged from 20 to 90. The mean score for this subsample on the CRMTG was 53.89 with a standard deviation of 18.45. The median score was 48.00 with a mode of 42. The interquartile range was 39.00 to 73.00. The scores were positively skewed (.34) with kurtosis remaining at –1.18.

Research Question 1

Chi square analysis for divergent validity of the subsample was used to evaluate the dichotomous pass/fail outcomes of the five sample groups (Ferguson, 1976; Lewin, 1987). A significant association was found ($\chi^2 = 104.56$, $df = 4$, $p < .001$). A one-way ANOVA was computed for the subsample comparing the CRMTG raw scores of participants of all five groups. A significant difference was found among groups ($F(4,190) = 71.88$, $p < .001$). Regarding internal consistency and reliability, Cronbach's

alpha for the CRMTG was calculated for this subsample to be .95. The results for question one remained similar to those for the original analysis.

Research Question 2

A chi-square test of independence was recalculated comparing the pass/fail rate for genetic (combined Groups A, B, & C) and non-genetic nurses (combined Groups D & E) in the subsample. Due to the low values in some cells, the Yates' correction for continuity was calculated ($\chi^2 = 99.88$, $df = 1$, $p < .001$). A significant association was found. A one-way ANOVA comparing the subsample raw scores of genetic (combined Groups A, B, & C) and non-genetic nurses (combined Groups D & E) was also computed. This analysis revealed that participants in the subsample who had genetic experience scored significantly higher than participants who did not have genetic experience ($F(1,193) = 280.89$, $p < .001$). The results for question two remained similar to those for the original analysis.

Research Question 3

A chi-square test of independence was calculated comparing the pass/fail rate for subsample participants with nursing Masters' (combined Groups A, B, & D) and those nurses without a Masters' in nursing (combined Groups C & E). Due to the low values in some cells, the Yates' correction for continuity was calculated ($\chi^2 = 4.95$, $df = 1$, $p = .026$). There was a significant difference in pass/fail outcomes when all nursing Masters' prepared nurses (advanced practice nurses) were compared to those nurses who did not have a Masters' degree in nursing for the subsample, whereas this difference had not been significant in the overall analysis reported earlier. A one-way ANOVA comparing the raw

scores of these nurses was also computed ($F(1,193) = 6.26$, $p = .013$). These analyses revealed that participants in the subsample who were advanced practice nurses did have a significantly higher pass rate on the CRMTG than nurses who did not have a Masters' degree in nursing.

Research Question 4

A chi-square test of independence using the Yates correction for continuity was calculated comparing the pass/fail rate for GNCC credentialed nurses (Group B) and non-credentialed nurses with comparable educational and clinical backgrounds (Group A) ($\chi^2 = .00$, $df = 1$, $p = 1.000$) in the subsample. A one-way ANOVA comparing the raw scores of the members of these groups was also computed ($F(1,43) = .17$, $p = .685$). Similar to results of analyses of the entire sample, no significant differences were found when comparing the pass/fail outcomes or raw scores of genetic nurses with a Masters' degree in nursing and those with the APNG credential.

Summary

The CRMTG was analyzed for validity and reliability using a sample of 356 nurses. Content and criterion-related validity were established using content experts and pilot study data. The CRMTG content validity index was calculated to be .88. Construct validity was evaluated using item analysis including level of difficulty which ranged from 1.60 to 96.40 % ($M = 54.39\%$) and discrimination index values which ranged from -.01 to .38 ($M = .22$) for each item. Divergent validity was evaluated using Chi square ($\chi^2 = 225.29$, $df = 4$, $p < .001$) and ANOVA ($F(4,351) = 36.34$, $p < .001$) analyses to compare among groups of nurses with different educational preparations and different clinical

practice areas. Reliability was evaluated using item-to-total and test-retest correlational analyses. The Cronbach's alpha for the CRMTG was calculated to be .99. Based on the data analysis for this study, the Criterion-Referenced Measurement Tool for Genetics (CRMTG) was found to be a valid and reliable indictor of knowledge of advanced practice nursing in genetics. It was also found that nurses who practice in a genetic setting had a significantly higher pass rate on the CRMTG than nurses who do not practice in a genetic setting. Analysis of data for research questions 3 and 4 led to inconclusive results, although when the subsample of participants who answered at least 85% of the items on the CRMTG were analyzed there was a significant difference between advanced practice nurses and those nurses who did not have their Masters' degree in nursing. Research question 5 could not be addressed in this study.

CHAPTER FIVE: DISCUSSION, CONCLUSIONS, IMPLICATIONS, & RECOMMENDATIONS

Introduction

This chapter relates the results of the study to the identified research questions and theoretical frameworks of the study. Methodological issues that were encountered during the course of the study will be discussed. Conclusions, limitations, implications, and recommendations will also be presented.

Discussion

Research Question 1

The first research question was: To what extent is the Criterion-Referenced Measurement Tool for Genetics (CRMTG) a valid and reliable indicator of knowledge of advanced practice nursing in genetics? The CRMTG was found to have overall validity and reliability based on this study. Validity refers to the degree that an instrument is or is not useful for the intended purpose (Nunnally, 1970). This study evaluated the CRMTG for content, criterion-related, construct and divergent validity (Table 17). Reliability measures the consistency and accuracy of an instrument. This study evaluated the CRMTG for internal consistency and consistency over time.

Content validity

Content validity measures the accuracy and relevancy of the information tested (Flynn & Reese, 1988; Gronlund, 1973; Waltz et al., 2005). Content validity also refers to the accuracy of the key. Each item should have only one clearly correct answer and the distracters should be clearly incorrect. The CRMTG was found to have content validity,

as determined by the pilot study and content expert reviewers (Appendices E, G, H, I, & J). There was continuity between the pilot study and expert comments related to the test items; however, the number of pilot study participants was very small. Item analysis of the study data found that 21 of the items on the final version of the CRMTG did not have the key as the most frequently selected answer in this study (Appendices B, J, & U). All five content experts agreed upon the key for each item. These items will need to be reviewed and rewritten for future versions of the CRMTG to ensure greater accuracy.

The most challenging element of the construction and revision of CRMTG items was creating multiple-choice test items that were sufficiently difficult for advanced practice nurses. "Multiple-choice items have been criticized on the grounds that they tend to test only at the knowledge or recall level" (Waltz et al., 2005, p. 290). The multiple-choice format was selected in order to be consistent with the standard for credentialing examinations. Case studies were included in order to increase the cognitive level of CRMTG items. The case study items scored well when analyzed for content validity. The difficulty in writing, analyzing, and maintaining a credentialing examination was one of the factors considered by the GNCC in selecting the portfolio process in lieu of an examination (Monsen et al., 2005).

Another method of assessing construct validity is factor analysis. Factor analysis would assess the links between the concepts being measured and items on the instrument and could be used to validate the test blueprint. Factor analysis is based on interval or ratio data (Comrey & Lee, 1992; Froman, 2001; Kim & Mueller, 1978). Because the

CRMTG consists of multiple-choice items that produce dichotomous data, a factor analysis could not be computed for this type of instrument.

Criterion-related validity

Criterion-related validity measures the degree that an instrument measures a concept as compared to established criteria (Cronbach, 1971; Williamson, 1999). Reviewers evaluated the items of the CRMTG for consistency with the criteria for advanced practice genetic nurses described in the International Society of Nurses in Genetics, Inc. (ISONG) Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998). The CRMTG was found to have criterion-related validity, as determined by the content expert reviewers.

Construct validity

Construct validity evaluates whether an instrument measures a defined concept or theoretical construct (Linn & Gronlund, 2005, Streiner & Norman, 2003, Waltz et al., 2005). Content validity evidence supports construct validity, but is not the only evidence used to determine construct validity (Waltz et al., 2005). "The major focus of construct validation for criterion-referenced measures is to establish support for the measure's ability to accurately categorize phenomena in accordance with the purpose for which the measure [is] being used" (Waltz et al., p. 178).

An item analysis was conducted to evaluate the level of difficulty and discrimination index of each item. The item analysis results were limited based on the number of respondents who skipped items. Of the 356 total respondents in the study, each CRMTG item had between 123 and 184 respondents who skipped that item (Appendix

B). It can be noted that the number of respondents who skipped each item increased as the test continued. This increase in skipped items could be attributed to lack of interest, time, and fatigue. The Surveymonkey © site did not allow for randomization of the order of items for each participant. The number of survey participants were determined by the number of log-ins onto the CRMTG website. Some participants appeared to have logged on, began entering demographic data, and then quit the survey either prior to beginning the CRMTG or after answering some, but not all, of the items. The greatest number of drop-outs on the CRMTG were from Group E probably due to frustration at the content area and level of difficulty of the items. Due to the large number of skipped items, a subsample of participants was created excluding any participants who did not complete at least 85% of the items on the CRMTG (Table 22).

Difficulty level did not seem to be a factor in skipping items as the items with the lowest number skipped had difficulty levels of 51.5% and 56.2%, which is on the more difficult end of the acceptable range. The item with the greatest number skipped had a difficulty level of 48.8%, which is very near the acceptable range. Between 132 and 166 respondents skipped items that had a difficulty level in the 90% range. Between 145 and 173 respondents skipped items that had a difficulty level in the < 25% range.

Item level of difficulty

When evaluating the level of difficulty, the higher the difficulty percent value, the easier the item. For example, if an item is calculated as having an 82% difficulty, 82% of the individuals answering the item responded correctly and 18% were incorrect. The optimum difficulty level for a four-alternative, multiple-choice item is 62%, with an

acceptable range being from 50% to 90% (De Ayala & Kelley, 1987). On the CRMTG, 55 of the 100 items were within the acceptable range of difficulty, 5 items were very easy, and 40 items were very difficult. Twelve of the very difficult items were at or below the guessing level of 25%; therefore they may need to be rewritten or discarded for future versions of the CRMTG (Appendices U & W). When compared to the expert panel's difficulty rating the calculated item level of difficulty results were quite variable (Appendix X); however, mean values did reflect expert ratings. Items that the experts rated as "easy" were calculated to have difficulty levels ranging from 14.8 to 96.4 ($\underline{M}$ = 60.53). Expert-rated "moderate" items had difficulty levels ranging from 1.6 to 86.6 ($\underline{M}$ = 49.71). Expert-rated "difficult" items had difficulty levels ranging from 25.4 to 44.4 ($\underline{M}$ = 34.9).

The variance in the expert difficulty rating and the calculated difficulty levels might reflect the bias of the experts related to their level of genetic knowledge. Future difficulty ratings for CRMTG items should be based on the actual calculated results. This bias may also have affected the cut score. For this reason, not only was the Chi-square analysis computed using the established cut score, but the one-way ANOVA was also computed using raw scores to determine if the cut score was affecting the results of the study. ANOVA results supported Chi-square results.

Due to the large number of participants who did not respond to at least 85% of the items on the CRMTG, a subsample was analyzed. The percent of subsample participants who answered the item correctly was calculated for each CRMTG item. Appendix X displays the results, for each CRMTG item, of this calculation along with the expert

difficulty rating and item level of difficulty calculated using the whole sample. The percent of subsample participants (n = 195) who answered each item correctly closely parallels the calculated item level of difficulty based on the whole sample (N = 356).

Item discrimination index

Item discrimination index values should be greater than .20 in order to provide information about the differences in participant knowledge levels of genetics (De Ayala & Kelley, 1987). Of the 100 items on the CRMTG, 64 items had item discrimination index greater than .20, the highest being .38 (Appendix U). Only one item had a negative item discrimination index, item 76, with a value of -.01. This item has been identified as problematic and will need to be discarded.

Divergent validity

Divergent validity, conversely, evaluates the instrument for poor correlations between the scores of those groups expected to score well and those expected to score poorly. Divergent validity is analyzed by comparing and contrasting responses for diverse groups or constructs (Graziano & Raulin, 2000). Increased validity would be measured when different groups or constructs have a low correlation.

The pass rates on the CRMTG, based on the cut score of 65 as established by the expert reviewers during Stage 6, were compared and contrasted between diverse groups. The consistent findings on the chi-square, which evaluated pass/fail outcomes, and the ANOVA, which evaluated raw scores, support the relative accuracy of the computed cut score. Cut score can also be calculated mathematically, once the instrument has been tested. The cut score for an instrument is typically set at one standard deviation below the

mean (Chadima, 2006). Based on this formula, the cut score for the CRMTG would be 2.64 (M=31.56, SD=28.92), which is obviously skewed.

The genetic nurses comprising Groups A, B, and C had much higher pass rates, 52.6%, 60%, & 63.6%, respectively, than the non-genetic nurses comprising Groups D and E with pass rates of 4.1% and 2.2%, respectively (Table 18). The content experts and pilot study findings were supported in the inclusion of pertinent genetic content. In terms of the fourth question of the COPA Model (Lenburg, 1999a; Lenburg, 1999b; Redman et al., 1999), these findings support the use of the CRMTG as an effective way to document genetic nurses' knowledge.

The genetic nurses who did not have a Masters' or Doctorate in nursing, Group C, had the highest pass rate of all of the groups studied at 63.6%. It should be noted that two-thirds of these nurses, although not having a Masters' or Doctoral degree in nursing, identified themselves as having a degree in genetics which could be presumed to positively influence their knowledge of genetics. Twelve members of Group C indicated that they had a Masters' degree in genetics which clearly would have given them an advantage on the CRMTG. The GNCC credentialed nurses, Group B, had the second highest pass rate, 60% (Table 18). The non-credentialed advanced practice level nurses practicing in genetics settings had the lowest pass rate of all genetic nurses at 52.6%.

In terms of the fourth question of the COPA Model (Lenburg, 1999a; Lenburg, 1999b; Redman et al., 1999), What are the most effective ways to document that learners and/or practitioners have achieved the required competencies?, these mixed findings do not fully support the use of the CRMTG as an effective way to document advanced

practice level nurses' genetic knowledge. These findings may be explained through further demographic exploration. The demographic questions included with the CRMTG asked for the highest level of nursing education; however, when genetic education data was requested, 66.7% of Group C, 49.1% of Group A, and no one in Group B had degrees in genetics (Table 16). These degrees would be at the graduate level, but not necessarily in nursing. The demographic form for the CRMTG will need to be rewritten to more clearly reflect genetic education in addition to nursing education for future research.

Internal consistency

An item-to-total correlation is used to evaluate the internal consistency of an instrument. This analysis evaluates the degree that each item in the instrument measures the same concept (Northern Arizona University, 1997). Of the 100 items on the CRMTG, 81 had item-to-total correlations above .50, the highest being .81. This suggests that most of the items contributed to the measurement of the concept of genetic knowledge. One item, number 76, was found to have a poor item-to-total correlation of .013 ($p = .805$) and should be deleted from future versions of the CRMTG. The other 99 items had significant item-to-total correlations ($p < .05$).

Internal consistency can also be evaluated using factor analysis and the Cronbach's alpha coefficient. As stated previously, factor analysis is not appropriate when analyzing dichotomous data. The alpha coefficient was calculated for the CRMTG to be .99.

Consistency over time

One of the measures of an instrument's reliability is consistency or stability over time. One method to evaluate this stability is the test-retest method. The test and retest scores should be similar, showing that, when given to the same group twice, the responses stay the same (Northern Arizona University, 1997).

The investigator noted that there was no way of identifying the first respondent to the CRMTG aside from a random code assigned by Surveymonkey ©. This was excellent for insuring anonymity of the participants, but would not allow for matching of test and retest scores. Upon identifying this problem, the investigator immediately redesigned the website to include an opening page with a request for email address for test-retest purposes only. This provided the participants the option to provide an identifier. This provided the investigator with a means to identify matching entries for the retest and to contact drawing winners.

Fifty-three participants logged on to the instrument website a second time to provide test-retest data. The only subjects included in the paired t-test, were those who scored above a zero on the instrument. Seventeen subjects' test and retest raw scores were both zero; therefore, these subjects were removed prior to the data analysis.

All 53 participants were included in the drawing for the amazon.com gift certificates. The email addresses were pulled from a hat to select the four winners. Serendipitously, the winners all had a raw score greater than zero. The investigator and amazon.com notified winners once their gift certificate was purchased and registered.

The participants were contacted one month after they had completed the initial test. They were contacted using the email address that they provided for test-retest purposes. Participants who did not indicate their email address were not contacted. The one-month time lapse may have affected the test-retest results because some respondents may have looked up the content area(s) that they felt weak in to try and improve their scores. However, the raw scores of the test-retest participants did not indicate any significant improvement in the results of the retest compared to the results of the initial test (Appendix V).

Only the 36 participants who received raw scores greater than 0 on both the test and the retest were used for statistical analysis. Based on a paired t-test of raw scores, there was no significant difference between test and retest scores, when retesting of the CRMTG was performed one month after initial testing. Participants in the retest ($N = 36$) represented Group A ($n = 14$), Group C ($n = 7$), Group D ($n = 11$), and Group E ($n = 4$). None of the Group B participants retested. A limitation of the study is that, even with the drawing as an incentive, there were so few retest participants.

Research Question 2

The second research question was: Do nurses who practice in a genetic setting have a higher pass rate on the CRMTG than nurses who do not practice in a genetic setting? According to the COPA model, evaluation of competency-based performance requires a criterion-referenced, summative evaluation of a participant's ability to meet a predetermined set of performance standards (Lenburg, 1999b). The CRMTG is a criterion-referenced instrument, which measures participants' knowledge of genetics

based on established standards of practice. Chi-square and ANOVA analyses of pass/fail outcomes and raw scores, respectively, demonstrated a significantly higher pass rate on the CRMTG for nurses who self-declared that they practiced in a genetic setting, Groups A, B, and C, than those nurses who self-indicated that they did not practice in a genetic setting, Groups D and E.

Research Question 3

The third research question was: Do advanced practice nurses have a higher pass rate on the CRMTG than nurses who do not have their Masters' degree in nursing? The CRMTG was designed to not only measure genetic knowledge, but measure genetic knowledge at an advanced practice level. Advanced practice nurses are defined as nurses with a Masters' degree in nursing. Advanced practice nurses in genetics provide genetic education, monitor genetic legislation, and conduct genetic research (ISONG, 2003b). Many genetic nurses have learned their specialty using an apprenticeship model. There are only five graduate level nursing programs in the United States with a specialty in genetics (ISONG, 2003a).

No significant association was found, using Chi Square analysis, between nursing education level and pass/fail outcomes among all nurses. Genetic nurses who had a Masters' or higher degree in nursing (Groups A and B) did not have significantly higher raw scores on the CRMTG than those having nursing education less than a Masters' degree in nursing (Group C). The limited availability of graduate genetic nursing programs, and the in-the-job training of most genetic nurses, regardless of education level may explain the findings of this study. Experienced genetic nurses may have more

knowledge of genetics than newer APNs. In addition, two-thirds of the genetic nurses had degrees in genetics. When the subsample of participants who responded to 85% or more items was analyzed, the genetic nurses with a Masters' or higher degree in nursing had significantly higher scores on the CRMTG than those having less than a Masters' degree in nursing. Clearly the different educational and experiential backgrounds of these genetic nurses affect the results of this study and make it difficult to discriminate among them using the CRMTG as the only means of evaluation.

The education item on the demographic form asked for highest level of nursing education. The investigator failed to anticipate the volume of participants who had non-nursing degrees and the implications of these educational preparations, particularly the non-nursing degrees in genetics.

