

SYSTEMATIC PROGRAM EVALUATION OF ONLINE NURSING EDUCATION
AT THE MASTER'S DEGREE LEVEL: CURRENT PRACTICES

by

EVA MAE HORNE

(Under the Direction of Lorilee R. Sandmann)

ABSTRACT

Online education is a relatively recent innovation in nursing education and warrants being evaluated for its effectiveness and impact on teaching and learning. The purpose of this study was to investigate current practice and use of systematic program evaluation of master's degree level online education in schools of nursing. An 84 item questionnaire was developed to measure: (1) to what extent are schools of nursing systematically evaluating their online education activities at the master's degree level; (2) what are the sources of evaluation data; (3) what are the areas of evaluation; (4) to what extent are the evaluation results utilized in schools of nursing; and (5) to what extent do institutional and program characteristics affect evaluation. The population sample included 383 schools of nursing with master's degree level programs accredited by the National League for Nursing Accrediting Commission and the Commission on Colleges in Nursing Education. One hundred seven (107/31%) schools of nursing responded to the Internet or mail and paper survey.

Using frequencies, means, standard deviations, and simple bivariate analysis, data analysis showed the most common source of evaluation data is from students ($M = 2.80$).

Respondents reported evaluation practices focused on process most frequently than any other area ($M = 2.63$). Frequency results indicated that all utilization activities were being done to some extent of agreement, that is, no respondents reported zero for any response. Finally, more mature master's degree level nursing programs are more apt to seek a variety of sources of evaluation data.

Four conclusions are drawn from this study. First, systematic program evaluation appears to be a pervasive practice in schools of nursing. Second, there is a discrepancy between espoused utilization of evaluation data and reported practices related to the predominant source of evaluation data (students) and primary focus of evaluation (process). Third, two major stakeholder groups, employers who hire program graduates and staff members who help implement the program, are not included in program evaluation practices and utilization activities. Finally, use of two program evaluation approaches to frame and conceptualize the study resulted in a comprehensive and coherent measurement of evaluation practice and utilization.

INDEX WORDS: Evaluating online education, web-enhanced nursing education, web-based nursing education, utilization-focused evaluation, CIPP program evaluation model

SYSTEMATIC PROGRAM EVALUATION OF ONLINE NURSING EDUCATION
AT THE MASTER'S DEGREE LEVEL: CURRENT PRACTICES

by

EVA MAE HORNE

B.S.N., The Medical College of Georgia, 1978

M.N., Emory University, 1994

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial
Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2013

© 2013

Eva Mae Horne

All Rights Reserved

SYSTEMATIC PROGRAM EVALUATION IN NURSING EDUCATION AT THE
MASTER'S DEGREE LEVEL: CURRENT PRACTICES

by

EVA MAE HORNE

Major Professor: Lorilee R. Sandmann

Committee: Thomas Valentine
Janette Hill
Marsha Davis

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
May 2013

DEDICATION

To

Willie B. Clark

My Granddaddy

For whom literacy was a challenge

He always wanted more for his grandchildren

and

William Henry Horne and Ruby Nell Horne

My parents

For your consistent encouragement

For your continuous support

ACKNOWLEDGEMENTS

I could not have accomplished this journey without much support, encouragement, patience, and prayers. First words of appreciation go to my committee. I will always be grateful for your willingness to work with me. I extend many thanks to my major professor, Dr. Lorilee Sandmann. I am amazed by your knowledge, but particularly appreciate your encouragement, patience, and graciousness throughout this journey. To Dr. Thomas Valentine, thank you for your diligence in getting me from one step to the next in instrument development without too much angst. Thanks to Dr. Janette Hill and Dr. Marsha Davis for your guidance and always having words of encouragement.

I want to thank my parents, siblings, in-laws, nieces, and nephews for praying for me and being a continuous source of support and presence. I want to extend a special thanks to my sister, Terri Yolanda Jones, for her steadfastness in helping me put together survey documents for mail-out during a particularly low point on this journey.

I want to thank my church family, my pastor, and his wife for praying for me and their faith in calling those things that aren't as though they were. Finally, thanks to my friends and co-workers who were always there with a smile and offers of help.

Above all I want to give special thanks and acknowledgement to my Heavenly Father for this opportunity. Without His grace, help, and hope to sustain me I would not have been able to complete this work.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	v
LIST OF TABLES	ix
LIST OF FIGURES	xi
CHAPTER	
1 INTRODUCTION	1
Background of the Problem	1
Statement of the Problem.....	8
Purpose of the Study	10
Significance of the Study	11
Terminology.....	12
2 REVIEW OF THE LITERATURE	14
Method for the Review	14
Overview of Online Education Program Evaluation Research.....	16
Literature Review Conclusions.....	30
Ensuring Quality in Program Evaluation.....	32
Conceptual Models for this Research Study	37
Summary	51
3 METHODOLOGY	53
Conceptual Framework.....	53

Instrument Development.....	59
Study Sample	76
Data Collection	77
Data Preparation.....	90
Data Analysis	94
Limitations	96
4 FINDINGS.....	99
Findings Related to Research Question 1.....	99
Findings Related to Research Question 2.....	105
Findings Related to Research Question 3.....	106
Findings Related to Research Question 4.....	107
Findings Related to Research Question 5.....	112
5 SUMMARY, CONCLUSIONS, AND IMPLICATIONS.....	115
Summary of the Study.....	115
Conclusions and Discussion.....	120
Implications for Practice.....	128
Implications for Future Research.....	132
Summary.....	134
REFERENCES	136
APPENDICES	148
A MAIN STUDY SURVEY INSTRUMENT.....	149
B MAIN STUDY SUPPORTING DOCUMENTS.....	166
C INITIAL ITEM POOL BY SOURCE AND AREA.....	176

D	INITIAL PROTOTYPE ITEMS	183
E	REDUCED ITEM POOL FOR EXPERT REVIEW	190
F	EXPERT REVIEW CRITIQUE SUMMARY	198
G	PROTOTYPE INSTRUMENT FOR PROSPECTUS	201
H	PILOT STUDY SURVEY WITH SUPPORTING DOCUMENTS.....	209
I	PILOT DISTRIBUTION AND RELIABILITY OF KEY MEASURES.....	224
J	RESPONDENT ADMINISTRATIVE JOB TITLES.....	226
K	FREQUENCY TABLE – EVALAUTING PRACTICE ITEMS.....	230

LIST OF TABLES

	Page
Table 2.1: Summary of Literature Using Frameworks and Models	21
Table 2.2: Comparison and Adaptation of the CIPP and Systems Approach Models.....	45
Table 3.1: Evaluating Practice Constructs by Area and Source	56
Table 3.2: Evaluating Use Construct	58
Table 3.3: Steps to Instrument Development.....	61
Table 3.4: Response Scale to Measure Evaluating Practice	69
Table 3.5: Response Scale to Measure Evaluating Utilization	69
Table 3.6: Approach to Data Collection for Pilot Study.....	73
Table 3.7: Summary of Participation and Response Rate for Pilot	74
Table 3.8: Approach to Data Collection for Main Study.....	78
Table 3.9: Summary of Responses by Method of Data Collection.....	85
Table 3.10: Participation and Response Rate Calculations	86
Table 3.11: Personal Characteristics of Study Respondents.....	88
Table 3.12: Institutional Characteristics of Study Respondents.....	89
Table 3.13: Distribution and Reliability of Scales and Sub-Scales.....	92
Table 3.14: Sub-Scales for Key Study Sub-Scales.....	93
Table 3.15: Types of Schools of Nursing in the U.S.....	98
Table 4.1: Top Ten Rank Ordered Means for Evaluating Practice.....	102
Table 4.2: Bottom Ten Rank Ordered Means for Evaluating Practice.....	104

Table 4.3: Source Sub-Scales Mean Scores, SD, Mean of Means.....	105
Table 4.4: Rank Ordered Means for Sub-Scales by Source.....	106
Table 4.5: Area Sub-Scales Mean Scores, SD, Mean of Means.....	107
Table 4.6: Rank Ordered Means for Sub-Scales by Area.....	108
Table 4.7: Frequency Table- Evaluating Use Practices.....	110

LIST OF FIGURES

	Page
Figure 3.1: Conceptual Framework	60

CHAPTER 1

INTRODUCTION

The Internet and online learning are now commonly cited as disrupting higher education as it has been known (Christensen & Eyring, 2011). This is evidenced in the transformation of adult distance education activities. It has certainly become an innovative tool utilized in nursing education programs over the past decade. Internet-based education has become standard practice for delivering training, educational courses, and even entire educational programs to students enrolled in higher education institutions across the United States and globally. Institutions of higher education have embraced the use of online instruction as a strategy for meeting the demands of students who prefer or find it more convenient to take classes in a web environment (Bangert, 2006; U.S. Department of Education, National Center for Education Statistics [NCES], 2008). However, with the rush to stay competitive in online adult education universities, colleges, and particularly schools of nursing are neglecting to systematically evaluate these programs. Online adult education activities have outpaced online adult education program evaluation.

Background of the Problem

Use of the Internet in Adult Education

Distance education has been a mode for adult education delivery since the early 1800s. Previous methods of delivery included correspondence courses via the postal system, radio broadcasting, television, computers, teleconferencing, interactive and

compressed video, and direct-beamed microwave signals (Thurmond, 2002). In 1953, the University of Houston became the first postsecondary institution in the U.S. to offer college courses for academic credit via electronic media to students not enrolled on campus. Over the next fifty years many more colleges and universities would also offer distance-based education courses via some electronic means (NCES, 2008)

During the 1990s, after the creation of the World Wide Web (WWW), the establishment of distance education programs greatly accelerated (Jackson, 2000). With the creation of the WWW, the use of the Internet as a medium, transformed distance education, particularly adult education, and attracted the attention of students, educators, and other stakeholders (Jackson, 2000). Where traditionally institutions had required students enrolled in courses to attend campus for content delivery in a face-to-face or classroom setting, the Internet greatly enhanced the delivery of online educational programs away from institutions and the classroom setting. For purposes of this research study online education is defined as a form of distance education using the Internet and will be used in reference to web-enhanced and web-based education.

According to the NCES (2008) during the 2006-2007 academic year, 66% of the 4,160 two-year and four-year Title IV degree-granting postsecondary institutions in the nation offered some form of online instruction. The most common factors cited by postsecondary institutions as affecting distance education decisions to a major extent were meeting student demand for flexible schedules (68%), providing access (67%), making more courses available (46%), and increasing student enrollment (45%) (NCES, 2008). A follow-up national survey of senior campus officials responsible for managing online and distance education programs conducted by the Western Cooperative for

Educational Telecommunications (WCET) and The Campus Computing suggested economics, experience, infrastructure, and evidence are driving the growth of online education (Green, 2009). This model of education delivery will continue to see significant growth as adults experience how an online degree can positively affect their careers, allow them time to carry out family and work responsibilities, and obtain a degree that merits the same respect as those obtained traditionally from “brick and mortar” institutions.

Use of the Internet in Schools of Nursing

Internet-based education has also transformed nursing education. Distance education in nursing is newer than in other disciplines. The first online or web-based nursing clinical offering was a RN-BSN program started by a major northeastern school of nursing in 1998. It began with traditional courses and became fully online by 1999 (American Association of Colleges of Nursing [AACN], 1999). Since this time, schools of nursing have embraced the use of the web or Internet-based technology to deliver online instruction in order to increase opportunities for nursing students to access undergraduate and graduate degrees. This is occurring predominantly at the graduate versus the undergraduate level.

Ali, Hodson-Carlton, and Ryan (2002) stated “distance education for nurses, most generally in the form of online education, has grown at a phenomenal rate in the last decade and may help to meet the growing demand for nurses” (p. 111). Nauman (2006) predicted that by the year 2015 there will be fewer nurses than are needed, thereby reinforcing the need for new and better modes of teaching and learning that can enhance student enrollment. Online education in nursing is also significant for nurses from

medically underserved rural and urban areas. Nurses who leave these areas for advanced education frequently do not return. Consequently, distance education increases the likelihood nurses will remain in the local area (Leasure, Davis, & Thievon, 2000).

Nauman noted that along with the current shortage of registered nurses in the U.S. there is also a shortage of nurse faculty to teach qualified applicants being denied access to programs. Nauman stated that “the nurse workforce shortage, coupled with a growing need for faculty, calls for redesign, restructuring, and recognition that the flexibility and availability of technology offer nursing education enormous opportunities for innovation” (p. 12).

The Issues

As great as this pedagogical innovation seems, there have been issues with online education in general and particularly in the nursing profession. First, systematic program evaluation of online education is uncertain. There is a paucity of documentation in the literature that systematic program evaluation of online education is being done routinely. Although online course offerings provide greater access to education there has been concern among adult educators and other stakeholders that online education courses have outpaced the evaluation of the quality of these courses, particularly at the program level (Avery, Cohen, & Walker, 2008; Bangert, 2006; Billings, 2000). Quality assurance procedures for design and delivery of web-based courses have often been ignored as universities and colleges rush to offer online education programs that will allow them to maintain and increase enrollments (Bangert & Easterby, 2008). Specifically, nursing educators noted that online education of adults has been adopted rapidly, but less

attention has been given to examining the quality and efficacy of this innovation compared to traditional educational practices (Leners, Wilson, & Sitzman, 2007).

Second, when online nursing programs are regularly evaluated we do not understand how the data gathered is utilized. Utilization of evaluation data can assist stakeholders as they make decisions regarding online program development, improvement, or continuation. But there is a dearth of research literature on systematic program evaluation data utilization of online nursing education activities.

Why Do Program Evaluation?