An additional explanation of the findings would be that genetic nurses practice in a vast variety of settings, and each nurse has learned specific knowledge and skills necessary to function at the highest possible level in that setting. This variety of experience and preparation was one of the factors considered by the GNCC when a portfolio format was selected as the credentialing evaluation method (Monsen et al., 2005).

The Masters' prepared nurses in Group D did have higher pass rates than the non-Masters' nurses in Group E; however, all of these nurses were non-genetic and, therefore, not the primary focus of this study. The results of this group were important to this study because they provided divergent validity data. It is interesting to note that, even among

the non-genetic nurses, those with graduate degrees in nursing had higher pass rates than those without graduate degrees in nursing.

Research Question 4

The fourth research question was: Do nurses credentialed by the GNCC have a higher pass rate on the CRMTG than non-credentialed nurses? GNCC credentialed nurses were identified as Group B. Although the pass rate for Group B was slightly higher than Group A and slightly lower than Group C, there were no significant differences in pass rates among these groups (Table 18). The small number of nurses credentialed as an advanced practice nurse in genetics (APNG) might account for the pass rate for Group B (APNGs) being slightly higher than Group A (Masters' prepared genetic nurses), as the best of the best have sought the credential (R. Monsen, personal communications, June 11, 2002, June 25, 2002, July 1, 2002, July 4, 2002, April 17, 2006; Moyer, 2002). The APNGs, Group B, had lower pass rates than the non-Masters' in nursing genetic nurses, Group C. This finding may reflect the rationale identified above for the high pass rates of non-Masters' in nursing prepared genetic nurses.

Research Question 5

The fifth research question was: Are GNCC portfolio pass/fail outcomes congruent with CRMTG pass/fail outcomes? The five study participants in Group B were all successful portfolio applicants to GNCC. A 100 % pass rate would be anticipated for such an elite group of credentialed specialists since they could be considered the most qualified genetic nurses; however, the pass rate on the CRMTG for the credentialed nurses was only 60% for the entire sample (Table 18) and 75% for the subsample (Table

22). Even an evaluation of items specific to advanced practice nurses showed no consistent advantage for the APNGs who participated in this study (Table 21, Appendices J & W). This finding would lead one to voice concern regarding the validity and reliability of the GNCC portfolio credentialing process. Preliminary evidence suggests that the process may be too lenient; however, this study should be repeated using both pass and fail portfolio participants. A kappa analysis could then be computed comparing actual pass/fail rates for both the CRMTG and the GNCC portfolio.

There is no way of knowing if any unsuccessful portfolio applicants responded to the invitation to participate in the study. Participants were not specifically asked if they had failed the portfolio process on the CRMTG demographic form. Contacts through GNCC would have reached only the credentialed nurses, those who had passed the portfolio review. Unsuccessful GNCC portfolio applicants may have been contacted through other organizations such as ISONG and NSGC, but there was no mechanism on the CRMTG to identify them. Therefore, this study was not able to answer whether portfolio pass/fail outcomes are congruent with CRMTG outcomes.

Theoretical Frameworks

COPA Model

The COPA model (Lenburg, 1999a, 1999b) was particularly helpful, in this study, for direction during the literature review. After reading the model, the exploration of competence and the development of the research questions seemed to flow logically from the model. Lenburg's Eight Core Practice Competencies were congruent with the International Society of Nurses in Genetics, Inc. (ISONG) Statement on the Scope and

Standards of Genetics Clinical Nursing Practice (1998), the criterion from which the CRMTG was developed. When credentialing is an outcome, competence is imperative.

Waltz, Strickland, and Lenz Model

The eight stages of the Waltz and colleagues model (2005) for the development of criterion-referenced measures were integral to this study. These stages were the roadmap that the investigator used to systematically work through the instrument development process. The stages were clearly defined and easily operationalized. The 1991 and 2005 texts by these authors provided strategies for developing measurement instruments, both criterion- and norm-referenced. Especially helpful were the discussions of reliability and validity measurements for each of these types of measures that included suggested statistical analyses. One table included in the 2005 text delineates reliability procedures and lists the statistical analyses appropriate for norm- or criterion-referenced measures. The validity chapter also compares and contrasts validity procedures for both types of measures.

Methodological Issues

Internet Data Collection

The use of Internet data collection for this study supported the findings of the literature. A geographically diverse sample was obtained through this method of data collection (Gosling et al., 2004). Participants were fairly comfortable with "paper-less" data collection (Faculty of Social Sciences, 2002).

Multiple issues arose during the data collection process. There was some confusion associated with inclusion criteria, particularly for those subjects contacted to

provide divergent validity data. Also, as mentioned by Hayslett and Wildemuth (2004), many participants selected "other" as a default which made inclusion criteria more difficult to determine. The CRMTG survey response rate was very disappointing, particularly in light of the believed number of contacts made. Hayslett and Wildemuth addressed the difficulty of low response rates. Non-coverage may also be a factor in the low response rate as noted by Strickland and colleagues (2003). Another difficulty appeared when the retest was attempted and the participants were unable to access the site a second time due to cookies placed on their computers. Fortunately, many participants with questions or concerns were very open with their issues and the investigator was able to provide clarification. Technological errors and interface difficulties were mentioned in the literature (Huffman, 2006), but not the specific issue of cookies placement by the computer program which is the difficulty that was faced in this study.

This method of data collection was very cost effective (Huffman, 2006). In this study, no money was spent on copying instruments, mailings, and follow-up postcards. The Surveymonkey © site does have a cost associated with it for advanced survey analysis. The site fee was $29 per month. The investigator has maintained this site account for the past 24 months. Some of the organizations did charge from $285 to $2,500 for access to their membership; however, the cost of mailing this large instrument to the volume of potential participants would have been prohibitive. Listserv access, easier coding, and quicker response times were very positive results of Internet data collection (Evans & Mathur, 2005; Strickland et al., 2003; Waltz et al, 2005). However, low response rates and incomplete data made data interpretation difficult.

Cookies

One of the issues that emerged during the Internet administration of the CRMTG was the cookies that the Surveymonkey © site placed on participants' computers. In order for participants to take the retest, they needed to delete the cookies or log onto the website using a different computer. The cookies were placed to allow the participants to resume the CRMTG at the point where they left off and to prevent them from going back and changing answers.

Initially the investigator was not aware of the cookie issue. Early participants notified the investigator when they attempted to re-access the site when requested to retest. Once the investigator realized the situation, the instructions to remove the cookies or use a different computer were included in the retest request letter (Appendix T). This may have contributed to the low retest response rate (10%). A pilot of the Surveymonkey © instrument for the retest component would have unveiled the cookie issue.

Genetic Knowledge

Another issue was the confusion that potential participants had regarding inclusion criteria for the study. Several early participants emailed the investigator stating that they were not eligible because they were not advanced practice genetic nurses. The investigator responded to them that they were very important to the study for divergent validity and encouraged them to participate. In response to that information, the investigator added the following paragraph to the request letter (Appendix P):

"You do not need to have any knowledge of genetics to participate in this study. In fact, I need a large number of nurses who do not practice in a genetics setting in order to assess instrument validity and reliability."

Response Rate

This study had a very low response rate (2%). The response rate may be related to the genetic knowledge issue stated above. Genetic content can be very intimidating for non-genetic nurses (Collins, 2002; Saunders, 1998). The response rate may be related to the use of the Internet for contact and data collection. Internet surveys have been documented as having lower response rates as compared to mailed surveys (Kittleson, 1995). Nurses have specifically been identified as a difficult group to capture through the Internet (Ehrenberger & Murray, 1998; Im et al., 2006).

Non-coverage may be a factor in the low response rate as well. Non-coverage refers to missing potential participants. Although the numbers of members listed in Table 6 was provided by the organizations, not all of the nurses in an organizations' membership may have received the request to participate. Members may have missed the request for study participants for a number of potential reasons as follows: 1) organizations may not have the most current contact information for their members; 2) some organization members may elect not to subscribe to organization-wide listservs; 3) many email providers use email filters to prevent "spam" emails, viruses, and/or spyware (Graham, 2002; Sorkin, 2005; Stone & Weil, 2003); 4) individuals may delete emails from people that they don't know; 5) a variety of technical or human errors may have occurred such as errors in keying in email addresses either at the organization level or by

the investigator; and/or 6) the duplication of contact numbers due to nurses having multiple memberships in nursing organizations, such as being a member of the ISONG, GNCC, and NSGC; thus, three contacts were actually only one person (Faculty of Social Sciences; Im et al., 2006; Strickland et al., 2003).

It is impossible to accurately calculate the actual response rate due to the issues raised above. There is no way to count the actual number of nurses contacted for this study. Review of the numbers in Table 6 would suggest that, with a zero percent response rate, members of the Emergency Nurses Association (ENA) and the American Medical Surgical Nurses (AMSN), did not receive notification of the study. The one participant from each organization may have been contacted through another means, i.e. personal contact, and marked the organization(s) as a member. The recalculated response rate for this study, excluding these two organizations would be fourteen percent. The highest response rates were from members of genetic organizations and personal contacts. These people had a vested interest in either the investigator, the study, or both. Advanced practice nurses would, reasonably, have a greater interest in supporting and participating in nursing research. Of particular interest to this group may be the advancement of nursing specialty credentialing that would explain the relatively high level of participation of members from the National Association of Clinical Nurse Specialists (NACNS).

Concern could be raised regarding adequate representation of the population with such a low response rate (2%), small sample size ($\underline{N}$ = 356), and number of participants who answered only a small percentage of items on the CRMTG. The response rate could be falsely low due to reduced numbers of potential participants who actually received the

request to participate. The low response rate and small sample size may not have had a large affect on this study's results based on findings of Tambor and colleagues (1993). Tambor and colleagues conducted a study to address small sample size and low response rates. They conducted a pilot study (response rate 19.6%, N = 69) and then underwent a more rigorous and labor-intensive recruitment of participants for their final study (response rate 64.8%, N = 1140). These researchers compared the results of their pilot to the results of their final study. There were no significant differences in study results between the small, low response rate, pilot study and the 16 times larger, three times greater response rate, final study. It is encouraging that, based on the Tambor and colleagues study cited, the results of this study may be reflective of the population contacted and not just the population studied.

Strengths of the Study Methodology

As mentioned above, the Waltz and colleagues model (2005) for criterion-referenced instrument development was incredibly helpful for organizing and conducting this study. The stages were well identified, logical, and clearly defined. Each stage could be implemented efficiently.

The expert reviewers that the investigator used for this study provided invaluable information for each test item. The experts spent an extended amount of time carefully reviewing, editing, critiquing, and rating each item in the total CRMTG item pool, which allowed the investigator to compile the best 100 items into the final CRMTG.

The Internet survey procedure was very cost effective for the investigator. A minimal monthly payment to access the website has been on-going from initial data

collection through completion of this study. Some of the organizations contacted charged the investigator for access to and/or contact information for their membership.

Surveymonkey © was very easy to set up. Having gone through this experience, the investigator would use this site again, but would set the instrument up differently. Changes would be made to the demographic form.

Weaknesses of the Study Methodology

The data collected on the demographic form was not exactly what was needed to accurately compare the results of the groups. Education questions were restricted to highest level of nursing education. This was an appropriate question; however, a follow-up question regarding other, non-nursing degrees would have been helpful, particularly when evaluating the pass/fail results of Group C as the demographic data indicates that two-thirds of the respondents who did not have a graduate degree in nursing, did have some level of genetic education. These data would also have helped in the discussion of research questions three and four.

The investigators' criteria for the genetic sub-specialties included on the CRMTG may need to be added to the demographic form. These definitions may have decreased the number of "other" write-ins under genetic sub-specialty on the demographic form. The investigator considered many of the areas that were written in to have fit into existing categories.

A larger number of GNCC credentialed nurses would have added to the value of this study. A focused study limited to GNCC applicants to compare pass/fail outcomes between portfolios and the CRMTG is recommended once the reliability and validity of

the CRMTG has been established. These data would have helped in the discussion of research question five.

A test bank was used for the initial pool of test items for the CRMTG. Preliminary psychometric data were not available for these items. This test bank was developed as a faculty aid to accompany a genetic text. Test items that were developed by the investigator tended to have greater difficulty indices and higher taxonomy levels than the test bank items. Having a greater percentage, if not all, of the CRMTG items created by the investigator would have been preferable, however, labor intensive. Having a larger pool of expert reviewed items, and more pilot study data would also have been beneficial.

The pilot study of the paper and pencil version of the CRMTG was very helpful for content validity data. A larger pilot study sample size would have given more psychometric data on each item. The initial pilot study was important to conduct prior to converting the CRMTG to an online format. A second pilot study of the instrument of the full study process once the CRMTG was converted to the online format would have been beneficial in identifying the test-retest cookies issue.

The limitations of Internet data collection, as identified in the literature, were supported in this study. Access to sample, low response rates, and technical issues were the primary weaknesses in this study.

Conclusions

In relation to the previously identified research questions for this study the results indicate that

1) The CRMTG is a valid and reliable indicator of knowledge of advanced practice nurses in genetics.
2) Nurses who practice in a genetic setting had a higher pass rate on the CRMTG than nurses who do not practice in a genetic setting.
3) Advanced practice nurses did not have a higher pass rate on the CRMTG than nurses who do not have their Masters' degree in nursing.
4) Nurses credentialed by the GNCC did not have a higher pass rate on the CRMTG than non-credentialed nurses.
5) It was not feasible, in this study, to compare the pass/fail outcomes on the CRMTG with the GNCC portfolios outcomes, since there were no participants who had failed the portfolio process.

Limitations

This study was limited to registered professional nurses in the United States who had computer access to the CRMTG. The primary limitation of this study is the low response rate leading to a small sample size. The sample of 356 did exceed the minimum requirement of 300 participants for item analysis. Two hundred ninety-three participants failed the CRMTG, 117 of whom had zero correct answers. Once all of the participants who answered less than 85% of the items on the CRMTG were removed from the data bank, the final sample size was 195, which is below the minimum threshold for this instrument. This reduces the power of the results of this study; however, the item analysis findings of this study could be used to improve the instrument for future instrument development and replications of this study.

Knowledge of how missing data are handled is important in interpreting results of a study about instrument development. For this study a skipped answer was marked incorrect to be consistent with standard test grading practice. This provided inconsistent data, because, when overall scores were reviewed it was impossible to identify if a given participant logged onto the website and never answered a question, quit completing the CRMTG part way through, or answered all or most of the questions, but answered incorrectly. However, once the raw data was reviewed and the subsample was identified, the results of the subsample data analysis were consistent with the overall results for the first three research questions. The results of the subsample analysis for the fourth research question showed significantly higher CRMTG pass rates for Masters' in nursing prepared genetic nurses as compared to non-Masters' in nursing genetic nurses.

The use of the Internet for both survey distribution and data collection, while being cost effective, limited the number of participants in the study. The electronic contact with potential participants made it difficult to know how many individuals were reached and, therefore, made it difficult to calculate accurate response rates. The electronic data collection made it difficult to adjust for missing data. The use of the Internet also affected the testing environment. The investigator had no control over the testing conditions. Participants may have logged on at work or home, had multiple interruptions, or even used genetic resources to aid them in competing the CRMTG.

The small number of genetic nurses available in the United States and the even smaller number of credentialed genetic nurses adds to the limitations of this study. With

such a small pool of nurses, even with reasonable response rates, the overall sample would be very small.

There was no significant difference between Masters' and non-Masters' prepared genetic nurses. This is not a surprising result, in that many genetic nurses have been trained in an apprentice-type model. There have been very few Masters' in nursing programs with the focus on the genetic specialty. Most genetic nurses have taken a variety of continuing education courses, received degrees in genetics, engaged in various forms of independent study, and/or participated in supervised clinical education.

The low Bloom's taxonomy levels of multiple-choice test items limit the ability of the CRMTG to evaluate advanced practice nursing knowledge. The addition of open-ended case study items would increase the cognitive level of items on the CRMTG; however, this style of item is more difficult and more subjective to score when compared to multiple-choice items. Traditionally, multiple-choice items are used for credentialing examinations. The disparity between the use of multiple-choice items and the evaluation of higher cognitive levels was the rationale behind the GNCC's design of the portfolio process for credentialing.

Implications

Nursing Practice

Portfolios have been used for both personal and program evaluation in a variety of professional settings. Nursing specialty organizations are exploring the use of portfolios for credentialing nurses. There is a limited amount of data available to psychometrically support the use of portfolios for the awarding of professional credentials (Dennison,

2005; Friedman & Marr, 1995; Hayes, Chandler, Merriam, & King, 2002; Johnson, 2002; Lenburg, 2000; Moyer, 2002; Seguin, 2005; Weddle, Himburg, Collins, & Lewis, 2002). The CRMTG can be used to validate the Genetic Nurses Credentialing Commission (GNCC) portfolio process for the awarding of credentials, thus adding to the professional literature in this exploratory area. The CRMTG may also be used to supplement or replace the portfolio process as determined by the GNCC board of directors.

This procedure for the development of an instrument to measure specialty nursing knowledge based on identified scope and standards can provide a framework for other specialty nursing groups. These groups can use this procedure to develop and maintain their own credentialing examinations. These groups may choose to model the genetic nurses portfolio process and use the procedure illustrated in this study to provide psychometric data for their portfolio process or they may choose to develop their own, unique process for the awarding of credentials. The procedure used in this study would allow them to evaluate the validity and reliability of whatever process they select by comparing it to the current standard, the credentialing examination (Cary & Smolenski, 2005).

Nursing Science

This study provides an opportunity to implement one aspect of the COPA model (Lenburg, 1999a, 1999b) in a non-educational setting. The flexibility of this model strengthens its usefulness within educational settings as well as opening an arena for use in professional clinical settings. This study has implemented the COPA Model in an advanced practice venue, which expands the model's scope beyond undergraduate

educational applications and minimum competency levels into post-graduate practice and national credentialing.

The implementation of the Waltz, Strickland and Lenz Stages for the Development of Criterion-Referenced Measures (1991, 2005) further supports the function and application of this model. Additionally, operationalizing the ISONG Statement on the Scope and Standards of Genetics Clinical Nursing Practice (1998) into an instrument provides a model for additional measurement of professional competence using other professional scope and standards documents.

Nursing Education

The specialty of genetic nursing requested a credential for professional recognition and monetary reimbursement of practitioners in many genetic settings. The Genetic Nurses Credentialing Commission (GNCC) created this credential. The GNCC uses a portfolio review method for the credential process (Moyer, 2002). The CRMTG, an instrument for the evaluation of genetic nursing knowledge, can be made available to the GNCC for inclusion in its credentialing process as a valid and reliable measurement of genetic knowledge. The instrument does not strongly differentiate between advanced practice and non-advanced practice nurses; however, the GNCC portfolio process includes transcript review which will provide that data. Review of Neural Net scores, a quantitative component of the GNCC portfolio process, would provide actual portfolio scores for candidates. The CRMTG can be used to provide psychometric data for the GNCC. Pass/fail portfolio results can be compared, using Kappa analysis, to CRMTG

pass/fail results. This study provides an instrument to measure genetic nursing knowledge and to evaluate and validate the GNCC portfolio review process.