Program evaluation is a component of every major program planning theory or framework used in adult education (Caffarella, 2002; Cervero & Wilson, 2006; Forester, 1989; Tyler, 1949). Therefore, program evaluation is an important component of successful online pedagogy. Fitzpatrick, Sanders, and Worthen (2011) defined program as “a complex of people, organization, management, and resources that collectively make up a continuing endeavor to reach some particular educational, social, or commercial goal” (p. 54). Evaluation is defined as “the identification, clarification, and application of defensible criteria to determine an evaluation object’s value (worth or merit) in relation to those criteria” (Fitzpatrick et al., 2011). Therefore, program evaluation is the systematic collection of information about the activities, characteristics, and outcomes of programs to make judgments about the program, improve or further program effectiveness, increase understanding, and/or inform decisions about future programming (Patton, 2008).

There are two basic types of evaluation. Formative evaluation focuses on program improvement and often provides information to judge the merit or worth of a part of a program. Summative evaluation focuses on providing information for decision making;

making judgments about program adoption, continuation, or expansion (Fitzpatrick et al., 2011). While the program evaluation process utilized by schools of nursing reveals the success or failure of programs to meet predetermined goals and standards, it does not always focus on program improvement (Chapman, 2006). This finding demonstrates a lack of comprehension and quality in the evaluation process both formative and summative.

Purpose of Program Evaluation in Online Education

The purpose of program evaluation of educational initiatives that include online activities will vary with the informational needs of each stakeholder or stakeholder group. These may include (a) justification of investment, (b) measuring progress toward program objectives, (c) measuring quality and/or effectiveness, (d) providing a basis for improvement, and (e) informing decision making (Thompson & Irele, 2007). Furthermore, program evaluation of online nursing education is needed to answer questions of interest to stakeholders, i.e., students, faculty, and administrators. For example, the constant upgrades in educational technology have increased stakeholder expectations and students are becoming increasingly knowledgeable and discerning consumers of adult education (Ruhe & Zumbo, 2009). Finally, formal review process evaluations utilized in higher education focus more on outcomes and not necessarily program improvement. And though all academic programs have similar purposes for their evaluation, there are certain characteristics of online programs that are not addressed in standardized review processes (Chapman, 2006).

Up until a few decades ago, the impact of distance education programs on institutions was small and somewhat marginalized. Resources devoted to distance

education evaluation were also small and the results were of interest only to see that distance education did not detract from the reputation given to more traditional educational programs. Past evaluation studies focused on comparatively demonstrating that distance education programs were as good as face-to-face institution programs or the “real thing” (Thompson & Irele, 2007). Only more recently has research in evaluation of online education branched out to include other concerns such as systematic program evaluation.

Subsequently, there is a scarcity of literature that reported systematic program evaluation using an evaluation model and/or theoretical approach (Ali et al., 2002; Avery et al., 2008; Lindsay, Jeffrey, & Singh, 2009; Mills, 2007; Singh, Jeffrey, & Lindsay, 2008). Most of the literature in online evaluation in nursing education of teaching and learning has been aimed at the individual course level and program level evaluation was scarce (Ali et al., 2002; Avery et al., 2008). Online nursing graduate education research and observations has focused on the success, failure, and use of a variety of technology software (Little, Passmore, & Schullo, 2006), comparison between online and traditional courses (Woo & Kimmick, 2000), teaching effectiveness (Billings, 2000; Billings, Connors, & Skiba, 2001; Seiler & Billings, 2004), and course design and curriculum (Avery et al., 2008). These studies and the different models and frameworks used in systematic program evaluation will be discussed in more detail in chapter 2.

With the power and reach of the Internet, distance education and online learning has moved from the margins to the mainstream. It is only through evaluation and evaluation research that educational institutions can determine whether particular online education programs and activities can continue to contribute positively to education

programming and educational quality (Thompson & Irele, 2007). Furthermore, evaluation research contributes to the building of theoretical frameworks and models.

Statement of the Problem

Online education is a relatively recent innovation in nursing education and warrants being evaluated for its effectiveness and impact on teaching and learning. Davidson (as cited in Ruhe & Zumbo, 2009, p. 7) stated that “every time we try something new, it is important to consider its value.” But a search of the literature shows scarcity of program evaluation research of the online component of nursing education. Anecdotal evidence suggests that higher education institutions such as schools of nursing are familiar with program evaluation of their more traditional educational offerings in order to ensure quality and to maintain accreditation. Standards, key elements, and competencies must be met for nursing programs to continue a positive accreditation status from their state, region, and either of the two professional accreditation agencies; the Commission on Collegiate Nursing Education (CCNE) or the National League for Nursing Accrediting Commission (NLNAC) (Story et al., 2010). Accreditation is extremely important in addressing whether programs are meeting established quality standards, and whether they should be certified and approved. However, accreditation evaluation studies fall short of being candid and the role of stakeholders is not often considered. The Internet has rapidly changed education’s context (Thompson & Irele, 2007). Thompson and Irele (2007) stated that “programs, institutions, and societies must make significant decisions as to how they wish to influence or shape these changes, and/or be shaped by them” (p. 419).

With the introduction of the Internet into nursing education, two concerns emerge. First, systematic program evaluation of online nursing education is questionable. Although online course offerings provide greater access to education, there has been increasing concern among adult educators that the prevalence of online education courses has outpaced the evaluation of the quality of these courses, particularly at the program level (Avery et al., 2008; Bangert, 2006; Billings, 2000). Little is known about the outcomes and how web technology and its learning tools contribute to quality teaching, learning, access, and cost benefit in online nursing education. The research that has been done in evaluation of online nursing education has mainly centered on quality components of individual courses and little has focused on systematic evaluation at the program level. Second, when and where online nursing education is regularly evaluated at the program level it is unclear how these results are used to make informed decisions about program improvement, revision, and/or continuation.

Therefore, distance education activities should be part of systematic program evaluation plans to show delineation in education delivery. Evaluation standards and criteria from the various accrediting or reviewing organizations, such as the Commission on Collegiate Nursing Education or the National League for Nursing Accrediting Commission, and specialty and regional organizations can be studied and used as guidelines to create a single integrated, comprehensive evaluation plan that includes traditional, web-based, and web-enhanced course offerings (Suhayda & Miller, 2006). But in order to provide valuable information about flaws in course design, in implementation of technical systems, and to generate specific directions for program improvement systematic evaluation is needed (Ruhe & Zumbo, 2008).

For purposes of this study, two program evaluation models that have been empirically tested and validated in the literature were adapted for use. These two theoretical frameworks informed both study concerns of evaluation practice and evaluation use of coordinated and systematic program evaluation in schools of nursing. The first is Rovai's (2003) systematic approach to program evaluation of online education. Rovai's (2003) framework is adapted from Stufflebeam's (1971, 2001) four types of evaluations; context, input, process, and product or better known as CIPP approach to evaluation. Where systematic program evaluation was performed in the literature, the CIPP model, or a modification thereof, was used most frequently. A detailed discussion of the CIPP model is presented in chapter 2.