The CRMTG may also be useful for nurse educators in graduate genetics nursing programs. Having a valid and reliable measure of advanced practice genetic nursing knowledge would provide a test pool for the development of graduate level examinations. The CRMTG may also be useful as an exit examination for these programs. Because the graduates of these programs are potential GNCC credentialing applicants, the applications of the CRMTG must be carefully considered and monitored.

Nursing Research

The development and testing of the CRMTG provides a model for evaluating an innovative nursing education and practice strategy, the portfolio; an example of credentialing research; an application of the Waltz and colleagues (2005) stages of instrument development; and a demonstration of the use of the Internet for the collection of data.

This study provides a description and methodology for instrument development. The instrument was developed using published clinical scope and standards. The instrument was developed in order to compare the psychometric results of a portfolio process for credentialing to the standard credentialing process, an examination (Nichols, 1991; Barnum, 1997). This study provides a model for the development of a criterion-referenced instrument that can be used to provide psychometric data and/or supplement a portfolio method of competency evaluation.

Credentialing research is a growing field of research both inside and outside of the nursing profession. As more organizations take on a variety of credentialing models, psychometric data are needed to assure the public that the credentialed professionals have met minimum standard criteria. This study demonstrates the beginning stages in the psychometric analysis of an innovative credentialing process.

The stages of the Waltz and colleagues (2005) model were clearly defined and easily operationalized. This study provides an example of the application of these well-articulated stages.

The Internet was used for seeking approval from organizations, contacting study participants, instrument administration, and raw data collection. Although the response rate was less than hoped, the study provides important information for future researchers interested in Internet data collection. Further refinement of the process would include a second pilot once the instrument was loaded onto the website. A practice run through all of the stages of data collection would help to identify potential weaknesses in the process.

Recommendations

Prior to making the CRMTG available to the GNCC, further refinement of the instrument is required.

1) Further testing using the CRMTG would continue to provide item analysis data that would allow for continued refinement of the instrument.

2) Item 76 needs to be replaced and, based on item analysis, some of the other items could be re-written to enhance the quality of the distracters.

3) Development of more case study items or other styles of items to increase the difficulty and cognitive levels of the CRMTG test items is recommended.

4) A larger pool of test items, with content, construct, and pilot data for each would also strengthen the final instrument.

5) Further exploration of the validity of the 65% cut score may need to be performed. This is a low value compared to the standard cut score of 80% for criterion-referenced measures (Wilde & Sockey, 1995). A change in cut score would alter study results.

In future studies it would be beneficial to expand in two ways.

1) Future research in genetics could include other questions from the COPA model (Lenburg, 1999a). The Essential nursing competencies and curricula guidelines for genetics and genomics (Jenkins et al., 2005) have been identified; however, the most effective way to learn these competencies has yet to be established, particularly a method of reaching and teaching practicing nurses and other health care providers.

2) Further research needs to be done, once the CRMTG is revised, to determine the extent of validity and reliability of the GNCC portfolio process for the awarding of genetic nursing credentials. Current credentialing processes using multiple-choice examinations are confirmed to be very labor intensive, expensive, and difficult to maintain relevancy and accuracy. It can be easily understood how many

organizations from the very large, such as the ANCC, to small, such as the forensic nurses, are considering the role of portfolios in the credentialing process.

A future recommendation for use of the CRMTG may be to add it to the GNCC credentialing process as a supplement to the portfolio. Some critics have raised concerns that the portfolio does not evaluate basic knowledge of genetics and the CRMTG has been established as being a valid and reliable indicator of genetic knowledge.

Summary

The CRMTG was found to be valid and reliable based on the findings of this study. The CRMTG pass rates indicated a distinct difference between genetic and non-genetic nurses. The CRMTG pass rates for the whole sample were not found to be useful for differentiating between Masters' in nursing and non-Masters' in nursing prepared nurses. However, the pass rates for the subsample of participants who answered at least 85% of the items on the CRMTG did indicate a distinct difference between Masters' in nursing and non-Masters' in nursing prepared nurses.

This study provides a model for the application and operationalization of nationally approved nursing practice scope and standards into a measurement instrument, the application of both COPA and Waltz and colleagues frameworks, and the use of the Internet for data collection. This study also provides psychometric data for an instrument that may be useful to the GNCC as a validation and/or supplement to their portfolio credentialing process.

References

Airasian, P. W. (1994). The impact of the taxonomy on testing and evaluation. In L. W. Anderson & L. A. Sosniak (Eds.), Bloom's taxonomy; A forty-year retrospective (pp. 82-102). Chicago: The University of Chicago Press.

Alabama State Nurses' Association (ASNA). (2000). ANCC to make certification available to all registered nurses through Open Door 2000 program. Wyoming Nurse, 12(3), 17.

Alsop, A. (2001). Competence unfurled: Developing portfolio practice. Occupational Therapy International, 8, 126-131.

American Nurses Association (ANA). (1996). ANA Policy Paper: ANA Response to Pew Commission Report. Washington, DC: author.

American Nurses Credentialing Center (ANCC). (2003a). Frequently asked questions. Retrieved January 10, 2003, from http://www.nursingworld.org/ancc/certify/faqs.htm

American Nurses Credentialing Center (ANCC). (2003b). Exam information and updates. Credentialing News, 6(4), 7.

American Nurses Credentialing Center (ANCC). (2004). Acute Care Nurse Practitioner. Retrieved December 3, 2004, from http://www.nursingworld.org/ancc/certification/cert/exams/tcos/acnptco.html

American Nurses Credentialing Center (ANCC). (2006). Frequently asked questions. Retrieved April 7, 2006, from http://www.nursingworld.org/ancc/cert/certfaqs.html

Anderson, C. A. (1997, September 12). Response to Barnum, B. S. (1997, August 13). Licensure, certification, and accreditation. Online Journal of Issues in Nursing, 2(4). Retrieved November 7, 2002, from http://www.nursingworld.org/ojin/letters/tlc5.htm

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives. New York: Longman.

Bargagliotti, T., Luttrell, M., & Lenburg, C. B. (1999, September 30). Reducing threats to the implementation of a competency-based performance assessment system. Online Journal of Issues in Nursing, 4(2). Retrieved November 7, 2002, from http: //www.nursingworld.org/ojin/topic10/tpc10_5.htm

Barnum, B. S. (1997, August 13). Licensure, certification, and accreditation. Online Journal of Issues in Nursing, 2(3). Retrieved November 7, 2002, from http://www. nursingworld.org/ojin/tpc4/tpc4_2.htm

Barriball, K. L., & White, A. E. (1999). Non-response in survey research: A methodological discussion and development of a explanatory model. Journal of Advanced Nursing, 30, 677-686.

Bastable, S. B. (1997). Nurse as educator: Principles of teaching and learning. Sudbury, MA: Jones and Bartlett.

Beck, C. T., & Gable, R. K. (2001). Ensuring content validity: An illustration of the process. Journal of Nursing Measurement, 9, 201-215.

Bell, M. D., Greig, T. C., Kaplan, E., & Bryson, G. (1997). Wisconsin Card Sorting Test dimensions in schizophrenia: Factoral, predictive, and divergent validity. Journal of Clinical and Experimental Neuropsychology, 19, 933-41.

Ben-David, M. F. (2000). The role of assessment in expanding professional horizons. Medical Teacher, 22, 472-477.

BenDebba, M., Heller, J., Ducker, T. B., & Eisinger, J. M. (2002). Cervical spine outcomes questionnaire: Its development and psychometric properties. Spine, 27, 2116-2123.

Biedrzycki, B. (2002, July). Gene therapy: Rhetoric, risks and research. Educational session (SC096) presented at the American Nurses Association 2002 Biennial Convention, Philadelphia.

Bloom, B. S., Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R. (1956). Taxonomy of educational objectives: The classification of educational goals, handbook 1. New York: David McKay.

Blumenreich, G. A. (2002). Fairness in credentialing and the certified registered nurse anesthetist. American Association of Nurse Anesthetists Journal, 70, 347-350.

Bronstein, L. R. (2002). Index of interdisciplinary collaboration. Social Work Research, 26, 113-123.

Brown, M-A. (1996). Primary care nurse practitioners: Don't blend the colors in the rainbow of advanced practice nursing. Online Journal of Issues in Nursing, 1. Retrieved November 2, 2005, from http://www.nursingworld.org/ojin/tpc1/tpcl_6.htm

Burns, N., & Grove, S. K. (2005). The practice of nursing research: Conduct, critique and utilization (4th ed.). Philadelphia: Saunders.

The Carnegie Foundation for the Advancement of Teaching. (2000). Carnegie classification. Retrieved July 23, 2004 from http://www.carnegiefoundation.org/Classification/CIHE2000/defNotes/Definitions.htm

Carol, R. (2003, winter). The future is now. Minority Nurse, 22-27.

Carson, A. R., & Hieber, K. V. (2001). Adult pediatric patients. American Journal of Nursing, 101(3), 46-54.

Cary, A. H. (2001). Certified registered nurses: Results of the study of the certified workforce. American Journal of Nursing, 101(1), 44-52.

Cary, A. H., & Smolenski, M. C. (2005). Overview of competency and the methods for evaluating continued competence. In R. B. Monsen (Ed.), Genetics nursing portfolios: A new model for credentialing (pp. 1-10). Silver Spring, MD: American Nurses Association.

Center for Nursing Education and Testing (C-NET). (1998). Improving test construction skills. (Available from C-NET, 601 Pavonia Avenue, Suite 601, Jersey City, NJ 07306).

Chadima, D. (2006, December). Planning for successful general education outcomes assessment. Presented at the annual conference of the Middle States Commission on Higher Education, Philadelphia, PA.

Chambers, R. (2002, June 17). Adopt APD to make appraisal simpler. General Practitioner, 44.

Chang, L. (1999). Judgemental item analysis of the Nedelsky and Angoff standard-setting methods. Applied Measurement in Education, 12, 151-165.

Child, D. (2004, July 7). Flesch-Kincaid reading level. Retrieved December 22, 2004, from http://www.ilovejackdaniels.com/php/flesch-kincaid-function

Clemans, W. V. (1971). Test administration. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 188-201). Washington, DC: American Council on Education.

Clinton, J., Denyes, M. J., Goodwin, J. O., & Koto, E. M. (1977). Developing criterion measures of nursing care: Case study of a process. Journal of Nursing Administration, 7(7), 41-45.

Collins, F. S. (2002, July). Our genetic future. Plenary session I (SC100) presented at the American Nurses Association 2002 Biennial Convention, Philadelphia.

Comrey, A. L., & Lee, H. B. (1992). A first course in factor analysis. Hillsdale, NJ: Lawrence Erlbaum.

Conley, Y. P., & Gorin, M. B. (2003). The genetics of age-related macular degeneration. MEDSURG Nursing, 12, 238-241, 259.

Cook, C. B. (1999, September 30). Initial and continuing competence in education and practice: Overview and summary. Online Journal of Issues in Nursing, 4(2). Retrieved November 7, 2002, from http://www.nursingworld.org/ojin/topic10/tpc10ntr.htm

Cooper, C. D. (2000). The use of electronic mail in the research process. Nurse Researcher, 7(4), 24-30.

Critical Care Nurse Specialtist (CCNS) Exam. (2004). Retrieved December 3, 2004, from http://www.certcorp.org/certcorp/certcorp.nsf/edcfc72ba47aaa708825666b0064bdcf/a7120ea6f671658b882566c000796d3c?OpenDocument

Cronbach, L. J. (1971). Test validation. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 443-507). Washington, DC: American Council on Education.

Daley, E. M., McDermott, R. J., Brown, K. R. M., & Kittleson, M. J. (2003). Conducting Web-based survey research: A lesson in Internet design. American Journal of Health Behavior, 27, 116-124.

Data imputation for missing values. (n.d.). Quantitative Methods in Public Administration. Retrieved February 26, 2007, from http://www2.chass.ncsu.edu/garson/pa765/missing.htm

De Ayala, R. J., & Kelley, H. P. (1987). MEC item analysis. Austin, TX: University of Texas. Retrieved May 15, 2006 from http://www.utexas.edu/academic/mec/scan/scanitem.html

Demetrulias, D. A. M., & McCubbin, L. E. (1982, Autumn). Constructing test questions for higher level thinking. Nurse Educator, 7, 13-17.

Dennison, R. D. (2005). Portfolios: Current uses. In R. B. Monsen (Ed.), Genetics nursing portfolios: A new model for credentialing (pp.25-39). Silver Spring, MD: American Nurses Association.

Denyes, M. J. (1980). Development of an instrument to measure self-care agency in adolescents. Dissertation Abstracts International, 4(5), 1716B. University Microfilms No. 8025672.

Diekelmann, N. L., & Ironside, P. M. (2005). Prologue: Using portfolios in genetics nursing. In R. B. Monsen (Ed.), Genetics nursing portfolios: A new model for credentialing (pp. vii-xi). Silver Spring, MD: American Nurses Association.

Dillman, D. A., Tortora, R. D., & Bowker, D. (1999). Principles for constructing web surveys (1st ed.). Washington, DC: Pullman.

Dolan, J. A. (1968). History of nursing (12th ed.). Philadelphia: W.B. Saunders.

Donabedian, A. (1981). Criteria, norms and standards of quality: What do they mean? American Journal of Public Health, 71, 409-412.

Donen, N. (1998). No to mandatory continuing medical education, yes to mandatory practice auditing and professional educational development. Canadian Medical Association Journal, 158, 1044-1046.

Duffy, M. (2002). Methodological issues in web-based research. Journal of Nursing Scholarship, 34(1), 83-88.

Dyne, P. L., Strauss, R. W., & Rinnert, S. (2002). Systems-based practice: the sixth core competency. Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine, 9, 1270-1277.

Ebel, R. L. (1979). Essentials of educational measurement. Englewood Cliffs: Prentice-Hall.

Ebel, R. L., & Frisbie, D. A. (1991). Essentials of educational measurement (5th ed.). Englewood Cliffs: Prentice-Hall.

Ehrenberger, H., & Murray, P. J. (1998). Issues in the use of communications technologies in nursing research. Oncology Nursing Forum, 25(10, Suppl.), 11-15.

Erickson, J. I., Duffy, M. E., Gibbons, M. P., Fitzmaurice, J., Ditomassi, M., & Jones, D. (2004). Development and psychometric evaluation of the professional practice environment (PPE) scale. Journal of Nursing Scholarship, 36, 279-285.

Evans, J. R., & Mathur, A. (2005). The value of online surveys. Internet Research: Electronic Networking Applications and Policy, 15, 195-219.

Faculty of Social Sciences, University of Ljubljana, Slovenia. (2002, June 6). Standards for internet survey data collection. Retrieved July 14, 2003 from www.websm.org /fp6/Eol_Internet_surveys_integratedproject_18.pdf

Fairman, J., & Lynaugh, J. (1998). Critical care nursing: A history. Philadelphia: University of Pennsylvania.

Fallon, L. (2004). Genetic Counseling. Retrieved July 5, 2004, from http://www.findarticles.com

Farmer, T. (1998). Using the internet for primary research data collection. Journal of Computer-Mediated Communication, 5(3), 124-132.

Feetham, S., Thomson, E. J., & Hinshaw, A. S. (2005). Nursing leadership in genomics for health and society. Journal of Nursing Scholarship, 37, 102-110.

Ferguson, G. A. (1976). Statistical analysis in psychology & education (4th ed.). New York: McGraw-Hill.

Fickeissen, J. L. (1990). 56 ways to get certified. American Journal of Nursing, 90(3), 50-57.

Flynn, M. K., & Reese, J. L. (1988). Development and evaluation of classroom tests: A practical application. Journal of Nursing Education, 27(2), 61-65.

Frankenburg, W. K., Dodds, J. B., & Fandal, A. W. (1970). Denver developmental screening test manual, revised edition. Denver: University of Colorado Medical Center.

Friedman, S., & Marr, J. (1995). A supervisory model of professional competence: A joint service/education initiative. Nurse Education Today,15, 239-244.

Frisbie, D. A. (1983). Testing achievement beyond the knowledge level. Journal of Nursing Education, 22, 228-231.

Froman, R. D. (2001). Elements to consider in planning the use of factor analysis. Southern Online Journal of Nursing Research, 5(2). Retrieved July 23, 2004 from http://www.snrs.org/members/SOJNR_articles/iss05vol02.pdf

Gaberson, K. B. (1996). Test design: Putting all the pieces together. Nurse Educator, 21(4), 28-33.

Gaskins, S., Dunn, L., Forte, L., Wood, F., & Riley, P. (1996). Student perceptions of changing answers on multiple choice examinations. Journal of Nursing Education, 35(2), 88-90.

Genetic Nursing Credentialing Commission. (2001, December). Portfolio evaluation and scoring training manual. Hot Springs, AR: Author.

The Genomics Policy Unit & The Medical Genetics Service for Wales. (2003). Fit for practice in the genetics era: A competence based education framework for nurses, midwives and health visitors. Pontypridd, UK: University of Glamorgan.

Gessaroli, M. E., & Poliquin, M. (1995). Competency-based certification project phase III: Content validity, standard setting and criterion-related validity. The Canadian Journal of Medical Radiation Technology, 26 (1), 17-21.

Gillis, A., & Jackson, W. (2002). Research for nurses: Methods and interpretation. Philadelphia: F. A. Davis.

Girot, E. A. (2000). Assessment of graduates and diplomates in practice in the UK – are we measuring the same level of competence? Journal of Clinical Nursing, 9, 330-337.

Glaser, R. (1963). Instructional technology and the measurement of learning outcomes: Some questions. American Psychologist, 18, 519-521.

Glaser, R & Nitko, A. J. (1971). Measurement in learning and instruction. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 625-670). Washington, DC: American Council on Education.

Goeden, S. (1999, Jan/Feb). Should registered nurses have to periodically take an exam to be relicensed as proof of their continued competence? No. The American Nurse. Retrieved November 7, 2002, from http://www.nursingworld.org/tan/99janfeb/kaleid.htm

Goodwin, L. D. (1999). Relations between observed item difficulty levels and Angoff minimum passing levels for a group of borderline examinees. Applied Measurement in Education, 12(1), 13-28.

Gosling, S. D., Vazire, S., Srivastava, S., & John, O. P. (2004). Should we trust web-based studies?: A comparative analysis of six preconceptions about Internet questionnaires. American Psychologist, 59, 93-104.

Graham, P. (2002). Will filters kill spam?. Retrieved July 31, 2006, from http://www.paulgraham.com/wfks.html

Graziano, A. M., & Raulin, M. L. (2000). Research methods: A process of inquiry (4th ed.). Boston: Allyn & Bacon.

Greco, K. E. (2003). Nursing in the genomic era: Nurturing our genetic nature. MEDSURG Nursing, 12, 307-312.

Greco, K. E., & Mahon, S. M. (2003). Genetics nursing practice enters a new era with credentialing. The Internet Journal of Advanced Nursing Practice, 5 (2). Retrieved March 10, 2003, from http://www.ispub.com/ostia/index.pnp.xmlFilePath=journals/ijanp/vol5n2/genetics.xml

Green, A., & Ogden, B. S. (1999, March/April). Three state nursing boards examine continued competency. The American Nurse. Retrieved November 7, 2002 from http://www.nursingworld.org/tan/99marapr/competen.htm

Gronlund, N. E. (1973). Preparing criterion-referenced tests for classroom instruction. New York: Macmillan.

Gronlund, N. E. (1985). Measurement and evaluation in teaching (5th ed.). New York: Macmillan.

Guilford, J. P. (1954). Psychometric methods (2nd ed.). New York: McGraw-Hill.

Gullickson, B. (1988). Examining simulated clinical performance. In Strickland, O.L. & Waltz, C.F. (eds.). Measurement of nursing outcomes, vol. 2, measuring nursing performance: Practice, education, and research. New York: Springer.

Hair, J. E., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). Multivariate data analysis (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Halpern, R. S., Thompson, C. B., & Schaffer, J. (2001). Comparison of mailed vs. Internet applications of the Delphi technique in clinical informatics research. Retrieved July 14, 2003 from www.amia.org/pubs/symposia/D2001120.pdf

Hambleton, R. & Eignor, D. (1978). Guidelines for evaluating criterion-referenced tests and test manuals. Journal of Educational Measurement, 15, 321-327.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). MMSS: Fundamentals of item response theory. Newbury Park, CA: Sage Publications.

Hart, K. E., & Sciutto, M. J. (1996). Criterion-referenced measurement of instructional impact on cognitive outcomes[F*]. Journal of Instructional Psychology, 23, 26-34.

Hawk, D. (1999, Jan/Feb). Should registered nurses have to periodically take an exam to be relicensed as proof of their continued competence? Yes. The American Nurse. Retrieved November 7, 2002, from http://www.nursingworld.org/tan/99janfeb/kaleid.htm

Hayes, E., Chandler, G., Merriam, D., & King, M. C. (2002). The master's portfolio: Validating a career in advanced practice nursing. Journal of the American Academy of Nurse Practitioners, 14, 119–125.

Hayslett, M., & Wildemuth, B. (2004). Pixel or pencils? The relative effectiveness of web-based versus paper surveys. Library and Information Science Research, 26(3), 73-93.

Hayter, J. (1983). Educational taxonomies revisited. Journal of Nursing Education, 22, 339-342.

Health Resources and Services Administration (HRSA). (n.d.). Preliminary findings: 2004 National sample survey of registered nurses. Retrieved August 18, 2006, from http://bhpr.hrsa.gov/healthworkforce/reports/rnpopulation/preliminaryfindings.htm

Hegyvary, S. T. (2005). Genomics, tsunamis, and other frontiers of knowledge. Journal of Nursing Scholarship, 37, 95-96.

Henig, R. M. (2004, October 10). The genome in black and white (and gray). The New York Times Magazine, 47-51.

Henrysson, S. (1971). Gathering, analyzing, and using data on test items. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 130-159). Washington, DC: American Council on Education.

Hood, L., & Leddy, S. K. (2005). Leddy & Pepper's conceptual bases of professional nursing (6th ed.). Philadelphia: Lippincott, Williams & Wilkins.

Horne, C., & Medley, F. (2001). Community-based strategies: an interdisciplinary community-based clinical experience for beginning students. Nurse Educator, 26, 120-121.

Huffman, I. (2006). Online questionnaire software advantages/disadvantages. Unpublished master's thesis, University of North Carolina, Chapel Hill.

Hull, A. L., Hodder, S., Berger, B., Ginsberg, D., Lindheim, N., Quan, J. et al. (1995). Validity of three clinical performance assessments of internal medicine clerks. Academic Medicine, 70, 517-22.

Hurtz, G. M., & Hertz, N. R. (1999). How many raters should be used for establishing cutoff scores with the Angoff method? A generalizability theory study. Educational and Psychological Measurement, 59, 885-897.

Im, E-O., & Chee, W. (2003). Issues in Internet research. Nursing Outlook, 51, 6-12.

Im, E-O., Chee, W., Lim, H. J., Bender, M., Tsai, H-M., Yang, S-O. et al. (2006). Recruitment of oncology nurses for Internet research: Issues and future directions. Oncology Nursing Forum, 33, 249-254.

International Society of Nurses in Genetics, Inc. (ISONG). (2003a). Academic programs. Retrieved May 18, 2006 from http://www.isong.org/resources/academicprograms.cfm

International Society of Nurses in Genetics, Inc. (ISONG). (2003b). What is a genetic nurse?. Retrieved May 18, 2006 from http://www.isong.org/resources/what_is.pdf

International Society of Nurses in Genetics, Inc. & American Nurses Association. (1998). Statement on the scope and standards of genetics clinical nursing practice. Washington, DC: American Nurses Association.

International Society of Nurses in Genetics, Inc. & American Nurses Association. (2007). Genetics/genomics nursing: Scope and standards of practice. Silver Spring, MD: American Nurses Association.

Jasper, M. (2001). The role of the nurse manager in ensuring competence – the use of portfolios and reflective writing. Journal of Nursing Management, 9, 249-251.

Jenkins, H. M. & Michael, M. M. (1986). Using and interpreting item analysis data. Nurse Educator, 11(1), 10-14.

Jenkins, J., Calzone, K., Lea, D., & Prows, C. (2005). Essential nursing competencies and curricula guidelines for genetics and genomics. Washington, DC: National Human Genome Research Institute.

Jenkins, J. & Collins, F. (2003). Are you genetically literate? American Journal of Nursing, 103(4), 13.

Jenkins, J. F., Dimond, E., & Steinberg, S. (2001). Preparing for the future through genetics nursing education. Journal of Nursing Scholarship, 33, 191-195.

Jenkins, J., Grady, P. A., & Collins, F. S. (2005). Nursing and the genomic revolution. Journal of Nursing Scholarship, 37, 98-101.

Jenkins, J. F., & Lea, D. H. (2005). Nursing care in the genomic era: A case-based approach. Sudbury, MA: Jones and Bartlett.

Jensen, G. M., & Saylor, C. (1994). Portfolios and professional development in the health professions. Evaluation & the Health Professions, 17, 344-357.

Johnson, S. (2002). Development of education competencies and the professional review process. Journal for Nurses in Staff Development, 18(2), 92-102.

Kelly, L. Y. & Joel, L. A. (1996). The nursing experience: Trends, challenges, and transitions (3rd ed.). New York: McGraw-Hill.

Kelly, L. Y. & Joel, L. A. (1999). Dimensions of professional nursing (8th ed.). New York: McGraw-Hill.

Kelly, L. Y. & Joel, L. A. (2003). Dimensions of professional nursing (9th ed.). New York: McGraw-Hill.

Kennedy, M. R. (2005). Informing content- and concept-based image indexing and retrieval through a study of image description. Unpublished master's thesis, University of North Carolina, Chapel Hill.

Kim, J., & Mueller, C. W. (1978). Factor analysis: Statistical methods and practical solutions. Beverly Hills, CA: Sage.

Kittleson, M. J. (1995). An assessment of the response rate via the Postal Service and e-mail. Health-Values, 19(2), 27-39.

Knight, C. F., & Knight, P. F. (1992). Developing a certificate course for hospice nurses: A delphi survey of subject areas. The Hospice Journal, 8(3), 45-57.

Krathwohl, D. R. & Payne, D. A. (1971). Defining and assessing educational objectives. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 17-45). Washington, DC: American Council on Education.

Lackman, L., Nieto, R. D., & Gliem, R. (1997). Instrument development for low literacy audiences: Assessing extension program personnel teaching effectiveness. Journal of Extension, 35 (1). Retrieved July 23, 2004, from http://www.joe.org/joe/1997february/rb1.html

Langenbucher, J. W., Labouvie, E., Martin, C. S., Sanjuan, P. M., Bavly, L., Kirisci, L. et al. (2004). An application of item response theory analysis to alcohol, cannabis, and cocaine criteria in DSM-IV. Journal of Abnormal Psychology, 113, 72-80.

Lashley, F. R. (2005). Clinical genetics in nursing practice (3rd ed.). New York: Springer.

Latvala, E. (2002). Developing and testing methods for improving patient-oriented mental health care. Journal of Psychiatric and Mental Health Nursing, 9, 41-47.

Lea, D. H., Anderson, G., & Monsen, R. (1998). A multiplicity of roles for genetics nursing: Building toward holistic practice. Holistic Nursing Practice, 12, 77-78.

Lea, D. H., Jenkins, J. F. & Francomano, C. A. (1998). Genetics in clinical practice: New directions for nursing and health care. Sudbury, MA: Jones and Bartlett.

Lea, D. H. & Monsen, R. B. (2003). Preparing nurses for a 21st century role in genomics-based health care. Nursing Education Perspectives, 24(2), 75-80.

Lenburg, C. B. (1979, April). Emphasis on evaluating outcomes: The New York Regents External Degree Program. Peabody Journal of Education, 212-221.

Lenburg, C. B. (1999a, September 30). Redesigning expectations for initial and continuing competence for contemporary nursing practice. Online Journal of Issues in Nursing, 4(2). Retrieved November 7, 2002, from http://www.nursingworld.org/ojin/topic10/tpc10_1.htm

Lenburg, C. B.(1999b, September 30). The framework, concepts and methods of the Competency Outcomes and Performance Assessment (COPA) model. Online Journal of Issues in Nursing, 4(2). Retrieved November 7, 2002, from http://www.nursingworld.org/ojin/topic10/tpc10_2.htm

Lenburg, C. B. (2000). Promoting competence through critical self-reflection and portfolio development: The inside evaluator and the outside context. Tennessee Nurse, 63(3), 11+.

Lenburg, C. B. (2002). The influence of contemporary trends and issues on nursing education. In B. Cherry & S. R. Jacob, Contemporary nursing: Issues, trends, & management (2nd ed., pp. 65-94). St Louis: Mosby.

Lenz, E. R. & Hardin, S. B. (2000). Doctoral education: Approaches to achieving quality when resources are limited. In Chaska, N.L. (Ed.), The nursing profession: Tomorrow and beyond (pp. 233-249). Thousand Oakes, CA: Sage.

Lenz, E. R. & Ketefian, S. (1995). Promoting scientific integrity in nursing research part I: Current approaches in doctoral programs. Journal of Professional Nursing, 11, 213-219.

Lewin, M. (1987). Understanding psychological research: The student researcher's handbook. Malabar, FL: Robert E. Krieger.

Lewis, R. (1999). Instructor's manual test item file to accompany human genetics (3rd ed.). Boston: WCB/McGraw-Hill.

Lewis, R. (2005). Human genetics: Concepts and applications (6th ed.). Boston: McGraw-Hill.

Linn, R.L. & Gronlund, N.E. (2005). Measurement and assessment in teaching (9th ed.). Upper Saddle River, NJ: Prentice-Hall.

Loblaw, D. A., Bezjak, A., & Bunston, T. (1999). Development and testing of a visit-specific patient satisfaction questionnaire: The Princess Margaret Hospital Satisfaction with Doctor Questionnaire. Journal of Clinical Oncology,17, 1931.

Loescher, L. J., & Merkle, C. J. (2005). The interface of genomic technologies and nursing. Journal of Nursing Scholarship, 37, 111-119.

Luttrell, M. F., Lenburg, C. B., Scherubel, J. C., Jacob, S. R., & Koch, R. W. (1999). Competency outcomes for learning and performance assessment. Nursing and Health Care Perspectives, 20, 134-141.

Lynn, M. R. (1986). Determination and quantification of content validity. Nursing Research, 35, 382-385.

Martuza, V. R. (1977). Applying norm-referenced and criterion-referenced measurement in education. Boston: Allyn and Bacon.

Mason, D. J. (2001). RN,C: What's in a letter? American Journal of Nursing, 101(1), 7.

Middleton, L. (2002, January). Portfolio scoring process: The cookbook version. Paper presented at the meeting of the Genetic Nursing Credentialing Commission, Inc., Baltimore, MD.

Miles, K., Penny, N., Power, R., & Mercey, D. (2003). Comparing doctor- and nurse-led care in a sexual health clinic: Patient satisfaction questionnaire. Journal of Advanced Nursing, 42(1), 64-72.

Miller, J. & Hogg, A. (2000). Internet vs. telephone data collection: Does method matter? Burke White Paer Series, 2(5). Retrieved July 14, 2003 from http://www.burke.com/whitepapers/PDF/B.WhitePaperVol2-2000-Iss5.pdf

Moloney, M. M. & Schwirian, P. M. (1998). Professionalization of nursing: Current issues and trends (3rd ed.). Philadelphia: Lippincott.

Monsen, R. (1999). America's nurses need ISONG! International Society of Nurses in Genetics, 10(1), 4.

Monsen, R. B., Cook, S. S., Middelton, L., & Kase, R. (2005). Developing a credential based upon portfolio evidence. In R. B. Monsen (Ed.), Genetics nursing portfolios: A new model for credentialing (pp. 55-68). Silver Spring, MD: American Nurses Association.

Moyer, J. E. (2002). The APNG(c): A preliminary look at credentialing nurses through portfolio review. Newborn and Infant Nursing Reviews, 2, 254-258.

Munoz, C., & Hilgenberg, C. (2005). Ethnopharmacology. American Journal of Nursing, 105(8), 40-49.

Munro, B. H. (2004). Statistical methods for health care research (5th ed.). Philadelphia: Lippincott, Williams & Wilkins.

Murrells, T. (2002). Assessing competence to practice in nursing: A review of the literature. Research Papers in Education, 17(1), 51-77.

National Council of State Boards of Nursing (NCSBN). (1998). Assuring competence: A regulatory responsibility. Retrieved May 5, 1998, from http://www.ncsbn.org/files /positions/contcomp.htm

National Council of State Boards of Nursing (NCSBN). (2003, May). Principles of test construction for nurses preparing NCLEX-RN examination and NCLEX-PN examination test items. Educational session presented at Chicago, IL.

Navigating specialty certification. (1999, November). Nursing Management, 30, 28.

Neely, M. A., & Schuley, M. R. (1980, June). Turning experience into competency statements. The Personnel and Guidance Journal, 663-665.

New Target. (2005). Divergent validity. Retrieved November 27, 2005, from http://www. statistics.com/content/glossary/d/divergvalid.php

Nichols, E. F., Lenburg, C. B., & Soehnlen, J. K. (2000). Evaluation of a collaborative articulation mobility project using escrow and transition course methods. Nursing and Health Care Perspectives, 21, 188-195.

Nichols, J. O. (1991). A practitioner's handbook for institutional effectiveness and student outcomes assessment implementation. New York: Agathon Press.

Nicol, M. J. (2003). The variation of response to pharmacotherapy: Pharmacogenetics – A new perspective to 'the right drug for the right person.' MEDSURG Nursing, 12, 242-49.

Nkanginieme, K. E. O. (1997). Clinical diagnosis as a dynamic cognitive process: Application of Bloom's taxonomy for education objectives in the cognitive domain. Med Educ Online, 2(1). Retrieved February 26, 2007 from http://www.ut mb.edu/meo/

Northern Arizona University. (1997). Reliability and validity. Retrieved May 15, 2006 from http://jan.ucc.nau.edu/"mezza/nur390/Mod4/reliability/lesson.html

Nunnally, J. C. (1970). Introduction to psychological measurement. New York: McGraw-Hill.

Nunnally, J. C. (1978). Psychometric theory (2nd ed.). New York: McGraw-Hill.

Nunnally, J. C., & Bernstein, I. H. (1994). Psychometric theory (3rd ed.). New York: McGraw-Hill.

Olsen, S. J., Feetham, S., Jenkins, J., Lewis, J., Nissly, T. Sigmon, H. et al. (2003). Creating a nursing vision for leadership in genetics. MEDSURG Nursing, 12, 177-183.

Palmer, R. H. (1991). Considerations in defining quality in health care. In R.H. Palmer, A. Donabedian, & G.J. Povar, Striving for quality in health care: An inquiry into policy and practice (pp. 3-58). Ann Arbour, MI: Health Administration.

Parker, J. (1994). Development of the American Board of Nursing Specialties (1991-93). Nursing Management, 25, 33-35.

Perkins, G. H. (2004). Will libraries' web-based survey methods replace existing non-electronic survey methods? Information Technology and Libraries, 23(3), 123-125.

Persing, B. F. & Cheek, D. J. (2000). Pharmacogenomics. The Nursing Clinics of North America, 35, 975-980.

Pett, M. A., Lackey, N. R., & Sullivan, J. J. (2003). Making sense of factor analysis: The use of factor analysis for instrument development in health care research. Thousand Oaks: Sage.

Pew Health Professions Commission. (1995). Critical challenges: Revitalizing the health professions for the twenty-first century. San Francisco: UCSF Center for the Health Professions.

Pitkow, J. E., & Recker, M. M. (1995). Using the web as a survey tool: Results from the second WWW user survey. Computer Networks & ISDN Systems, 27, 809-822.

Pitts, J., Coles, C., Thomas, P., & Smith, F. (2002). Enhancing reliability in portfolio assessment: Discussions between assessors. Medical Teacher, 24, 197-201.

Polit, D. F. (1996). Data analysis & statistics for nursing research. Upper Saddle River, NJ: Prentice Hall.

Polit, D. F., & Beck, C. T. (2005). Essentials of nursing research methods, appraisal, and utilization (6th ed.). Philadelphia: Lippincott Williams & Wilkins.

Popham, W. J. (1978). Criterion-referenced measurement. Englewood Cliffs, NJ: Prentice Hall.

Porter-O'Grady, T. (2001). Profound change: 21st century nursing. Nursing Outlook, 49, 182-186.

Prows, C. A., Glass, M., Nicol, M. J., Skirton, H., & Williams, J. (2005). Genomics in nursing education. Journal of Nursing Scholarship, 37, 196-202.

Prows, C. A., Hetteberg, C. J., Latta, N., Lovell, K., Saal, A., Warren, H. M. et al. (2003). Outcomes of a genetics education program for nursing faculty. Nursing Education Perspectives, 24(2), 81-85.

Read, C. Y., Perry, D. J., & Duffy, M. E. (2005). Design and psychometric evaluation of the Psychological Adaptation to Genetic Information Scale. Journal of Nursing Scholarship, 37, 203-208.

Readability formulas. (2004). Retrieved December 22, 2004, from http://csep.psyc.memphis.edu/cohmetrix/readabilityresearch.htm

Redman, R. W., Lenburg, C. B., & Walker, P. H. (1999, September 30). Competency assessment: Methods for development and implementation in nursing education. Online Journal of Issues in Nursing, 4(2). Retrieved November 7, 2002, from http://www.nursingworld.org/ojin/topic10/tpc10_3.htm

Reliability and validity. (n.d.). Retrieved July 23, 2004 from http://seamonkey.e.asu.edu/~alex/teaching/assessment/reliability.html

Rhodes, S. D., Bowie, D. A., & Hergenrather, K. C. (2003). Collecting behavioral data using the world wide Web: Considerations for researchers. Journal of Epidemiology and Community Health, 57, 68-73.

Rizzolo, M.A. (1987). Guidelines for creating test question banks. Computers in Nursing, 5(2), 65-69.

Roberts, D. (2001). It's a 'career credential.' Orthopaedic Nursing, 20(5), 11.

Routledge, J., & Wilson, M. (1997). Reflection on the development of a reflective assessment. Medical Teacher, 19, 122+. Retrieved May 30, 2001, from EBSCOhost (Academic Search Elite) on the World Wide Web: http://ehostvgw18.epnet.com

Saunders, C. S. (1998). Accessing genetics resources. Patient Care Nurse Practitioner, 1(9), 39-48.