The second approach is utilization-focused evaluation (UFE). The primary focus of the UFE framework is intended use or utilization of evaluations by intended users. UFE is defined as "the systematic collection of information about the activities, characteristics, and results of programs to make judgments about the programs, improve or further develop program effectiveness, inform decisions about future programming, and /or increase understanding" (Patton, 2008, p. 39). A more detailed discussion of UFE framework, and the rationale behind the decision for the term use versus utilization for this study, is presented in chapter 2.

Purpose of the Study

The purpose of this quantitative study is to investigate to what extent schools of nursing are currently practicing systematic program evaluation of online education and how are they using the results. The research questions are:

1. To what extent are schools of nursing systematically evaluating their online education activities at the master's degree program level?
2. What are the sources of evaluation data?
3. What are the areas of evaluation?
4. To what extent are evaluation results utilized by schools of nursing?
5. To what extent do institutional and program characteristics affect evaluation (a) data source; (b) area; and (c) utilization in schools of nursing online master's degree level programs?

Significance of the Study

Examining the process and utilization of systematic evaluation of online nursing education has practical, theoretical, and research significance. Practically, this study will attempt to contribute to the knowledge base by investigating to what extent is program evaluation of online nursing education taking place in schools of nursing. The information provided from studies such as this, to the extent of sampling adequacy, will be a measure of state of the art online program evaluation. Schools of nursing who wish to establish or refine their evaluation system will be able to see what is being done at a majority of schools and what is being done at particularly excellent schools. Likewise, evaluation experts who might interact with the nursing programs will have a sense of what matters.

Furthermore, I hope to raise awareness among nurse educators of the need for systematic program evaluation from a fiscal accountability standpoint. Higher education, including schools of nursing, is being held accountable for their cost, efficacy, and educational value. Administrators and stakeholders want to see a return on their investment as well as technological cost benefits. Schools of nursing are accountable to

stakeholders for their performance and to ensure identification and implementation of ongoing process improvement efforts (Lindsay et al., 2009). Evaluation research can help point out best practices in assessing cost benefit and cost effectiveness of online nursing education.

Theoretically this research study offers an empirical test of two theoretical evaluation models of program evaluation in online nursing education. These theoretical models are UFE and an adapted version of CIPP for online education evaluation. The components, premises, and guidelines of the models will hopefully reveal their potential usefulness and value in program planning of online education. Furthermore, the use of these two models together will advance and test a new combination model. The survey instrument developed using this combo model can be used by schools of nursing to assess the quality of their evaluation process of online education.

Research information from systematic program evaluation data contributes to evidence-based nursing education and to “building a science of nursing education” (Lindsay et al., 2009, p. 181). Program evaluation can inform curriculum, pedagogy, and outcomes of online nursing education. Research in program evaluation at the various levels of online higher education contributes not only to evidence-based education, but theory development and best practices in online education and evaluation.

Terminology

Terms used in this study, with definitions, are listed below:

- Master’s degree level education in nursing – a registered nurse that has usually completed baccalaureate level training and furthers learning in analytical skills, integrating theory with practice, and increases knowledge

and competence in a nursing specialty, i.e., clinically, administratively, educationally (AACN, 1999).

- Online education – a form of distance education; education delivered by web-based or Internet-based technologies; terms often used interchangeably in the literature.
- Web-based education – no face-to-face meetings with the instructor and the teaching and learning experience is totally online (O’Neil, Fisher, & Newbold, 2004).
- Web-enhanced education – teaching and learning events that combine aspects of online and face-to-face education (O’Neil et al., 2004).

CHAPTER 2

REVIEW OF THE LITERATURE

The purpose of this study is to investigate to what extent schools of nursing are currently practicing systematic program evaluation of online education and how they use the results. This chapter provides a review of the literature relating to the proposed research topic. Specifically, the focus is on five main areas. First, the criteria used in selecting the literature for review. Second, the initial part of the literature review will highlight in general what has been done in program evaluation of online nursing education at the graduate level. Third, the review will take a more narrow focus on what is being done in systematic program evaluation research pertaining to online graduate nursing programs. Fourth, a summary will be made of the literature findings. Finally, a discussion of different models of systematic program evaluation; based on a review of these models from the literature present rationale for why the selected models were chosen and how they informed and guided this research project.

Method for the Review

This review is concerned with systematic program evaluation of online graduate nursing programs at the master's and post master's degree level. Graduate level programs were chosen because more distance education nursing courses are offered at this level (NCES, 2008; Potempa et al., 2001) and schools of nursing would have had more opportunity for data collection in terms of longevity and experience for program

evaluation. The online component included both web-based and web-enhanced technology delivered instruction models of teaching learning using the Internet or web.

The primary criterion for inclusion in this review was that the evaluations had to be based on research and/or theory development in systematic program evaluation of online nursing education. Dissertations, research presentations, and technical articles were also included. The following data bases were searched for this literature review: Web of Knowledge, CINAHL, PubMed, Academic Search Complete, JSTOR, Education Research Complete, and ERIC. Teaching strategies without a systematic approach to evaluation, practice oriented, and anecdotal literature, were omitted.

The terms used in the search began with “program evaluation” (summative and/or formative) which garnered evaluations of entire programs or specific outcomes and constructs measured for formative evaluation and program improvement. After the initial search protocol, the additional phrase “graduate nursing programs” was used to focus and narrow the results. Finally, the following key words, some defined in chapter 1 of this document, were also employed in the search: distance education/ learning, e-learning, web-based education, web-enhanced education, Internet-based education/learning, and online education/ learning, since these terms are used interchangeably in the literature. Since the Internet was not widely in use in the public domain until 1989, only research articles between 1989 to present were included (Jackson, 2000).

Much of the evaluation literature on online nursing education has been aimed at the individual course level and do not include program evaluation in its entirety. Course level research was included to show what was being evaluated. Only five articles were

found that claimed evaluation at the program level using an evaluation approach or theoretical model.

Overview of Online Education Program Evaluation Research

In general, a great deal of online evaluation research not specific to nursing has explored the effectiveness of online tools (Spatariu, Hartley, & Bendixen, 2004), assessed aspects of courses (Roblyer & Ekhaml, 2000), addressed evaluating effective online instruction (Graham, Cagiltay, Lim, Craner & Duffy, 2001; Wentling & Johnson, 1999), assessed the value of online courses in specific fields of study (Carmichael, 2001), and compared face-to-face learning (Hoben, Neu, & Castle, 2002; Russell, 1999).

Russell's (1999) "no significant difference phenomenon" is perhaps the most quoted in the literature. Russell (1999) reviewed 355 studies on distance education produced from 1928 to 1998. Some of the early studies examined correspondence courses, but most compared instruction over videotape with interactive video and satellite with on campus in person courses. Consistently, based on statistical test, "no significant difference" between the comparison groups was found. But, only 40 of the 355 studies specifically included web-based instruction, and the study was completed prior to the proliferation of courses using the web, especially in nursing. In spite of this, Russell's study helped to validate and legitimize distance education. The study results continue to provide a scholarly boost to online education.