Scherubel, J. C. (2002). Nursing licensure and certification. In B. Cherry & S.R. Jacob, Contemporary nursing: Issues, trends, & management (2nd ed., pp. 95-110). St Louis: Mosby.

Schunlev, M. (2001). Coordinating research surveys with email and the web (1st ed.). New York: Rand Corporation.

Schwarz, P. A. (1971). Prediction instruments for educational outcomes. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 303-331). Washington, DC: American Council on Education.

Seguin, J. T. (2005). The concept of the portfolio. In R. B. Monsen (Ed.), Genetics nursing portfolios: A new model for credentialing (pp. 11-24). Silver Spring, MD: American Nurses Association.

Sheehan, K. & Hoy, M. (1999). Using e-mail to survey internet users in the United States: Methodology and assessment. Journal of Computer-Mediated Communication, 4(3), 154-165.

Shirey, M. R. (2005). Celebrating certification in nursing. Nursing Administration Quarterly, 29, 245-253.

Siah, C. Y. (2005). All that glitters is not gold: Examining the perils and obstacles in collecting data on the Internet. International Negotiation, 10, 115-130.

Skirton, H., Patch, C., & Williams, J. (2005). Applied genetics in healthcare: A handbook for specialist practitioners. New York: Taylor & Francis.

Smith, C. (1997). Casting the net: Surveying an internet population. Journal of Computer-Mediated Communication, 3(1), 214-235.

Smolenski, M., Cary, A. H., & Monsen, R. B. (2002, July). Portfolio and evidence-based nursing competency. Educational session (SC076) presented at the American Nurses Association 2002 Biennial Convention, Philadelphia.

Smolenski, M.C., & Gagan, M. J. (2005). Credentialing, certification, and competence: Issues for new and seasoned nurse practitioners. Journal of the American Academy of Nurse Practitioners, 17, 201-204.

Sorkin, D. E. (2005). Spam laws: United States: State laws: Summary. Retrieved July 31, 2006, from http://www.spamlaws.com/state/summary.html

Stanley, J. C. (1971). Reliability. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 356-442). Washington, DC: American Council on Education.

Stanton, M. P. H. (1983). Objective test construction – A must for nursing educators. Journal of Nursing Education, 22, 338-342.

Streiner, D. L., & Norman, G. R. (2003). Health measurement scales: A practical guide to their development and use (3rd ed.). New York: Oxford University Press.

Strickland, O. L. (1997). Challenges in measuring patient outcomes. Nursing Clinics of North America, 32, 495-512.

Strickland, O. L., Burgess, A. W., Oberst, M. T., & Kim, H. S. (1987). Private sector support of nursing research. Nursing Research, 36, 253-56.

Strickland, O. L., Moloney, M. F., Dietrich, A. S., Myerburg, S., Cotsonis, G. A., & Johnson, R. V. (2003). Measurement issues related to data collection on the World Wide Web. Advances in Nursing Science, 26, 246-56.

Stone, B., & Weil, d. (2003, November 24). Soaking in spam. Newsweek, 142, 66.

Styles, M. M., & Affara, A. F. (1998). ICN on regulation: Towards 21st century models. Geneva, Switzerland: International Council of Nurses.

Sweet, E., & Russell, C. (1996). A discussion of data collection via the internet. Retrieved July 14, 2003 from http://www.census.gov/srd/www/abstract/sm96-09.html

Swoboda, S. J., Muehlberger, N., Weitkunat, R., & Schneeweiss, S. (1997). Internet surveys by direct mailing: An innovative way of collecting data. Social Science Computer Review, 15(3), 110-123.

Tambor, E. S., Chase, G. A., Faden, R. R., Geller, G., Hofman, K. J., & Holtzman, N. A. (1993). Improving response rates through incentive and follow-up: The effect on a survey of physicians' knowledge of genetics. American Journal of Public Health, 83, 1599-1603.

Task Force on Accreditation of Health Professions Education. (1999, June 23). New report calls accreditation system 'obsolete' for doctors, nurses, other health professionals. San Francisco: The Center for the Health Professions. Retrieved November 7, 2002, from http://futurehealth.ucsf.edu/press_releases/accred.html

Taylor, I., Thomas, J., & Sage, H. (1999). Portfolios for learning and assessment: Laying the foundations for continuing professional development. Social Work Education, 18, 147-160.

Thorndike, R. L. (1971). Reproducing the test. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 160-187). Washington, DC: American Council on Education.

Tinkelman, S. N. (1971). Planning the objective test. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 46-80). Washington, DC: American Council on Education.

Tinsley, H. E., & Tinsley, D. J. (1987). Uses of factor analysis in counseling psychology research. Journal of Consulting Psychology, 34 (4), 14-24.

Twycross, A,. & Shields, L. (2005). Reliability and validity in practice: Assessment tools. Paediatric Nursing, 17(9), 43.

Van Oppen, P. (1992). Obsessions and compulsions: Dimensional structure, reliability, convergent and divergent validity of the Padua Inventory. Behaviour Research and Therapy, 30, 631-7.

Van Ort, S. & Hazzard, M. E. (1985). A guide for evaluation of test items. Nurse Educator, 10(5), 13-15.

Varnhagen, C. K., Gushta, M., Daniels, J., Peters, T. C., Parmar, N., Law, D. et al. (2005). How informed is online informed consent? Ethics & Behavior, 15(1), 37-48.

Violato, C., Marini, A., & Lee, C. (2003). A validity study of expert judgment procedures for setting cutoff scores on high-stakes credentialing examinations using cluster analysis. Evaluation & The Health Professions, 26(1), 59-72.

Waltz, C. F., Strickland, O. L., & Lenz, E. R. (2005). Measurement in nursing and health research (3rd ed.). New York: Springer.

Web Trends Corporation. (2000). Evaluating internet data collection models: Technological advantages of log file analysis. Retrieved July 14, 2003 from http:// www.pi.danet.de/products/german/whitepapers/wt_data-collection.html

Weber, S. (2006). Developing nurse practitioner student portfolios. Journal of the American Academy of Nurse Practitioners, 18, 301-302.

Weddle, D. O., Himburg, S. P., Collins, N., & Lewis, R. (2002). The professional development portfolio process: Setting goals for credentialing. Journal of the American Dietetic Association, 120, 1439-1444.

Weis, D., & Schank, M. J. (2000). An instrument to measure professional nursing values. Journal of Nursing Scholarship, 32, 201-204.

Wesman, A. G. (1971). Writing the test item. In R. L. Thorndike (Ed.), Educational measurement (2nd ed., pp. 81-129). Washington, DC: American Council on Education.

White, J., Carey, L., & Dailey, K. (2001). Web-based instrumentation in educational survey research. WebNet Journal, 3(1), 46-50

Whittaker, S., Carson, W. & Smolenski, M. C. (2000, June 30). Assuring continued competence—Policy questions and approaches: How should the profession respond? Online Journal of Issues in Nursing, 5. Retrieved March 17, 2003, from http://www.nursingworld.org/ojin/topic10/tpc10_4.htm

Wilde, J. & Sockey, S. (1995). Evaluation handbook. Retrieved June 16, 2002 from http:/ /www.ncbe.gwu.edu/miscpus/eacwest/evalhbk.htm

Wilkinson, T. J., Challis, M., Hobma, S. O., Newble, D. I., Parboosingh, J. T., Sibbald, R. G., et al. (2002). The use of portfolios for assessment of the competence and performance of doctors in practice. Medical Education, 36, 918-924.

Williamson, K. (1999, March 18). Validity and reliability in assessment. Retrieved July 23, 2004 from http://www.siumed.edu/c2000/pervalidity.html

Wilmoth, M. C. (1995). Computer networks as a source of research subjects. Western Journal of Nursing Research, 17, 335-338.

Wright, K. B. (2005). Reserching Internet-based populations: Advantages and disadvantages of online survey research, online questionnaire suthoring software packages, ands web survey services. Journal of Computer-Mediated Communication, 10(3). Retrieved July 20, 2006, from http://jcmc.indiana.edu/vol 10/issue3/wright.html

Wright, S., & Neill, K. (1999). Using the world wide web for research data collection. Clinical Excellence for Nurse Practitioners, 3, 362-365.

Wu, T-Y. & Yu, M-Y. (2003). Reliability and validity of the mammography screening beliefs questionnaire among Chinese American Women. Cancer Nursing, 26, 131-142.

Yun, G., & Trumbo, C. (2000). Comparative response to a survey executed by post, e-mail, and web form. Journal of Computer-Mediated Communication, 6(1), 100-110.

Zieky, M. & Perie, M. (2006). A primer on setting cut scores on tests of educational achievement. Princeton, NJ: Educational Testing Service.

Appendix A
Final CRMTG Content/Bloom Mapping

Question	Genetics	S & S	Bloom	Difficulty	Relevance
1	Adult	Assessment	Comprehension	Moderate	Important
2	Adult	Assessment	Application	Moderate	Important
3	General	Planning	Application	Moderate	Acceptable
4	Pediatrics	Identification	Comprehension	Easy	Important
5	Pediatrics	Identification	Knowledge	Easy	Essential
6	Pediatrics	Diagnosis	Knowledge	Easy	Important
7	Reproductive	Identification	Application	Moderate	Important
8	Adult	Evaluation	Application	Moderate	Important
9	Adult	Planning	Application	Easy	Essential
10	Reproductive	Genetic Psychosocial Counseling	Comprehension	Moderate	Important
11	Reproductive	Genetic Psychosocial Counseling	Application	Moderate	Important
12	Adult	Outcome Identification	Comprehension	Moderate	Important
13	Adult	Identification	Comprehension	Easy	Essential
14	Adult	Identification	Knowledge	Easy	Essential
15	Adult	Identification	Application	Moderate	Essential
16	Adult	Identification	Comprehension	Easy	Important
17	General	Teaching & Research	Comprehension	Easy	Essential
18	General	Identification	Application	Easy	Essential
19	General	Teaching & Research	Knowledge	Easy	Essential
20	Reproductive	Teaching & Research	Comprehension	Easy	Important
21	Adult	Health Promotion/Disease Prevention	Application	Easy	Important
22	Reproductive	Assessment	Comprehension	Moderate	Important
23	General	Knowledge of Genetic Therapeutics	Comprehension	Moderate	Important
24	Adult	Knowledge of Genetic Therapeutics	Application	Moderate	Important
25	General	Teaching & Research	Application	Moderate	Essential

Question	Genetics	S & S	Bloom	Difficulty	Relevance
26	Reproductive	Outcome Identification	Analysis	Moderate	Important
27	Adult	Outcome Identification	Application	Moderate	Important
28	Reproductive	Knowledge of Genetic Therapeutics	Comprehension	Easy	Essential
29	Reproductive	Knowledge of Genetic Therapeutics	Knowledge	Easy	Essential
30	Reproductive	Knowledge of Genetic Therapeutics	Analysis	Moderate	Important
31	General	Diagnosis	Application	Moderate	Important
32	General	Assessment	Application	Moderate	Important
33	General	Identification	Knowledge	Moderate	Important
34	Adult	Teaching & Research	Comprehension	Easy	Important
35	General	Teaching & Research	Knowledge	Easy	Essential
36	Pediatrics	Genetic Psychosocial Counseling	Comprehension	Easy	Important
37	General	Identification	Knowledge	Easy	Essential
38	General	Identification	Knowledge	Moderate	Important
39	Pediatrics	Identification	Knowledge	Easy	Important
40	General	Identification	Knowledge	Easy	Important
41	Adult	Teaching & Research	Knowledge	Moderate	Important
42	Pediatrics	Diagnosis	Knowledge	Easy	Important
43	Adult	Planning	Knowledge	Moderate	Acceptable
44	Reproductive	Identification	Knowledge	Moderate	Acceptable
45	Reproductive	Genetic Psychosocial Counseling	Knowledge	Easy	Important
46	Reproductive	Identification	Comprehension	Easy	Important
47	Pediatrics	Identification	Knowledge	Moderate	Important
48	Adult	Health Promotion/Disease Prevention	Comprehension	Moderate	Essential
49	General	Identification	Knowledge	Moderate	Important
50	General	Identification	Knowledge	Moderate	Essential
51	General	Identification	Knowledge	Moderate	Important

Question	Genetics	S & S	Bloom	Difficulty	Relevance
52	General	Identification	Comprehension	Moderate	Important
53	General	Identification	Knowledge	Moderate	Important
54	General	Identification	Knowledge	Difficult	Acceptable
55	General	Identification	Knowledge	Difficult	Acceptable
56	General	Knowledge of Genetic Therapeutics	Knowledge	Easy	Important
57	General	Knowledge of Genetic Therapeutics	Comprehension	Moderate	Important
58	Adult	Knowledge of Genetic Therapeutics	Comprehension	Easy	Important
59	Adult	Identification	Knowledge	Easy	Acceptable
60	Reproductive	Genetic Psychosocial Counseling	Knowledge	Easy	Important
61	General	Case Coordination	Comprehension	Moderate	Essential
62	Adult	Genetic Psychosocial Counseling	Comprehension	Easy	Important
63	Adult	Knowledge of Genetic Therapeutics	Knowledge	Easy	Essential
64	Adult	Knowledge of Genetic Therapeutics	Comprehension	Moderate	Important
65	Pediatrics	Knowledge of Genetic Therapeutics	Application	Easy	Important
66	General	Knowledge of Genetic Therapeutics	Knowledge	Easy	Acceptable
67	Adult	Teaching & Research	Comprehension	Easy	Essential
68	Adult	Identification	Comprehension	Moderate	Important
69	General	Case Coordination	Analysis	Moderate	Important
70	General	Health Promotion/Disease Prevention	Comprehension	Easy	Important

Question	Genetics	S & S	Bloom	Difficulty	Relevance
71	General	Health Promotion/Disease Prevention	Knowledge	Easy	Important
72	Reproductive	Identification	Application	Easy	Essential
73	General	Outcome Identification	Evaluation	Moderate	Essential
74	General	Identification of Risk	Comprehension	Easy	Essential
75	Reproductive	Identification of Risk	Knowledge	Easy	Important
76	Pediatrics	Case Coordination	Comprehension	Moderate	Important
77	Pediatrics	Therapeutic Communication	Application	Moderate	Important
78	Pediatrics	Genetic Psychosocial Counseling	Application	Moderate	Important
79	Reproductive	Health Promotion/Disease Prevention	Knowledge	Easy	Essential
80	Pediatrics	Assessment	Analysis	Moderate	Important
81	Pediatrics	Diagnosis	Knowledge	Easy	Important
82	Pediatrics	Outcome Identification	Evaluation	Moderate	Important
83	Pediatrics	Planning	Application	Moderate	Important
84	Pediatrics	Evaluation	Analysis	Easy	Important
85	Pediatrics	Assessment	Knowledge	Easy	Important
86	Pediatrics	Identification	Knowledge	Easy	Important
87	Pediatrics	Outcome Identification	Analysis	Moderate	Important
88	Pediatrics	Planning	Application	Moderate	Acceptable
89	Pediatrics	Evaluation	Evaluation	Moderate	Important
90	Pediatrics	Assessment	Knowledge	Easy	Important
91	Pediatrics	Diagnosis	Analysis	Moderate	Important
92	Pediatrics	Outcome Identification	Evaluation	Moderate	Important
93	Pediatrics	Health Promotion/Disease Prevention	Analysis	Moderate	Important
94	Pediatrics	Teaching & Research	Analysis	Moderate	Acceptable
95	Pediatrics	Evaluation	Application	Moderate	Acceptable
96	General	Planning	Application	Easy	Essential
97	Adult	Assessment	Comprehension	Easy	Essential

Question	Genetics	S & S	Bloom	Difficulty	Relevance
98	Adult	Case Coordination	Application	Easy	Essential
99	Adult	Genetic Psychosocial Counseling	Analysis	Moderate	Important
100	Adult	Therapeutic Communication	Analysis	Moderate	Essential

Appendix B
Effectiveness of Distracters

Item	Percent of respondents who selected A	Percent of respondents who selected B	Percent of respondents who selected C	Percent of respondents who selected D	Number of respondents who skipped this item
1	8.6	8.6	31.3	51.5*	123
2	7.8	16.5	15.2	60.6*	125
3	5.2	56.9*	14.2	23.7	124
4	20.2	20.6	3.0	56.2*	123
5	41.7*	22.6	7.4	28.3	126
6	59.6*	22.8	9.6	7.9	128
7	15.8	19.0	28.1	37.6*	135
8	3.1	74.7*	3.1	19.2	127
9	73.7*	11.0	15.4	0.0	128
10	14.8	1.3	32.3*	54.1	127
11	8.6	14.4	23.0	55.0*	134
12	39.4	50.7*	8.6	1.4	135
13	0.9	96.4*	2.3	0.5	134
14	61.7*	8.6	12.2	18.5	134
15	4.5	14.5	5.5	75.5*	136
16	0.9	1.3	5.8	92.4*	132
17	1.3	6.3	60.1*	32.3	133
18	17.0	28.9	8.3	46.8*	138
19	47.5*	21.7	16.3	14.9	135
20	7.6	1.3	80.9*	10.7	131
21	0.9	92.8*	4.5	1.8	135
22	23.6	66.4*	7.7	2.7	136
23	15.6	7.8	16.5	60.1*	138
24	1.4	11.1	0.9	86.6*	139
25	33.6	11.1	51.6*	3.7	139
26	81.8*	0.9	7.3	10.0	136
27	20.4	68.5*	10.2	0.9	140
28	1.4	94.1*	5.4	0.5	135
29	12.0	3.2	2.8	82.0*	139
30	3.7	24.2	4.1	68.0*	137
31	5.8	3.8	25.0	65.9*	148
32	22.7*	33.6	37.9	5.7	145
33	44.7*	11.1	4.3	39.9	148
34	39.6	22.6	29.2*	8.5	144
35	58.3*	31.3	8.1	2.4	145
36	55.2	42.0*	1.4	1.4	144

Item	Percent of respondents who selected A	Percent of respondents who selected B	Percent of respondents who selected C	Percent of respondents who selected D	Number of respondents who skipped this item
37	18.4	8.7	15.0	60.7*	150
38	12.4	59.0*	17.1	11.4	146
39	4.3	22.6	10.1	63.0*	148
40	5.3	59.9*	2.9	34.8	149
41	24.6	64.8*	9.0	1.5	157
42	15.2	38.4*	36.9	9.6	158
43	14.6	18.8	41.1*	25.5	164
44	34.7	9.3	43.0	13.0*	163
45	71.1*	25.9	3.5	0.0	155
46	4.5	46.0	24.7*	26.8	158
47	7.1	38.9	11.6	45.5*	158
48	4.0	10.0	84.6*	1.5	155
49	37.9	8.6	52.0*	1.5	158
50	50.0*	18.9	15.8	15.3	160
51	11.6	55.1	29.3*	4.0	158
52	9.7	15.4	17.4	57.4*	161
53	24.0	14.8	30.6	30.6*	160
54	14.8	17.3	23.5	44.4*	160
55	29.5	44.0	25.4*	1.6	163
56	43.9*	19.9	24.0	12.2	160
57	38.4*	23.2	17.2	21.2	158
58	36.7	6.5	50.8*	6.0	157
59	9.9	69.3*	17.8	3.5	154
60	4.1	27.4*	58.4	10.2	159
61	27.3	23.5	17.1	33.7*	169
62	1.0	86.2*	1.0	12.2	160
63	15.3	60.8*	11.1	13.8	167
64	39.1*	47.4	1.0	13.0	164
65	20.9	18.9	8.2	52.0*	160
66	67.2*	18.5	1.5	12.8	161
67	53.8	1.5	39.5*	6.2	161
68	24.6*	28.3	27.8	19.3	169
69	10.6	14.3	56.6*	18.5	167
70	13.7	71.1*	12.6	2.6	166
71	31.6	7.5	38.5*	23.0	169
72	9.8	60.3*	21.2	9.8	172
73	67.2*	11.6	14.3	11.1	167
74	5.9	28.3	3.7	62.6*	169
75	14.8*	16.4	47.1	23.8	167

Item	Percent of respondents who selected A	Percent of respondents who selected B	Percent of respondents who selected C	Percent of respondents who selected D	Number of respondents who skipped this item
76	1.1	2.6	1.6*	95.8	167
77	39.2	34.9	7.4	19.0*	167
78	13.2	34.4*	1.1	54.0	167
79	94.7*	0.5	5.8	0.5	166
80	11.1	12.1	73.7*	4.7	166
81	85.3*	11.0	2.1	2.6	165
82	3.7	9.6	78.2*	8.5	168
83	2.6	31.2	7.9	58.2*	167
84	49.7	29.6*	4.2	16.4	167
85	75.8*	7.9	8.9	7.9	166
86	4.8	80.6*	2.7	11.8	170
87	18.8	14.0	62.4*	5.4	170
88	3.8	39.8	21.0	36.6*	170
89	19.8	9.1	8.0	63.6*	169
90	65.6*	12.4	15.1	7.5	170
91	33.5	11.6	30.1*	25.4	183
92	48.8*	14.0	30.8	6.4	184
93	19.9	14.2	21.6	46.6*	180
94	18.1	2.8	32.8	48.0*	179
95	18.5	46.6*	15.7	19.1	178
96	10.0	81.1*	7.8	1.7	176
97	79.5*	13.0	7.0	14.1	171
98	4.9	75.1	5.4	17.8*	171
99	11.0	35.4*	27.6	28.7	175
100	79.8*	7.1	8.2	7.1	173

* Key

Appendix C
Permission for test bank

From: "Allen, Deborah" [redacted]
To: "'Jeanine Seguin'" [redacted]
Subject: Permission
Date sent: Mon, 31 Mar 2003 15:02:26 -0600

In regard to your request to use test items for Human Genetics, third edition by Ricki Lewis, you have our permission to use the items you need for your dissertation. I am the developmental editor for McGraw-Hill responsible for the project and have checked with the Sponsoring Editor, Patrick Reidy. He has no objection to your use of the items. Ricki Lewis is
also happy for you to use the items. I have repeated your request below for your records. Thank you for contacting us and good luck with your dissertation.

Dear Dr. Lewis,
I am working on my dissertation and would like to use some items from your test item file for the 3rd edition of Human Genetics. I am developing a tool to evaluate the knowledge base of advanced practice nurses in genetics to be potentially used as an evaluation of the Genetic Nurses Credentialing Commission portfolio process.

A working abstract of my dissertation study is attached.

Do I have your permission to use/modify selected items from this test bank for my dissertation?

Thank you for your consideration,
Jeanine Seguin, DNSc(c), RN, A/GNP
Keuka College
Associate Professor of Nursing
Keuka Park, NY 14478
[redacted]

Deborah Allen
Senior Developmental Editor
McGraw-Hill Higher Education
[redacted]

Appendix D
Sample Items with answers from Criterion-Referenced Measurement Tool for Genetics
(1st draft)

Please select the best answer for each item. There is no penalty for guessing.

1. (d) Some people with polydactyly have more than 5 fingers, while others do not. This is an example of a phenotype that is
    a. codominant.
    b. pleiotropic.
    c. incompletely dominant.
    d. incompletely penetrant.
2. (d) A family has an autosomal dominant condition where the second toe is attached by webbing to the third toe and is longer than the big toe. Only some family members who inherit the mutant gene have the odd toe, and the extent of webbing varies. This phenotype is an example of
    a. codominant and pleiotropic.
    b. codominant and incompletely expressed.
    c. incompletely dominant and pleiotropic.
    d. incomplete penetrance and variable expressivity.
3. (b) For a genetic marker to be useful in tracking an allele known to cause a specific disease, it would have to be present in
    a. all members of the family.
    b. all the family members who show symptoms of the disease.
    c. family members from every generation.
    d. family members who are symptomless.
4. (a) Genomic mismatch scanning is a technique used to compare
    a. similarities in genomes.
    b. differences in genomes.
    c. mutations in similar genes in two species.
    d. wild type to mutant alleles.
5. (d) Sickle cell disease is caused by what type of error?
    a. Deletion.
    b. Translocation.
    c. Expanding triplet repeat.
    d. Point mutation.
6. (a) Which diagnosis is the result of a second-hit theory mutation?
    a. retinoblastoma
    b. neurofibromatosis
    c. Down syndrome
    d. Huntington disease

Appendix E

Sample Items from CRMTG Pilot Tool	Content Validity Indicators
1. Some people with polydactyly have more than 5 fingers, while others do not. This is an example of a phenotype that is a. codominant. b. pleiotropic. c. incompletely dominant. d. incompletely penetrant.	Regarding Item 1 Did you have difficulty responding to this item? Yes No Why? Do you have questions about this item? Yes No What are they? Please write in any revision suggestions.
2. A family has an autosomal dominant condition where the second toe is attached by webbing to the third toe and is longer than the big toe. Only some family members who inherit the mutant gene have the odd toe, and the extent of webbing varies. This phenotype is an example of a. codominant and pleiotropic. b. codominant and incompletely expressed. c. incompletely dominant and pleiotropic. d. incompletely penetrance and variable expressivity.	Regarding Item 2 Did you have difficulty responding to this item? Yes No Why? Do you have questions about this item? Yes No What are they? Please write in any revision suggestions.
3. Genomic mismatch scanning is a technique used to compare a. similarities in genomes. b. differences in genomes. c. mutations in similar genes in two species. d. wild type to mutant alleles.	Regarding Item 3 Did you have difficulty responding to this item? Yes No Why? Do you have questions about this item? Yes No What are they? Please write in any revision suggestions.

Appendix F
Demographics of Pilot Participants (N = 6)

		Number of Pilot Participants	Percent of Pilot Participants	Average Score on Pilot CRMTG
Highest Nursing Degree	Diploma	1	16.7	75.0
	Associate	0	0	n/a
	Baccalaureate	1	16.7	70.0
	Masters	2	33.3	70.5
	Doctorate	2	33.3	82.5
	Missing	0	0	n/a
Primary Clinical Practice Area	Reproductive Genetics	0	0	n/a
	Pediatric Genetics	0	0	n/a
	Adult Genetics	0	0	n/a
	Other Genetic	Prenatal Screening 1 All types of genetics 1	33.3	72.5
	Non-genetic	Midwifery 1 Adult Health 1 Psych Mental Health 1	50.0	77.7
	Non-clinical	0	0	n/a
	Missing	1	16.7	73.0
Years of Genetic Clinical Experience	None	0	0	n/a
	1-5	2	33.3	82.5
	6-10	1	16.7	70.0
	11-15	1	16.7	75.0
	16-20	0	0	n/a
	>20	1	16.7	73.0
	missing	1	16.7	68.0
Hours per Week in Genetics Setting	None	4	66.7	76.5
	1-10	0	0	n/a
	11-20	0	0	n/a
	21-30	0	0	n/a
	31-40	0	0	n/a
	>40	2	33.3	72.5
	Missing	0	0	n/a

		Number of Pilot Participants	Percent of Pilot Participants	Average Score on Pilot CRMTG
Genetic Education (may select all that apply)	On-the-job training	4	66.7	76.5
	Continuing education	6	100	75.2
	College courses	3	50.0	80.0
	Certification program	0	0	n/a
	Degree in genetics	0	0	n/a
	Missing	0	0	n/a
Geographic Region	Northwest	0	0	n/a
	Southwest	0	0	n/a
	Central	2	33.3	75.0
	Northeast	1	16.7	88.0
	Southeast	1	16.7	70.0
	Missing	2	33.3	71.5
Age	< 20	0	0	n/a
	20-29	0	0	n/a
	30-39	0	0	n/a
	40-49	3	50.0	75.3
	50-59	3	50.0	75.0
	60-69	0	0	n/a
	70 or above	0	0	n/a
	Missing	0	0	n/a
Race	Caucasian	5	83.3	76.2
	Hispanic	0	0	n/a
	African-American	0	0	n/a
	Native American	0	0	n/a
	Other	0	0	n/a
	Missing	1	16.7	70.0
Gender	Male	0	0	n/a
	Female	5	83.3	76.2
	Missing	1	16.7	70.0

Appendix G
CRMTG Pilot Data (116-item)

Item number	Percent of pilot participants who answered correctly	Percent of pilot CVI participants who had difficulty responding to question	Percent of pilot CVI participants who had questions about this item
1	100.0	50.0	0
2	83.3	50.0	0
3	66.7	33.3	0
4	0	83.3	16.7
5	83.3	50.0	0
6	100.0	33.3	0
7	83.3	50.0	16.7
8	66.7	83.3	16.7
9	83.3	66.7	0
10	0	66.7	0
11	100.0	33.3	0
12	100.0	33.3	0
13	33.3	66.7	16.7
14	66.7	66.7	33.3
15	66.7	100.0	16.7
16	83.3	16.7	0
17	100.0	0	0
18	100.0	0	0
19	100.0	0	0
20	83.3	16.7	16.7
21	66.7	50.0	0
22	100.0	0	0
23	100.0	33.3	0
24	83.3	16.7	16.7
25	100.0	0	0
26	83.3	83.3	16.7
27	66.7	33.3	0
28	100.0	33.3	0
29	100.0	0	0
30	83.3	50.0	0
31	83.3	16.7	16.7
32	100.0	0	0
33	83.3	16.7	16.7
34	100.0	0	0
35	83.3	0	33.3
36	83.3	16.7	0
37	33.3	50.0	0

Item number	Percent of pilot participants who answered correctly	Percent of pilot CVI participants who had difficulty responding to question	Percent of pilot CVI participants who had questions about this item
38	100.0	33.3	0
39	100.0	16.7	0
40	83.3	50.0	0
41	50.0	33.3	0
42	100.0	0	0
43	100.0	33.3	0
44	100.0	16.7	16.7
45	66.7	83.3	0
46	83.3	66.7	16.7
47	83.3	83.3	16.7
48	100.0	66.7	0
49	33.3	83.3	0
50	33.3	83.3	0
51	16.7	83.3	0
52	100.0	16.7	0
53	16.7	66.7	0
54	83.3	16.7	0
55	83.3	0	0
56	0	100.0	16.7
57	66.7	66.7	0
58	0	100.0	0
59	83.3	33.3	0
60	0	66.7	0
61	66.7	66.7	0
62	66.7	83.3	0
63	66.7	83.3	0
64	33.3	66.7	0
65	33.3	40.0	0
66	50.0	100.0	20.0
67	66.7	20.0	0
68	83.3	60.0	0
69	66.7	40.0	0
70	66.7	100.0	20.0
71	100.0	0	0
72	0	40.0	20.0
73	83.3	60.0	20.0
74	0	100.0	20.0
75	66.7	60.0	20.0
76	100.0	20.0	0
77	0	40.0	40.0

Item number	Percent of pilot participants who answered correctly	Percent of pilot CVI participants who had difficulty responding to question	Percent of pilot CVI participants who had questions about this item
78	83.3	60.0	0
79	0	100.0	0
80	33.3	75.0	0
81	100.0	0	20.0
82	50.0	60.0	0
83	66.7	60.0	0
84	66.7	40.0	0
85	66.7	40.0	0
86	83.3	60.0	0
87	0	80.0	0
88	83.3	60.0	0
89	83.3	0	0
90	33.3	60.0	0
91	33.3	80.0	0
92	0	80.0	20.0
93	16.7	100.0	0
94	0	100.0	0
95	83.3	60.0	0
96	66.7	60.0	0
97	50.0	60.0	0
98	33.3	100.0	0
99	83.3	80.0	0
100	100.0	0	0
101	50.0	40.0	0
102	100.0	20.0	0
103	66.7	80.0	20.0
104	60.0	80.0	0
105	50.0	50.0	0
106	83.3	100.0	0
107	66.7	100.0	40.0
108	50.0	60.0	0
109	66.7	80.0	20.0
110	66.7	80.0	20.0
111	83.3	80.0	0
112	33.3	80.0	20.0
113	100.0	100.0	25.0
114	33.3	100.0	0
115	66.7	80.0	0
116	100.0	75.0	25.0

Appendix H
Sample Item from CRMTG Content Expert Review Tool

32. A hurricane devastates a population living on an island in the Atlantic Ocean, and only a few individuals survive. With little else to do, they produce many children. Several generations later, the replenished population suffers from several inherited disorders that are very rare in other groups. This genetic event is best explained as
    a. a population bottleneck.
    b. genetic load.
    c. a founder effect.
    d. a meteorological mutation.

The key for this item is "a"

I have classified this item as general genetics/evaluation on the CRMTG content map. Do you agree? Yes No

If no, new classification:

Which Bloom Taxonomy level is tested by this question?

____ Knowledge
____ Comprehension
____ Application
____ Analysis
____ Synthesis
____ Evaluation

Please rate this item on the following using the Likert scale where 1 = strongly disagree and 4 = strongly agree

This item represents the content domain of genetic nursing.
1 2 3 4

The question is clearly written.
1 2 3 4

The question and answer are accurate.
1 2 3 4

The question is appropriate for this test.
1 2 3 4

What level of difficulty would you consider this question?

Easy Moderate Difficult

What level of relevance to advanced practice nursing in genetics would you consider this question?

Essential Important Acceptable Questionable

Comments:

Appendix I
Overall CRMTG Content Validity Expert Review Tool

Please rate the CRMTG on the following using the Likert scale where 1 = strongly disagree and 4 = strongly agree

The tool directions were clear.

1 2 3 4

The vocabulary on the tool was appropriate for the level of participants.

1 2 3 4

The length of the tool is appropriate for the subject being evaluated.

1 2 3 4

The arrangement of items is appropriate and logical.

1 2 3 4

The content of the CRMTG reflects advanced practice nursing in genetic knowledge.

1 2 3 4

Please complete the table below by entering the number in each cell which is representative of the percentage of items in the cell (by your item rating above) to which the minimally qualified advanced practice nurse in genetics should be able to respond correctly.

Ebel's grid for standard-setting

		Relevance			
		Essential	Important	Acceptable	Questionable
Difficulty	Easy	%	%	%	%
	Medium	%	%	%	%
	Hard	%	%	%	%

Comments:

Appendix J
Expert CVI Item Outcomes

Item number	Outcome	CRMTG Item Number
1	Rewritten	1
2	No change	2
3	Rewritten	3
4	Deleted	
5	Rewritten	4
6	No change	5
7	No change	6
8	Deleted	
9	No change	7
10	Deleted	
11	Rewritten	8
12	No change	9
13	Rewritten	10
14	Rewritten	11
15	Deleted	
16	Rewritten	12
17	No change	13
18	No change	14
19	Rewritten	15
20	No change	16
21	No change	17
22	Rewritten	18
23	No change	19
24	Rewritten	20
25	Rewritten	21
26	Rewritten	22
27	No change	23
28	No change	24
29	Deleted	
30	Rewritten	25
31	Rewritten	26
32	No change	27
33	Rewritten	28
34	No change	29
35	Rewritten	30
36	Rewritten	31
37	Rewritten	32
38	Rewritten	33
39	Deleted	
40	Deleted	

Item number	Outcome	CRMTG Item Number
41	No change	34
42	No change	35
43	No change	36
44	No change	37
45	No change	38
46	Rewritten	39
47	No change	40
48	No change	41
49	No change	42
50	No change	43
51	Rewritten	44
52	Rewritten	45
53	Rewritten	46
54	Rewritten	47
55	No change	48
56	Deleted	
57	No change	49
58	Deleted	
59	Rewritten	50
60	No change	51
61	No change	52
62	No change	53
63	No change	54
64	No change	55
65	Deleted	
66	Deleted	
67	No change	56
68	No change	57
69	No change	58
70	Deleted	
71	No change	59
72	Rewritten	60
73	No change	61
74	Deleted	
75	Deleted	
76	Rewritten	62
77	Deleted	
78	Deleted	
79	Deleted	
80	Deleted	
81	No change	63
82	Deleted	
83	Rewritten	64

Item number	Outcome	CRMTG Item Number
84	Deleted	
85	No change	65
86	Deleted	
87	Deleted	
88	Rewritten	66
89	No change	67
90	Deleted	
91	Deleted	
92	Deleted	
93	No change	68
94	No change	69
95	No change	70
96	Deleted	
97	Deleted	
98	Rewritten	71
99	Deleted	
100	No change	72
101	Rewritten	73
102	Rewritten	74
103	Rewritten	75
104	Rewritten	76
105	No change	77
106	No change	78
107	Deleted	
108	Rewritten	79
109	Deleted	
110	Rewritten	80
111	No change	81
112	Rewritten	82
113	No change	83
114	Rewritten	84
115	Rewritten	85
116	No change	86
117	Rewritten	87
118	No change	88
119	Rewritten	89
120	No change	90
121	Rewritten	91
122	Rewritten	92
123	No change	93
124	No change	94
125	Rewritten	95
126	No change	96

Item number	Outcome	CRMTG Item Number
127	Deleted	
128	Deleted	
129	Deleted	
130	No change	97
131	No change	98
132	No change	99
133	No change	100

Appendix K
Sample Items from the Criterion-Referenced Measurement Tool for Genetics
(Final version)

DiGeorge syndrome causes abnormal parathyroid glands, disrupting blood calcium levels; heart defects; and an underdeveloped thymus gland, impairing development of the immune system. About 85 percent of patients have a microdeletion of a particular area of chromosome 22. In one family, a girl, her mother, and a maternal aunt have very mild cases of DiGeorge syndrome, and they also have a reciprocal translocation involving chromosomes 22 and 2.

92. What physical assessment findings would be consistent with DiGeorge syndrome?
    a. Epicanthal folds
    b. Synophrys
    c. Malformed ears
    d. Rocker-bottom feet

93. What would be an appropriate nursing diagnosis for an affected individual in this family?
    a. Risk for Injury
    b. Risk for Disuse Syndrome
    c. Risk for Altered Nutrition: More than Body Requirements
    d. Risk for Violence: Directed at Others

94. What would NOT be an expected outcome for a client with DiGeaorge syndrome? The client will:
    a. remain free of falls.
    b. maintain a patent airway
    c. check her pulse rate for 1 minute every day.
    d. receive 2 hours of sunlight every day.

95. The health care team should monitor for all of the following, EXCEPT:
    a. Tetany
    b. Dysrhythmias
    c. Cataracts
    d. Lupus

96. The mother states that she has been doing some research on this syndrome. She asks you why people with the translocation are less severely affected than those with the microdeletion. The best response would be:
    a. "It really varies from family to family how the error is expressed."
    b. "The microdeletion may be more extensive than the deleted region in the translocation individuals."
    c. "A microdeletion means that an entire chromosome is missing."
    d. "A translocation may be more extensive than the deleted region of a microdeletion in many individuals."