The nursing literature also suggests there is no significant difference in learning outcomes between classroom and online students (Leasure, Davis, & Thievon, 2000; O'Neil & Fisher, 2008; Woo & Kimmick, 2000). Woo and Kimmick (2000) compared an online graduate nursing research course with a traditional face-to-face taught by the same

instructor. A total of 97 students (44 Internet; 53 traditional) participated. There was no significant difference in test scores and overall course satisfaction ($p > .05$). But the Internet students rated the lectures as more stimulating than those in the traditional course. One limitation was that the comparison was limited to the didactic content or lecture component of the course.

Other evaluation research topics identified in the literature were included in six broad categories: (a) the success, failure, and use of a variety of technology software (Little et al., 2006), (b) comparison between online and traditional courses (Beta-Jones & Avery, 2004; Cragg, Dunning, & Ellis, 2008; Woo & Kimmick, 2000), (c) teaching and learning effectiveness and strategies (Daroszewski, Kinser, & Lloyd, 2004; Huckstadt & Hayes, 2005); (d) teaching and learning outcomes (Edwards, 2005)), and (e) measurement of student perceptions and experiences of online nursing education (Billings, Skiba, & Connors, 2005; Fearing & Riley, 2005; Seiler & Billings, 2004; Wills & Stommel, 2002).

Overview of Online Education Program Evaluation Research in Nursing

Little et al. (2006) looked at the success of technology software measured by student satisfaction. The researchers undertook a pilot research project to incorporate synchronous classroom software into an ongoing online program. A convenience sample of 33 students enrolled in a RN-BS community/public health nursing clinical course participated. The participants were all female and ranged in age from 20 to 54. Demographic information and online educational experiences was evaluated using a researcher-developed web-based questionnaire. Students' reactions to the synchronous sessions and the instructor's teaching approaches were measure on a researcher-

developed, post-participation, web-based questionnaire. Twenty-six students responded to the final survey for a return rate of 86%.

Little et al. (2006) reported that the majority rated the synchronous technology as either posing no problem or minor problems when getting familiar with the software. The university also administered the standard end-of-semester course evaluation to assess student satisfaction which was an 11-item instrument using a 5-point Likert scale. Results showed ($n = 14$) students who participated in the online sessions reported significantly higher levels of satisfaction with the course than students who did not use the technology ($t = 2.183$, $df = 26$, $p = .038$). Based on positive responses from the pilot, the same study was repeated in another semester with similar results noted. The most negative comments in the latter study were related to technology issues such as initializing the setup wizard. Also, some students felt that the online sessions were too long being on average one hour and forty five minutes.

Billings et al. (2001, 2005) developed an instrument that measured students' perceptions to determine the best practices for web-based graduate and undergraduate nursing courses. Billing et al. (2001, 2005) developed the survey instrument on the basis of benchmarking. The first study (2001) was a pilot test. The final survey instrument, *Evaluating Educational Uses of the Web in Nursing* (EEUWIN) is a 57-item, Likert-type questionnaire with two open-ended or unstructured questions. Content validity was well established from a consensus panel of expert nurses, and reliability yielded a Cronbach's alpha of 0.85 to 0.94. Billings et al. (2001) selected the following variables for benchmarking:

- outcomes to include access, convenience, connectedness, preparation for real work, proficiency with technology use, socialization to the profession, and satisfaction;
- educational practices to include active learning prompt feedback, time on task, collaboration and interaction among peers, and student-faculty interaction; and
- uses of technology to include technology infrastructure and use of technology that promotes productive use of time.

Data analysis was by descriptive and inferential statistics. The samples for both studies were obtained from different schools of nursing (219 and 558) increased generalizability.

Billings et al. (2001) found from the pilot study that convenience, accessibility, and reliable technology are important for student satisfaction. Feedback and active learning are essential for success in online learning. Interaction could decrease and computer proficiency tended to improve in web-based courses. Billings et al. (2005) essentially concurred with the findings of their pilot study. However, graduate nursing students perceived spending more time on task in the web-based course than in a traditional campus course. Both graduate and undergraduate nursing students were willing to assume responsibility for their own learning, were satisfied with the web-based courses, and felt socialized to the profession. There were no obvious limitations to these studies.

Seiler and Billings (2004) obtained data regarding what worked well or needed improvement in courses offered fully online by once again using the EEUWIN

instrument. In this study the researchers analyzed the responses to the two open-ended questions using qualitative description and content analysis. The two questions asked the best thing about the course and how could the course be improved. Four hundred and fifty-eight students responded to the survey. Participants ranged in age from 20 to over 50 and most (97%) were female. One-third were enrolled in masters programs. Seiler and Billings (2004) found that technology should be reliable and promote productive use of time. Educational practices needed to recognize and highlight active participation, sufficient time, meaningful and timely feedback, positive interaction, diverse ways of learning, and expectations of students. Outcomes highlighted learning styles, access, convenience, connectedness, professionalism, satisfaction, and orientation to technology. The only limitations to the study was that of the sample, though large, was a convenience sample and that the responses were subjective (a characteristic of a qualitative approach) and may not have reflected the opinion of entire groups and thereby decreased generalizability.

Although these studies in online education have informed and made significant contribution to nursing and higher education, they did not indicate how the results fit into a larger program evaluation plan, how the research results could be used for program improvement, or how the evaluation was utilized by stakeholders in decision making. They did not identify an evaluation approach. Nor was this their focus.

Review of Research Using Theoretical Models of Evaluation

Five articles potentially met criteria for systematic program evaluation. See Table 2.1 for a summary of the five evaluation frameworks and models used in online education programs. Two of the articles (Lindsay et al., 2009; Singh et al., 2008) used a literature-

based evaluation and will be discussed first. These two articles were also more inclusive of components and guidelines used to evaluate programs found later in this chapter. The remaining three articles (Ali et al., 2002; Avery et al., 2008; Mills, 2007) are based on distance learning (DL) evaluation theory and will be presented alphabetically.

Table 2.1

Summary of Literature Using Evaluation Frameworks and Models

Frameworks and Models	Literature
CIPP model	<i>Paradox of a Graduate Human Science Curriculum Experienced Online: A Faculty Perspective.</i> Lindsay et al. (2009)
Participatory evaluation approach	<i>Isolated Learning for Caring Professionals: Advantages and Challenges.</i> Singh et al.
UFE (used for data analysis); Chickering and Ehrmann's education principles adapted to technology; Billings's framework for evaluating quality of online education	<i>Evaluation of an Online Graduate Nursing Curriculum: Examining Standards of Quality.</i> Avery et al. (2008)
Chickering and Ehrmann's education principles adapted to technology	<i>Web-based Professional Advanced Practice Nursing: A consumer Guide for Program Selection.</i> Ali et al. (2002)
Framework from the EDUCAUSE Center for Applied Research	<i>Evaluation of Online and On-site Options for Master's Degree and Post-master's Certificate Programs.</i> Mills (2007)

Lindsay et al. (2009) evaluated the experiences of nine nursing faculty who developed and taught in a web-based masters of science nursing program (MSN). Stakeholders were identified as students, faculty, administrative partners, and employers. Lindsay et al. (2009) did note that all stakeholders participated in the evaluation but this article focused on faculty experience. The purpose of the article was to evaluate graduate nursing faculty experience with developing and teaching in the MSN online program. The purpose of the evaluation was summative, formative, and for accountability in order to assess the implementation of the program to detect or predict defects in program design. Lindsay et al. (2009) used portions of Stufflebeam's CIPP program evaluation model as an approach to evaluation. Stufflebeam's CIPP model is considered an example of a management approach to program evaluation (Fitzpatrick et al., 2011). From this model the researchers used accountability indicators for process related to program implementation and product related to program outcomes to focus the evaluation. The stakeholders identified four evaluation questions that probed the faculty's experience in the implementation of this online graduate program.

Journaling and faculty focus groups comprised the qualitative data collection methods. Faculty focus group meeting transcripts and journal entries were analyzed to identify achievements and challenges related to online education. Several achievements were consistent throughout the analysis. First was being able to activate the master's degree program six weeks after obtaining approval to do so. Second was "establishing congruence between learning outcomes and content with course processes given the online mode of program delivery" (p. 183). Third was satisfaction with managing the workload requirement of transitioning to online methods of teaching and learning.

Finally, the enjoyment of bringing faculty research and experience to the preparation process was highlighted. The challenges included time commitment in course development, pedagogical challenges, and the time commitment required for course management and student engagement. Utilization of the evaluation resulted in plans to include employers as stakeholders; review of the curriculum to possibly add a thesis option; and to increase faculty development in technology. The Lindsay et al. (2009) program evaluation included all components of the process. However, the sample size was a convenience sample and small in number.

Singh et al. (2008) evaluated student perspectives and experiences with a web-based MSN degree in nursing. The stakeholders were identified as students, faculty, administrative partners, and employers. This arm of the evaluation was to document students' experiences with web-based education from the first day of orientation to program graduation from 2005-2007. The purpose of the study was to document graduate student experiences of an online MSN program from orientation to graduation and summative and formative evaluation for accountability of the program. This comprehensive evaluation focused on program process and outcomes from learning. A total of nine questions were identified and focused on process and outcomes of teaching and learning from a student's perspective. The researchers used participatory evaluation approach with a focus on utilization of results. Singh et al. (2008) supported the idea that processes embedded in programs are reflective of the participant's lives. It is also based on the assumption that inclusion and participation in the evaluation process ensures relevancy, validity, fairness, and the utility of the evaluation (Joint Committee on Standards for Education Evaluation, 1994; Singh et al., 2008).

The sample consisted of the first cohort of students that entered and completed the program (n = 11) over a 2 year period attending full or part time. Data collection methods included three questionnaires, journaling, focus group, and individual interviews. Qualitative data analysis from journaling led to a list of achievements, challenges, and recommendations. Achievements during the two year period in order of timing in the program included getting a successful start in the program, staying in the course, and supporting and helping peers. Student challenges included time management, balancing the demands of the program and other priorities, and shared methods of assessment of student work. Singh et al. (2008) reported utilization of results in three areas. Formative evaluation resulted in changes to program orientation and making sure that goals in courses online matched the goals of the program. Summative evaluation resulted in two modes of program delivery: web-enhanced and web-based. This addressed the demand for more face-to-face interaction from local students.

Though Singh et al. (2008) offered narrative results for the questionnaires, they did not share quantitative descriptive results. The researchers also did not offer results from the interviews or focus groups. A convenience sample was used and the size was small. However, this study contained all components of the evaluation process.

Ali et al. (2002) used a scenario approach to research how a mid-western nursing school developed, implemented, and continually evaluated web-based learning in a master's program. Stakeholders were not specifically identified but students and faculty seemed to be the focus. The purpose was to evaluate the development and implementation of a web-based master's program over a two year period from 1998-

2000. The program was evaluated from student perspectives assessing student satisfaction with the program.

Construction of the instrument was based on Chickering and Gamson's *Seven Principles of Good Education* restructured for distance education technologies. From these principles, criteria for evaluation were developed and included course content, interaction, participation, critical thinking, faculty preparation, communication skills, and technical skills. Continued internal consistency reliability of the tool during the first year of the program resulted in a nine-item survey tool with a high Cronbach's alpha ($\alpha = .91 - .94$). A 5-point Likert agree/disagree scale was used. In addition to the scale, three open ended questions sought feedback regarding what students liked and disliked about the course and suggestions for improvement. During the second year a total of eleven courses were delivered ($n = 126$) and evaluated. Results were calculated using descriptive statistics. Means for the nine-item scale ranged from 4.43 to 4.47 ($SD = .64 - .77$), with 5 being highest, in which students were satisfied with the experience of participating in web-based courses.

Specific areas of satisfaction were reported with content and the currency of content, critical thinking exercises, interaction between faculty and students, participation among students, and time to complete assignments. Negative responses included feedback not being timely and too many assignments. Findings from the open ended comments revealed a variety of learning styles, positive aspects included content and content currency, and that theory could be applied through case study. The authors did not specifically note what improvement activities or decisions resulted from this study (Ali et al., 2002).

The study by Ali et al. (2002) was program specific and the results were generalizable to web-based courses. The specific purpose was focused on students' perspectives of online courses. The context was identified. Data collection methods and results were appropriate for the study. On the other hand there were several items missing from the evaluation. First, the researchers did not specify specific questions for the evaluation. Second, the evaluation tool was not shared. Finally, utilization was questionable since the researchers did not share specific program improvements or decisions that came from the evaluation results. This could have been attributed to a deficiency in the theoretical framework.

Avery et al. (2008) conducted an evaluation of a web-enhanced 16 course program that was developed for three master's specialty areas of nursing at a large mid-western U.S. university. Evaluation stakeholders were not specifically named but nursing faculty seemed to be the focus. The purpose of the project was to evaluate the quality of the 16 web-enhanced courses. The faculty's shared beliefs on what should be evaluated were well grounded in the literature and resulted in four quality standards being selected; (a) course mechanics, (b) course organization, (c) student support, and (d) communication and interaction in online education.

The researchers employed methods from existing instruments found in the literature to develop an evaluation tool or instrument that would best measure the quality standards. Reliability and validity of the instrument were determined by using a pilot test, which focused on a graduate ethics course that had been taught several times by a single faculty member. The final instrument consisted of 20 items and one comment question. A 5-point Likert scale was used to rate the 20 items; higher numbers represented closer

adherence to the standard. Two faculty members from the school of nursing that had the most experience with online teaching and one educational technology person were used to collect data using the methods of peer review, interviews, and the developed instrument.

To promote utilization of evaluation results, Avery et al. (2008) used Patton's (2008) utilization-focused evaluation (UFE) approach in data analysis. UFE is sometimes sub-categorized as a participant approach to evaluation (Fitzpatrick et al., 2011).