97. A positive response from the mother to client education would be:
    a. "I will not have to worry about this syndrome carrying over to the men in the family."
    b. "I will make sure that my daughter has all of her ordered lab work drawn."
    c. "I will make sure that my daughter has a lot of friends around to help her feel good about herself."
    d. "I will not have to worry about my daughter passing this syndrome on to may grandchildren."

Appendix L
School Contact Letter

Dear Colleague,

I am a doctoral student at Widener University School of Nursing in Chester, PA. I am currently collecting data to complete my doctoral dissertation. My study involves the development of a tool to measure advanced practice nursing in genetics knowledge. I am writing to ask for your assistance with this study. I would like to have a request for participants sent to the recent graduates of your Masters in Nursing program who specialized in genetics.

Members who are willing to participate in this study will be asked to log on to a website and spend about one hour taking a test. Subjects will be asked to retake the test one month later, to calculate test reliability. Upon completion of the second test subjects may email me to be entered into a drawing for one of four $100 amazon.com gift certificates.

Benefits to subjects include the opportunity to further nursing research and the validation of the current genetic nursing credentialing process. There are no known risks to study participants. I will use the demographic information to describe the sample in aggregate only. Anonymity will be maintained. Any questions or concerns about this study can be sent to me at: [redacted] If participants would like results of this study, they can contact me at this email address.

My dissertation chairperson is Dr. Elizabeth Bayley, Professor of Nursing. Dr. Bayley can be contacted at [redacted] or [redacted]. If you have any questions about the rights of research participants you may contact Dr. Barbara Patterson, Chairperson of Widener University's Institutional Review Board at [redacted] or [redacted]. The Widener University Institutional Review Board has approved the solicitation of participants for this study.

Your cooperation is important in helping me to collect data to complete this study. Thank you for your time and consideration.

Sincerely,
Jeanine Seguin, DNSc (candidate), APRN, BC

Appendix M
Organization Contact Letter

Dear Colleague,

I am a doctoral student at Widener University School of Nursing in Chester, PA. I am currently collecting data to complete my doctoral dissertation. My study involves the development of a tool to measure advanced practice nursing in genetics knowledge. I am writing to ask for your assistance with this study. I would like to have a request for participants sent to the membership of your organization.

Participants do not need to have any knowledge of genetics to participate in this study. In fact, I need a large number of nurses who do not practice in a genetics setting in order to assess tool validity and reliability.

Members who are willing to participate in this study will be asked to log on to a website and spend about one hour taking a test. Subjects will be asked to retake the test one month later, to calculate test reliability. Upon completion of the second test subjects may email me to be entered into a drawing for one of four $100 amazon.com gift certificates.

Benefits to subjects include the opportunity to further nursing research and the validation of the current genetic nursing credentialing process. There are no known risks to study participants. I will use the demographic information to describe the sample in aggregate only. Anonymity will be maintained. Any questions or concerns about this study can be sent to me at: [redacted]. If participants would like results of this study, they can contact me at this email address.

My dissertation chairperson is Dr. Elizabeth Bayley, Professor of Nursing. Dr. Bayley can be contacted at [redacted] or [redacted]. If you have any questions about the rights of research participants you may contact Dr. Barbara Patterson, Chairperson of Widener University's Institutional Review Board at [redacted] or [redacted]. The Widener University Institutional Review Board has approved the solicitation of participants for this study.

Your cooperation is important in helping me to collect data to complete this study. Thank you for your time and consideration.

Sincerely,
Jeanine Seguin, DNSc (candidate), APRN, BC

Appendix N
Protection of Rights of Human Subjects Approval

Widener University

Office of the Provost

Memorandum

To: Jeanine Seguin, MS, APRN, BC

From: Dr. Barbara Patterson
Chairperson, Widener University Institutional Review Board

Date: April 15, 2005

RE: **Protection of Rights of Human Subjects Review**

This letter serves to inform you that your research, **A Psychometric Analysis of an Instrument that Tests Genetic Knowledge of Advanced Practice Nursing in Genetics Patterned on the International Society of Nurses in Genetics Scope and Standards of Genetic Clinical Practice (#27-05)** has been reviewed and **approved** by the Widener University Institutional Review Board (IRB) for the protection of rights of human subjects. You may begin data collection as proposed in your application.

If, for any reason, the approved research data collection method changes significantly, you are required to notify the IRB, in writing, of such changes. Please, remember that the IRB committee and Widener University accept no responsibility for liabilities associated with this study. Ultimately, responsibility rests with the investigator.

The approval of this study is in effect for one year from the date of approval and is eligible at that time for renewal. Upon completion of the study, a final written report of the research is to be submitted to the IRB.

The members of the IRB extend their best wishes for your successful completion of this research project. If you have any questions, please call me at [redacted]. Thank you.

[redacted]

Barbara Patterson, PhD, RN

CC: Dr. E. Bayley

Appendix O
Groups A, B & C Letter

Dear Colleagues,

I am a doctoral candidate at Widener University School of Nursing in Chester, PA. I am currently collecting data to complete my doctoral dissertation. My study involves the development of a tool to measure advanced practice nursing in genetics knowledge. I am writing to ask for your assistance with this study.

Your participation involves logging on to the tool at the following site: http://www.surveymonkey.com/s.asp?u=35731109750 The test should take approximately one hour to complete. You will be asked to retake the test one month later, to calculate test reliability. Upon completion of the second test you may email me at [redacted] to be entered into a drawing for one of four $100 amazon.com gift certificates.

Benefits to you include the opportunity to further nursing research and the validation of the current genetic nursing credentialing process. There are no known risks to you. I will use the demographic information to describe the sample in aggregate only. Anonymity will be maintained.

Accessing the survey website implies your consent to participate in this study. If you change your mind prior to completing the test, you may withdraw from the study by logging off of the website without submitting your responses. Once you submit your responses to the test items they will become part of the study database and cannot be withdrawn.

Any questions or concerns about this study can be sent to me at: [redacted] If you would like results of this study, please contact me at this email address.

My dissertation chairperson is Dr. Elizabeth Bayley, Professor of Nursing. Dr. Bayley can be contacted at [redacted] or [redacted]. If you have any questions about the rights of research participants you may contact Dr. Barbara Patterson, Chairperson of Widener University's Institutional Review Board at [redacted] or [redacted]. The Widener University Institutional Review Board has approved the solicitation of participants for this study.

If you are willing to participate in this study please log on to the tool at the following site: http://www.surveymonkey.com/s.asp?u=35731109750

Your cooperation is important in helping me to collect data to complete this study. Thank you for your time and consideration.

Sincerely,
Jeanine Seguin, DNSc. (candidate), APRN, BC

Appendix P
Groups D & E Letter

Dear Colleagues,

I am a doctoral candidate at Widener University School of Nursing in Chester, PA. I am currently collecting data to complete my doctoral dissertation. My study involves the development of a tool to measure advanced practice nursing in genetics knowledge. I am writing to ask for your assistance with this study.

You do not need to have any knowledge of genetics to participate in this study. In fact, I need a large number of nurses who do not practice in a genetics setting in order to assess tool validity and reliability.

Your participation involves logging on to the tool at the following site: http://www.surveymonkey.com/s.asp?u=35731109750 The test should take approximately one hour to complete. You will be asked to retake the test one month later, to calculate test reliability. Upon completion of the second test you may email me at [redacted] to be entered into a drawing for one of four $100 amazon.com gift certificates.

Benefits to you include the opportunity to further nursing research and the validation of the current genetic nursing credentialing process. There are no known risks to you. I will use the demographic information to describe the sample in aggregate only. Anonymity will be maintained.

Accessing the survey website implies your consent to participate in this study. If you change your mind prior to completing the test, you may withdraw from the study by logging off of the website without submitting your responses. Once you submit your responses to the test items they will become part of the study database and cannot be withdrawn.

Any questions or concerns about this study can be sent to me at: [redacted] If you would like results of this study, please contact me at this email address.

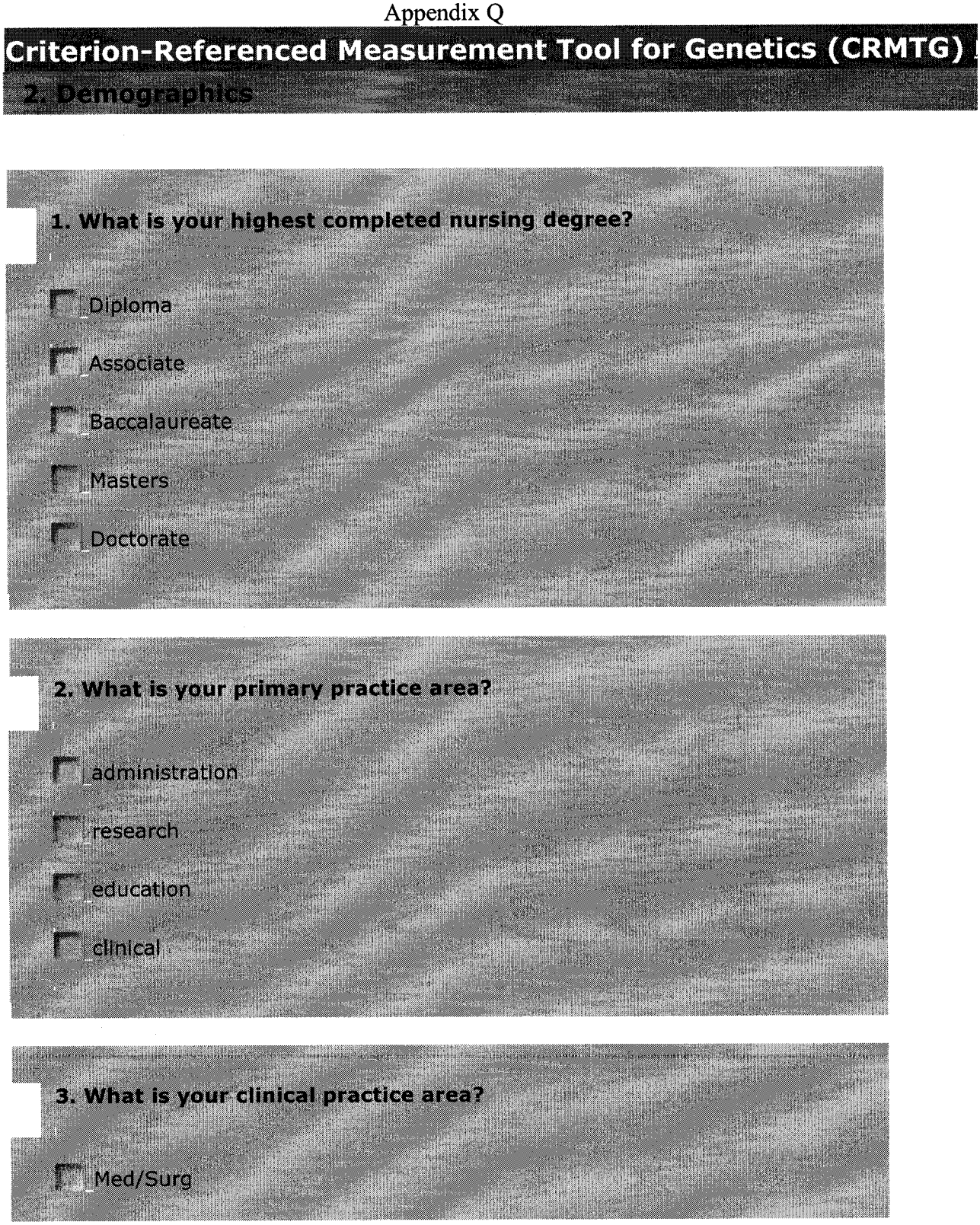
My dissertation chairperson is Dr. Elizabeth Bayley, Professor of Nursing. Dr. Bayley can be contacted at [redacted] or [redacted]. If you have any questions about the rights of research participants you may contact Dr. Barbara Patterson, Chairperson of Widener University's Institutional Review Board at [redacted] or [redacted]. The Widener University Institutional Review Board has approved the solicitation of participants for this study.

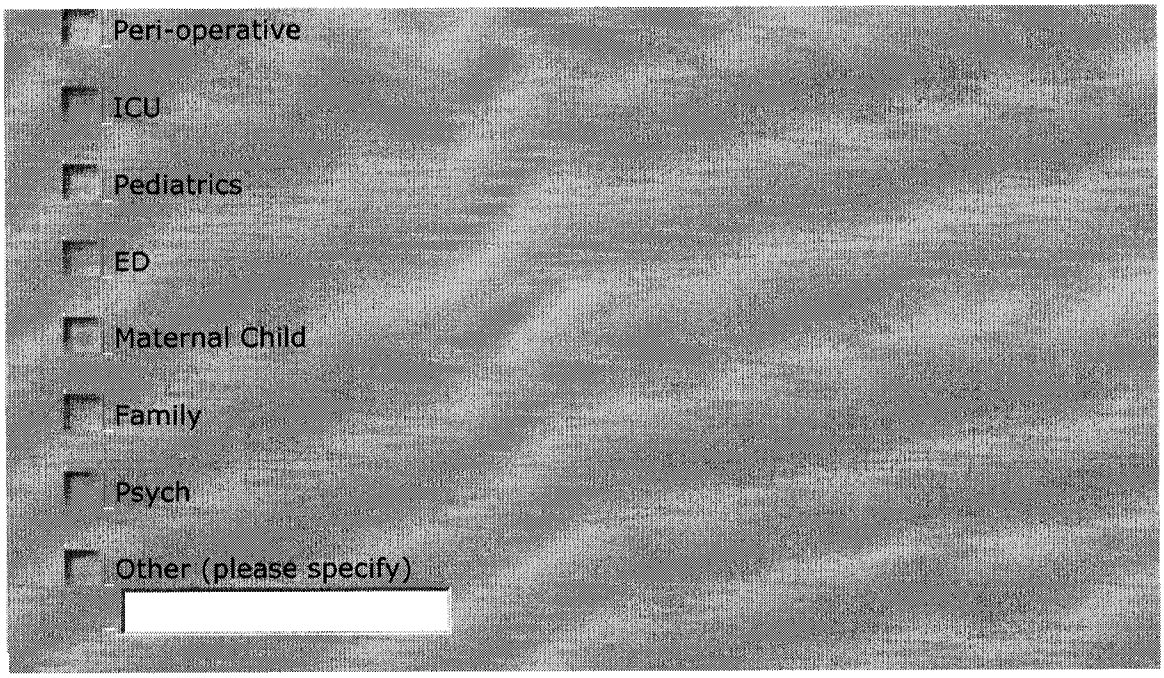
If you are willing to participate in this study please log on to the tool at the following site: http://www.surveymonkey.com/s.asp?u=35731109750

Your cooperation is important in helping me to collect data to complete this study. Thank you for your time and consideration.

Sincerely,
Jeanine Seguin, DNSc. (candidate), APRN, BC

Appendix Q





4. In what geographic region of the country do you practice?

- Northwest
- Northeast
- Southwest
- Southeast
- Central

5. How many years of nursing experience do you have?

6. What is your age?

7. What is your race?

- African-American
- Asian
- Caucasian
- Hispanic
- Native American
- Other (please specify)

8. What is your gender?

- Male
- Female

9. If you specialize in genetics, what is your primary clinical practice area?

- Reproductive genetics

- [ ] Pediatric genetics
- [ ] Adult genetics
- [ ] Other (please specify)

10. How many years of genetic nursing experience do you have?

11. How many hours per week do you practice in a genetics setting?

12. What genetic education have you received?

- [ ] on-the-job training
- [ ] certification program
- [ ] continuing education
- [ ] degree in genetics
- [ ] college courses
- [ ] Other (please specify)

Next >>

Appendix R

Criterion-Referenced Measurement Tool for Genetics (CRMTG) Exit this surv

1. Welcome

Dear Participant

Thank you for your interest in participating in the development of a tool to compare the results of a standardized exam to a portfolio review for the awarding of advanced practice genetic clinical nursing credentials. The test should take approximately one hour to complete. I will use the demographic information to code the data. No one outside the research committee will know your score. Confidentiality will be maintained.

There will be no compensation, other than your entry into a drawing at the completion of the second administration of the tool, to you for completing this study. You will, however, be providing valuable data to myself and furthering the efforts of advanced practice genetic nursing credentialing. When you click "next" to begin the tool, your consent will be implied. If you change your mind prior to completing the test, you may withdraw from the study by closing the tool without submitting your data. Once you complete the tool click "done."

You will be contacted again in one month and asked to retake the CRMTG. After the second submission, if you would like to enter the drawing for one of four $100 amazon.com gift certificates, you will be instructed how to enter the drawing.

Any questions or concerns about this study can be sent to me at: jeanineseguin1@adelphia.net. If you would like results of this study, please contact me at this email address.

My dissertation chairperson is Dr. Elizabeth Bayley, Professor of Nursing. Dr. Bayley can be contacted at [redacted] If you have any questions about the rights of research participants you may contact Dr.

Barbara Patterson, Chairperson of Widener University's Institutional Review Board at [redacted] The Widener University Institutional Review Board has approved the solicitation of participants for this study.

Your cooperation is important in helping me to collect data to complete this study. Thank you for your time and consideration.

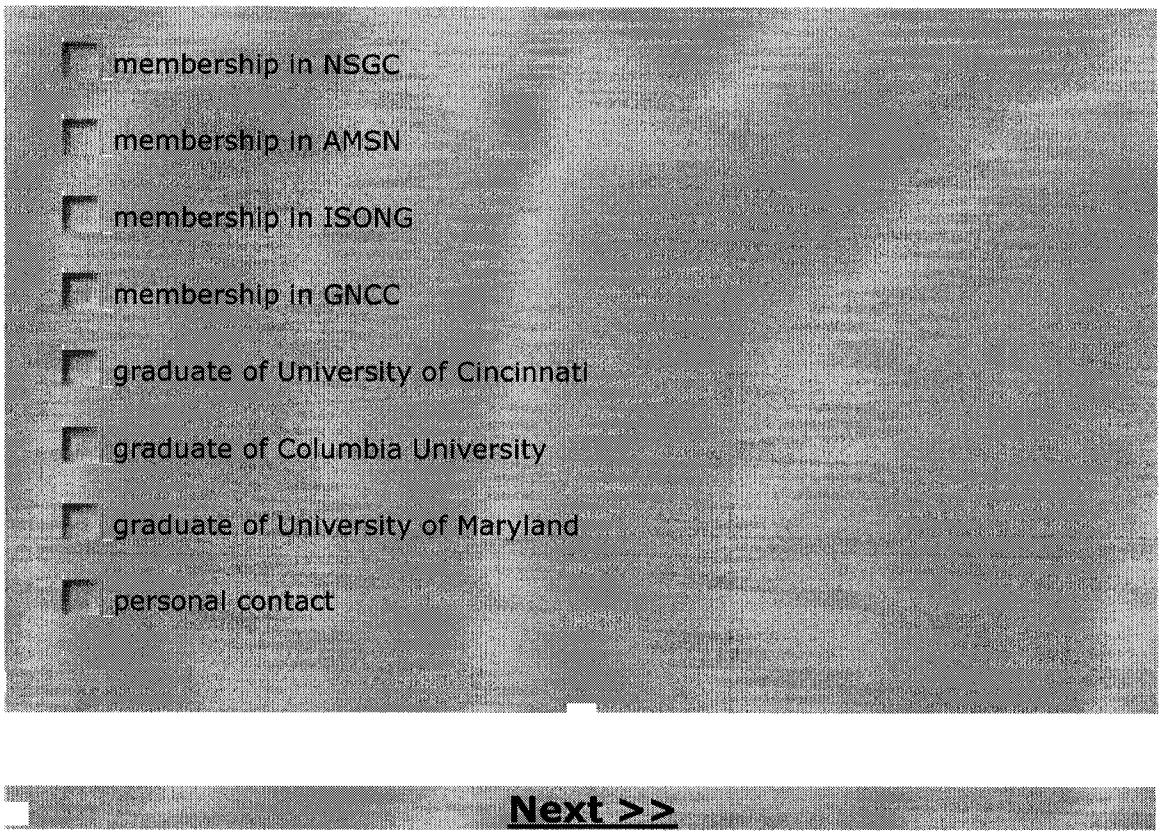
Thank you,

Jeanine Seguin, DNSc (candidate), APRN, BC

1. For test/retest contact and tracking purposes only, your preferred email address is:

2. You found out about this study through:

- membership in AACN
- membership in ONS
- membership in ENA
- membership in AANP
- membership in ACNP
- membeship in AONE
- membership in NACNS
- membership in APNA



Appendix S

Contacts and Participants by Month of Data Collection

	6/05	7/05	8/05	9/05	10/05	11/05	12/05	Total
Estimated Number of Contacts	250	23,286	1,757	2,550	12,529	0	0	40,372
Actual Number of Participants	11	34	88	71	60	90	2	356

Appendix T
Retest Request Letter

Dear Colleague,

Thank you so much for participating in this study. Please take the time to complete the CRMTG for a second time. The data from this repeat sampling will provide test-retest reliability data.