Researchers used descriptive statistics to report results. Qualitative analysis was accomplished by follow up faculty interviews reviewing each item on the tool. The overall mean score for all items was 3.91 on the 1-5 scale, 5 being the highest score. Goals and objectives appropriate to course level was rated the highest at 4.51, with 5 being the highest. That there is a written connection between the course objectives and learning activities was scored lowest at 2.88, with 5 being the highest. Themes derived from qualitative data included support for technology, support for different learning styles, interaction between student-student is critical, the need for course objectives to match learning activities, and the importance of student voice. The researchers reported the findings to the faculty as a whole for quality improvement of online courses and that faculty response was positive to the evaluation.

Avery et al. (2008) were specific about the purpose and goals for the evaluation. The quality standards were developed using appropriate literature. The tool designed by Avery et al. (2008) was very specific to online and blended delivery. Data collection and analysis were consistent with quantitative and qualitative methods and utilization-focused approach to program evaluation. Results were disseminated with utilization in mind.

Decisions regarding program improvement and instrument revision were specific to the program. From this study best practices in quality were developed.

The researchers did not share the instrument or questions used for data collection. This made it difficult to attach exact meaning to the statistical results without knowing the scale that was used or interpretation given to questions. Also, it could be argued that test-retest or calculation of Cronbach's alpha may have been a better determinant of reliability of the instrument. Input from students, alumni, and administrators would have added to the evaluation plan. The researchers noted, however, that broader overall program evaluation was beyond the study at this time.

Mills (2007) conducted a program evaluation using a comparative study approach of an online versus a traditional master's degree and post master's certificated program. All programs were six years old. Stakeholders were identified as students, faculty, and administrators at a mid-western school. The purpose and question of the summative evaluation was to determine whether their distance learning program should be continued aggressively as part of the strategic initiatives for the SON. Mills (2007) hypothesized that "student socio-demographic and admissions data and all student outcome measures and program performance would be comparable between distance learning and on-site students, with two exceptions" (p. 74). The exceptions were increasing student access, which was a goal of the online program and to demonstrate marketability of the post master's program for online delivery. The theoretical evaluation approach was based on a framework from the EDUCAUSE Center for Applied Research. Mills (2007) modified the approach to evaluate three of the constructs; student outcomes, program effectiveness, and organizational effectiveness. Mills did indicate that the other three possible

constructs, institutional transformation, institutional outcomes, and faculty related outcomes, were outside the scope of the program evaluation in terms of resources.

Archival or secondary data were collected from 17 courses and all students within those courses ($N = 270$). Mills examined the program for socio-demographic and student related outcomes (cumulative grade point averages and certification pass rates, etc.), program effectiveness (enrollment, retention, and completion rates, etc.) and organizational effectiveness (cost, etc.). Data was analyzed using both descriptive and qualitative measures. Master's DL students tended to be approximately six years older than on-site students. No differences existed in other demographics. A series of 2-way analysis demonstrated no significant differences ($P = .169$) in cumulative grade point average between on-site and online.

In examining program effectiveness, for the MSN degree there was significant difference ($P = .008$) when enrollment was compared with admissions or the yield rate between the two groups. Retention rates were significantly higher for DL groups than on-site ($P = .038$). Organizational effectiveness data showed that DL carried a higher cost in technology cost and faculty workload cost. But it was noted that enrollment shifted from on-site to DL courses and also that program enrollment also was increased overall. There were no tuition or fee structure differences for on-site or online. Administrative decisions to continue the DL program was based on evaluation data from the success of the program and organizational effectiveness (Mills, 2007).

There were some limitations to this study. Since the study used archival data, there was missing information for data collection and analysis. The sample was a convenience sample. Mills (2007) study addressed major components of the program

evaluation process to include utilization of results. Three areas of the original EDUCAUSE framework were not used; institutional transformation, institutional outcomes, and faculty related outcomes. This decision to omit these three areas limited the scope of the evaluation, and therefore could have potentially limited the completeness of the summative evaluation.

Literature Review Conclusions

A review of seven discipline appropriate databases, using ten key terms showed that there is a paucity of systematic program evaluation research of online master's education in the literature. Lack of a program approach to evaluation of nursing distance education activities continues to be evident in the literature despite over two decades of use. Ali et al. (2002) report evaluation results from as early as 1998. The majority of the literature of online nursing education has been aimed at the individual course level (Ali et al., 2002; Avery et al., 2008).

Types of evaluation, formative and/or summative, were sometimes noted in the review articles. Stakeholders are generally identified as students, faculty, administrators, and alumni. Where all stakeholders were not implicated the researchers did note the primary focus and purpose of the evaluation. In general, students and faculty perspectives seem most important to the evaluators, although Mills (2007) did include program performance and institutional effectiveness.

Two evaluations (Lindsay et al., 2009; Singh et al., 2008) used two different approaches respectively; (a) participatory approach focused on utilization, and (b) Stufflebeam's CIPP management approach focused on decision making. Suhayda and Miller (2006) also described the use of a management approach using Stufflebeam's

CIPP model to frame their SON program evaluation plan which includes traditional, web-enhanced, and web-based delivery. One evaluation (Mills, 2007) used a framework from the EDUCAUSE Center for Applied Research that focused on three of the six measures. Ali et al. (2002) and Avery et al. (2008) used Chickering and Gamson's *Seven Principles of Good Education* as a theoretical approach. This reiterates the fact that an evaluation approach depends on stakeholder needs, purpose of the evaluation, and evaluator expertise.

Data collection methods included questionnaires, focus groups, interviews, journaling, and archival data collection. Data analysis methods are both quantitative and qualitative. Research using multiple approaches to data collection did not always present all results. Utilization of evaluation results were used for program improvement and decision making. Program improvement included process and strategies aimed at student satisfaction and teaching effectiveness within the program. Decisions included continuation of programs, modifications, or additions of online courses within programs.

From the review four different systematic evaluation approaches were identified in evaluation online graduate nursing educational programs. Two were traditional models; CIPP and UFE. Two were distance learning models; the EDUCAUSE framework and Chickering and Gamson's *Seven Principles of Good Education*.

Before moving on to Chapter 3, in the final section of this chapter an overview of systematic program evaluation and the different program evaluation models will be presented; traditional and those developed specifically for online evaluation. Finally, the theoretical models chosen to inform this study will be given in-depth discussion.

Ensuring Quality in Program Evaluation

High quality program evaluations are the concern of the evaluation profession and stakeholders clients. But what constitutes a quality evaluation? The concern for quality evaluations has led to the development of evaluation standards and the push for evaluators to use theoretical models in practice. But by far the development of evaluation standards has been the greatest effort to ensure quality in the field of educational evaluation. During the 1960s evaluators began a discussion in the literature suggesting what constitutes a good or bad evaluation. Checklists of evaluation standards began to be exchanged informally and several evaluation scholars published their own guidelines and criteria. Also, clients were beginning to ask evaluators for more accountability of their reports (Fitzpatrick et al. 2011). Under the leadership of Daniel Stufflebeam, the first formal standards for evaluation were published by the Joint Committee on Standards in 1981 (Joint Committee on Standards for Educational Evaluation, 1994). Updated in 2011 (Yarbrough, Shulha, Hopson, & Caruthers, 2011), these standards are commonly agreed on characteristics of good evaluation practice from those in the field. The most recent revision of these standards call for quality evaluations to have five main features:

- utility – intended to ensure that an evaluation will serve the practical information needs of intended users;
- feasibility – intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal;
- propriety – intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results;

- accuracy – intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated; and,
- accountability – encourage adequate documentation of evaluations and metaevaluative perspective focused on improvement and accountability for evaluation processes and products (Yarbrough et al., 2011).