Your participation involves logging on to the tool at the following site: http://www.surveymonkey.com/s.asp?u=35731109750 In order to re-access the site you will need to either log on using a different computer or delete cookies. The test should take approximately one hour to complete. Upon completion of this second test you will automatically be entered into a drawing for one of four $100 amazon.com gift certificates.

Benefits to you include the opportunity to further nursing research and the validation of the current genetic nursing credentialing process. There are no known risks to you. I will use the demographic information to describe the sample in aggregate only. Anonymity will be maintained.

Accessing the survey website implies your consent to participate in this study. If you change your mind prior to completing the test, you may withdraw from the study by logging off of the website without submitting your responses. Once you submit your responses to the test items they will become part of the study database and cannot be withdrawn.

Any questions or concerns about this study can be sent to me at: [redacted]. If you would like results of this study, please contact me at this email address.

My dissertation chairperson is Dr. Elizabeth Bayley, Professor of Nursing. Dr. Bayley can be contacted at [redacted]. If you have any questions about the rights of research participants you may contact Dr. Barbara Patterson, Chairperson of Widener University's Institutional Review Board at [redacted] The Widener University Institutional Review Board has approved the solicitation of participants for this study.

If you are willing to participate in this study please log on to the tool at the following site: http://www.surveymonkey.com/s.asp?u=35731109750

Your cooperation is important in helping me to collect data to complete this study. Thank you for your time and consideration.

Sincerely,
Jeanine Seguin, DNSc. (candidate), APRN, BC

Appendix U
CRMTG Item Analysis

Item	Item-to-total correlation (significance)	Items not significant on item-to-item analysis	Item discrimination index	Item difficulty (%)
1	.632 (.000)	75, 76	.23	51.5*
2	.703 (.000)	76	.27	60.6*
3	.664 (.000)	76	.22	56.9*
4	.585 (.000)	75, 76, 92	.16	56.2*
5	.689 (.000)	76, 92	.32	41.7*
6	.695 (.000)	76	.25	59.6*
7	.466 (.000)	76, 77, 98	.12	37.6*
8	.740 (.000)	76	.22	74.7*
9	.689 (.000)	76	.15	73.7*
10	.415 (.000)	45, 68, 71, 76, 92, 93	.12	32.3
11	.514 (.000)	76, 77, 98	.08	55.0*
12	.655 (.000)	75, 76	.22	50.7*
13	.801 (.000)	76	.09	96.4*
14	.703 (.000)	76	.26	61.7*
15	.746 (.000)	76	.22	75.5*
16	.745 (.000)	76	.07	92.4*
17	.593 (.000)	76, 98	.13	60.1*
18	.707 (.000)	76, 92	.33	46.8*
19	.653 (.000)	76, 92	.26	47.5*
20	.756 (.000)	76	.17	80.9*
21	.774 (.000)	76	.11	92.8*
22	.743 (.000)	76	.27	66.4*
23	.596 (.000)	75, 76, 98	.15	60.1*
24	.572 (.000)	76	.19	86.6*
25	.654 (.000)	76, 77, 92	.24	51.6*
26	.722 (.000)	76	.11	81.8*
27	.756 (.000)	75, 76	.28	68.5*
28	.789 (.000)	76	.10	94.1*
29	.775 (.000)	76	.21	82.0*
30	.717 (.000)	75, 76	.24	68.0*
31	.770 (.000)	76	.34	65.9*
32	.440 (.000)	46, 75, 76, 84, 92, 98, 99	.14	22.7
33	.698 (.000)	76, 92	.30	44.7*
34	.620 (.000)	76, 92	.26	29.2
35	.741 (.000)	76	.32	58.3*
36	.701 (.000)	76, 92	.32	42.0
37	.752 (.000)	76	.35	60.7*
38	.713 (.000)	76	.29	59.0*

Item	Item-to-total correlation (significance)	Items not significant on item-to-item analysis	Item discrimination index	Item difficulty (%)
39	.740 (.000)	76	.30	63.0*
40	.620 (.000)	75, 76	.19	59.9*
41	.700 (.000)	76	.26	64.8*
42	.472 (.000)	76, 92, 98	.11	38.4*
43	.413 (.000)	75, 76, 77, 94, 98	.07	41.1*
44	.414 (.000)	76, 84, 92, 93	.11	13.0
45	.634 (.000)	10, 76	.16	71.1*
46	.385 (.000)	32, 75, 76, 77, 84	.12	24.7
47	.725 (.000)	76, 92	.35	45.5*
48	.812 (.000)	76	.28	84.6*
49	.582 (.000)		.17	52.0*
50	.721 (.000)	76	.34	50.0*
51	.585 (.000)	76, 92	.23	29.3
52	.760 (.000)	76	.37	57.4*
53	.629 (.000)	76, 92	.28	30.6*
54	.642 (.000)	76	.29	44.4*
55	.438 (.000)	76, 99	.15	25.4
56	.698 (.000)	76, 92	.32	43.9*
57	.658 (.000)	76, 92	.30	38.4*
58	.633 (.000)	76	.25	50.8*
59	.739 (.000)	76	.31	69.3*
60	.411 (.000)	76, 92, 98, 99	.13	27.4
61	.619 (.000)	76, 92	.26	33.7*
62	.782 (.000)	76	.25	86.2*
63	.764 (.000)	76	.38	60.8*
64	.543 (.000)	76, 98	.20	39.1
65	.646 (.000)	76, 98	.25	52.0*
66	.714 (.000)	76	.29	67.2*
67	.590 (.000)	75, 76	.24	39.5
68	.475 (.000)	10, 76, 92, 98	.16	24.6
69	.589 (.000)	76	.21	56.6*
70	.736 (.000)	76	.30	71.1*
71	.530 (.000)	10, 76	.19	38.5*
72	.758 (.000)	76	.38	60.3*
73	.680 (.000)	76, 98	.25	67.2*
74	.741 (.000)	76	.35	62.6*

Item	Item-to-total correlation (significance)	Items not significant on item-to-item analysis	Item discrimination index	Item difficulty (%)
75	.259 (.000)	1, 4, 12, 23, 27, 30, 32, 40, 43, 46, 67, 76, 77, 91, 92, 94, 95	.03	14.8
76	.013 (.805)	All except 49 & 77	-.01	1.6
77	.313 (.000)	7, 11, 25, 43, 46, 75, 84, 90	.31	19.0
78	.508 (.000)	76	.21	34.4
79	.794 (.000)	76	.22	94.7*
80	.758 (.000)	76	.30	73.7*
81	.783 (.000)	76	.26	85.3*
82	.793 (.000)	76	.32	78.2*
83	.712 (.000)	76	.32	58.2*
84	.403 (.000)	32, 44, 46, 76, 77, 98	.12	26.6
85	.742 (.000)	76	.27	75.8*
86	.753 (.000)	76	.26	80.6*
87	.676 (.000)	76	.26	62.4*
88	.673 (.000)	76, 92	.32	36.6
89	.669 (.000)	76	.26	63.6*
90	.671 (.000)	76, 77	.25	65.6*
91	.448 (.000)	75, 76, 92	.15	30.1
92	.342 (.000)	75, 76, 98, 99	.04	48.8*
93	.443 (.000)	10, 44, 76	.09	46.6*
94	.504 (.000)	43, 75, 76	.17	48.0*
95	.662 (.000)	75, 76	.31	46.6*
96	.739 (.000)	76	.26	81.1*
97	.700 (.000)	76	.22	79.5*
98	.299 (.000)	76, 84	.06	17.8
99	.429 (.000)	32, 55, 60, 76, 92	.11	35.4*
100	.716 (.000)	76	.24	3.8

* Most popular response

Appendix V
Test-Retest Scores by Participant

Participant	Test Score	Retest Score
1	36	42
2	54	62
3	68	69
4	31	33
5	38	39
6	25	79
7	88	89
8	79	81
9	84	82
10	87	90
11	71	80
12	36	45
13	78	81
14	80	77
15	83	79
16	81	82
17	64	67
18	89	84
19	76	73
20	39	26
21	45	53
22	52	42
23	46	48
24	51	51
25	49	56
26	33	34
27	31	34
28	37	28
29	21	29
30	49	49
31	35	45
32	42	48
33	38	45
34	46	36
35	49	44
36	27	46

Appendix W
CRMTG Item Analysis by Group

Item	Group A % Correct	Group B % Correct	Group C % Correct	Group D % Correct	Group E % Correct
1	63.2	100.0	69.7	27.2	10.9
2	77.2	100.0	72.7	31.4	15.2
3	78.9	80.0	72.7	30.2	8.7
4	66.7	80.0	69.7	33.1	10.9
5	73.7	60.0	75.8	11.8	6.5
6	73.7	80.0	72.7	31.4	14.1
7	38.6	80.0	45.5	20.7	7.6
8	77.2	100.0	72.7	46.2	21.7
9	78.9	100.0	75.8	44.4	19.6
10	5.1	20.0	42.4	18.9	7.6
11	42.1	40.0	48.5	38.5	16.3
12	59.6	80.0	66.7	24.9	10.9
13	80.7	100.0	69.7	65.1	32.6
14	77.2	100.0	66.7	33.1	10.9
15	73.7	100.0	69.7	46.2	19.6
16	71.9	100.0	66.7	66.9	28.3
17	40.4	20.0	63.6	42.6	18.5
18	70.2	100.0	72.7	15.4	7.6
19	63.2	100.0	60.6	22.5	6.5
20	75.4	100.0	72.7	52.7	22.8
21	75.4	80.0	69.7	63.3	30.4
22	71.9	100.0	72.7	35.5	17.4
23	56.1	80.0	60.6	36.7	14.1
24	70.2	100.0	72.7	57.4	23.9
25	64.9	20.0	63.6	26.6	8.7
26	70.2	80.0	66.7	53.8	25.0
27	71.9	100.0	42.1	34.9	20.7
28	77.2	100.0	66.7	66.9	28.3
29	78.9	100.0	72.7	51.5	18.5
30	68.4	100.0	69.7	38.5	18.5
31	73.7	100.0	63.6	33.1	14.1
32	22.8	60.0	36.4	9.5	4.3
33	61.4	60.0	57.6	17.8	6.5
34	50.9	20.0	57.6	5.3	4.3
35	68.4	100.0	69.7	26.0	13.0
36	59.6	60.0	72.7	10.7	10.9
37	68.4	80.0	69.7	24.9	16.3
38	63.2	100.0	66.7	27.8	15.2

Item	Group A % Correct	Group B % Correct	Group C % Correct	Group D % Correct	Group E % Correct
39	66.7	100.0	72.7	26.6	20.7
40	52.6	80.0	45.5	33.1	20.7
41	54.4	60.0	57.6	33.7	20.7
42	29.8	40.0	48.5	19.5	8.7
43	28.1	0.0	30.3	24.9	12.0
44	22.8	40.0	18.2	2.4	0.0
45	50.9	60.0	60.6	43.2	19.6
46	26.3	20.0	21.2	11.8	6.5
47	66.7	60.0	72.7	12.4	6.5
48	75.4	80.0	72.7	45.0	25.0
49	45.6	0.0	51.5	26.6	16.3
50	59.6	60.0	66.7	17.8	9.8
51	45.6	20.0	48.5	5.3	6.5
52	68.4	80.0	72.7	22.5	7.6
53	61.4	20.0	48.5	3.0	5.4
54	52.6	60.0	60.6	14.2	10.9
55	26.3	20.0	33.3	9.5	6.5
56	57.9	80.0	60.6	14.2	5.4
57	56.1	60.0	57.6	10.1	5.4
58	52.6	80.0	54.5	22.9	12.0
59	70.2	80.0	63.6	38.5	10.9
60	24.6	20.0	33.3	11.2	9.8
61	45.6	40.0	54.5	7.7	4.3
62	70.2	60.0	66.7	50.3	20.7
63	66.7	80.0	72.7	21.9	10.9
64	35.1	20.0	42.4	18.9	8.7
65	50.9	80.0	51.5	27.2	6.5
66	57.9	80.0	60.6	33.7	18.5
67	45.6	60.0	48.5	15.4	6.5
68	31.6	20.0	42.4	5.9	3.3
69	42.1	60.0	48.5	29.0	16.3
70	63.2	60.0	60.6	36.1	16.3
71	33.3	60.0	42.4	16.0	9.8
72	61.4	40.0	72.7	23.7	10.9
73	52.6	60.0	60.6	34.3	15.2
74	63.2	80.0	69.7	25.4	10.9
75	8.8	0.0	24.2	6.5	4.3
76	0.0	0.0	0.0	0.6	2.2
77	17.5	40.0	6.0	8.9	7.6
78	28.1	0.0	39.4	14.2	13.0
79	66.7	80.0	63.6	54.4	26.1
80	70.2	80.0	57.6	36.1	16.3

Item	Group A % Correct	Group B % Correct	Group C % Correct	Group D % Correct	Group E % Correct
81	66.7	80.0	72.7	43.8	23.9
82	66.7	80.0	69.7	39.1	17.4
83	61.4	80.0	60.6	25.4	8.7
84	22.8	0.0	36.4	13.6	8.7
85	66.7	80.0	60.6	37.9	19.6
86	61.4	60.0	66.7	40.2	23.9
87	54.4	40.0	48.5	29.6	18.5
88	59.6	40.0	54.5	6.5	3.3
89	54.4	60.0	60.6	30.2	14.1
90	54.4	60.0	48.5	34.9	14.1
91	24.6	40.0	33.3	11.2	6.5
92	19.3	40.0	6.1	34.3	12.0
93	33.3	20.0	24.2	24.3	14.1
94	31.6	20.0	39.4	24.3	13.0
95	49.1	60.0	51.5	15.4	9.8
96	56.1	40.0	60.6	43.2	20.7
97	56.1	80.0	60.6	42.0	21.7
98	21.1	0.0	12.1	7.1	5.4
99	31.6	20.0	27.3	16.0	7.6
100	56.1	60.0	69.7	43.2	16.3

Appendix X
Comparison of Expert Panel, Total Sample, and Subsample Difficulty Results by CRMTG Item

Item number	Expert difficulty rating	Item level of difficulty (N = 356)	Percent Correct for Subsample (n = 195)	Item number	Expert difficulty rating	Item level of difficulty (N = 356)	Percent Correct for Subsample (n = 195)
1	Moderate	51.5	51	51	Moderate	29.3	29
2	Moderate	60.6	61	52	Moderate	57.4	55
3	Moderate	56.9	57	53	Moderate	30.6	29
4	Easy	56.2	54	54	Difficult	44.4	44
5	Easy	41.7	43	55	Difficult	25.4	24
6	Easy	59.6	59	56	Easy	43.9	43
7	Moderate	37.6	36	57	Moderate	38.4	38
8	Moderate	74.7	76	58	Easy	50.8	49
9	Easy	73.7	72	59	Easy	69.3	68
10	Moderate	32.3	32	60	Easy	27.4	27
11	Moderate	55.0	57	61	Moderate	33.7	32
12	Moderate	50.7	51	62	Easy	86.2	87
13	Easy	96.4	95	63	Easy	60.8	58
14	Easy	61.7	61	64	Moderate	39.1	38
15	Moderate	75.5	73	65	Easy	52.0	52
16	Easy	92.4	92	66	Easy	67.2	66
17	Easy	60.1	63	67	Easy	39.5	39
18	Easy	46.8	46	68	Moderate	24.6	23
19	Easy	47.5	47	69	Moderate	56.6	54
20	Easy	80.9	82	70	Easy	71.1	69
21	Easy	92.8	92	71	Easy	38.5	37
22	Moderate	66.4	67	72	Easy	60.3	57
23	Moderate	60.1	58	73	Moderate	67.2	64
24	Moderate	86.6	91	74	Easy	62.6	59
25	Moderate	51.6	50	75	Easy	14.8	14
26	Moderate	81.8	81	76	Moderate	1.6	02
27	Moderate	68.5	69	77	Moderate	19.0	18
28	Easy	94.1	94	78	Moderate	34.4	33
29	Easy	82.0	80	79	Easy	94.7	91
30	Moderate	68.0	68	80	Moderate	73.7	71
31	Moderate	65.9	67	81	Easy	85.3	83
32	Moderate	22.7	22	82	Moderate	78.2	75
33	Moderate	44.7	45	83	Moderate	58.2	56

Item number	Expert difficulty rating	Item level of difficulty (N = 356)	Percent Correct for Subsample (n = 195)	Item number	Expert difficulty rating	Item level of difficulty (N = 356)	Percent Correct for Subsample (n = 195)
34	Easy	29.2	28	84	Easy	26.6	29
35	Easy	58.3	59	85	Easy	75.8	74
36	Easy	42.0	42	86	Easy	80.6	77
37	Easy	60.7	59	87	Moderate	62.4	59
38	Moderate	59.0	59	88	Moderate	36.6	35
39	Easy	63.0	63	89	Moderate	63.6	61
40	Easy	59.9	57	90	Easy	65.6	63
41	Moderate	64.8	65	91	Moderate	30.1	27
42	Easy	38.4	37	92	Moderate	48.8	43
43	Moderate	41.1	39	93	Moderate	46.6	42
44	Moderate	13.0	13	94	Moderate	48.0	44
45	Easy	71.1	70	95	Moderate	46.6	43
46	Easy	24.7	25	96	Easy	81.1	75
47	Moderate	45.5	44	97	Easy	79.5	75
48	Moderate	84.6	84	98	Easy	17.8	17
49	Moderate	52.0	50	99	Moderate	35.4	31
50	Moderate	50.0	48	100	Moderate	3.8	74