The use of program evaluation models is another way to ensure comprehension and quality in evaluation practice. Program evaluation models that have been empirically tested and validated can serve as road maps to guide evaluation practice; including important aspects, steps, and strategies (Ruhe & Zumbo, 2009). A definition for evaluation model is a set of beliefs about the “concepts and structure of evaluation work” (Ruhe & Zumbo, 2009, p. 10). An evaluation model can provide steps and guidelines for arriving “at defensible descriptions, judgments, and recommendations” (Madaus & Kellaghan, 2000, p. 20). While in practice some evaluators may not start the evaluation process with theory, using a theoretical model or framework for program evaluation can help to ensure that the evaluation plan is both comprehensive and coherent (Bevil, 1991). It is also general consensus among professional evaluators, scholars in the field, and past presidents of the American Evaluation Association (AEA) that theory or theoretical models should be used to guide practice (King, 2003). The challenge is to decide which approach or combination of approaches is most relevant to the evaluation and will produce a high quality evaluation.

Types and Classifications of Program Evaluation Models

There are varied and diverse models of program evaluation; to discuss them all would be outside the scope and range of this research project. Presented is a brief overview of one classification schema, to include one example for each classification. A classification often cited in the literature was developed by Fitzpatrick et al. (2011) that include:

- objectives-oriented in which the focus is on specifying goals and objectives and determining the extent to which they have been attained. The Tylerian evaluation approach is an example. This approach was named for Ralph Tyler (1949) who popularized this approach in primary education;
- management-oriented in which the central concern is on identifying and meeting the informational needs of managerial decision makers and stakeholders and for administrative accountability. For example, the CIPP model developed by Daniel Stufflebeam and Enon Guba. This model will be discussed in more detail later in this chapter;
- consumer-oriented in which the central issue is developing evaluative information on products with accountability to the consumer. For example, Michel Scriven's Key Evaluation Checklist approach;
- expertise-oriented which depend primarily on the direct application of professional expertise to judge the quality of whatever endeavor is evaluated. For example, the approach to evaluation used for formal professional review by accreditation bodies; and,

- participant-oriented in which involvement of participants or stakeholders are central in determining the values, criteria, needs, data, and conclusions for the evaluation. For example, Utilization-Focused Evaluation (UFE), also discussed in greater detail later in in this chapter.

Systematic program evaluation models contain similar general components, steps, and guidelines that include:

- identify stakeholders and clarify the purposes of the evaluation;
- analyze the context of the evaluation and set boundaries on what is to be evaluated;
- determine an evaluation approach or approaches;
- identify and select the evaluation questions and criteria;
- conduct the evaluation utilizing identified methods for data collection and analysis; and,
- interpret the results, report the results, and utilize the results (Fitzpatrick et al., 2011, Johnson, 1998; Patton, 2008; Rossi, Lipsey, & Freeman, 2004, Rovai, 2003).

Online Education Evaluation Approaches

There are many examples in the literature of the use of models in program evaluation of traditional or face-to-face adult education programs. But what does an evaluation approach or model look like for online program evaluation? Ruhe and Zumbo (2009) identified five characteristics from the literature of an approach that was modern and professional that could be applied to online program evaluation. First, the evaluation should be based on a theoretical model from professional program evaluation literature.

Models are guiding roadmaps. Second, the evaluation should provide a comprehensive assessment of values (merit and worth). Third, it should report on scientific evidence, underlying values, and unintended consequences. This will ensure comprehensiveness. Fourth, the approach should accommodate the use of mixed methods. Finally, the evaluation should use systematic, thorough, and rigorous methods and procedures.

In addition to the criteria offered by Ruhe and Zumbo (2009), there were other criteria identified from the literature applicable to this research project. First, the approach must be empirically tested and validated. Second, the evaluation model has to inform the research questions regarding evaluation practice and utilization. Third, the model needs to be easily adapted to encompass online nursing education. Fourth, clearly defined guidelines and steps need to be present from which to extrapolate items for survey development. Fifth, the model or models must meet the evaluation standards published by the Joint Committee on Standards in Evaluation. Finally, the evaluation model has to have been used in the context of online nursing education. No one model encompassed all the criteria.

Many of these evaluation models have been adapted from distance education models such as Chickering and Ehrmann (1996) model focusing on effective principles of education applied to the use of technology used in online education. The EDUCAUSE model (Newman, 2003) was also adapted from quality standards for online education. Both models were referred to in this literature review and have been applied across contexts.

There are other evaluation models that are specific to online education. Limitations of these models include lack of rigor and the tendency to rely on self-reports

and qualitative data (Rovai, 2003). Also, they tend not to meet all the criteria for an evaluation approach to online education noted by Ruhe and Zumbo (2009) above. For example, Kirkpatrick and Kirkpatrick's (2006) model focuses on four levels of evaluation; reaction, learning, behavior, and skills. The model is based on evidence collected in four different levels and does not include underlying values or unintended consequences (Ruhe & Zumbo, 2009). Therefore the model does not meet the criteria for comprehensiveness. In addition, since these models are intended for online education only, schools of nursing with hybrid or web-enhanced programs could be affected.

Some nursing researchers have developed benchmarks for online nursing education (Billings et al, 2001, 2005). These benchmarks have been used to develop evaluation frameworks for online nursing education (Avery et al., 2008). But, they tend to be context specific and the systematic approach to evaluation is questionable.

Conceptual Models for this Research Study

Using criteria developed by Ruhe and Zumbo (2009) and that was extrapolated from the literature, two program evaluation models were chosen to provide a conceptual framework for this research study; CIPP and UFE. Collectively, not only have these two models been empirically tested and validated but further lauded in the literature by other professional evaluators influential in the field of evaluation, scholars, and even past presidents of the American Evaluation Association. Daniel Stufflebeam (2001) and Patton (2008) undertook a comprehensive, exhaustive, and independent review of how 22 different evaluation approaches stack up against the standards. The research concluded that only nine out of the 22 could be identified as strongest and most promising for continued use and development. Decision and accountability models, such as CIPP,

scored highest. UFE was also listed among the strongest approaches, with highest rating for utility which aligned most readily with the standards.

Individually, the CIPP model has been around since 1969 and continues to have relevance and influence in the field of evaluation. Under the leadership of CIPP developer, Daniel Stufflebeam, the Joint Committee on Standards in 1981 published standards for evaluation (Joint Committee on Standards for Educational Evaluation, 1994). These standards are commonly agreed on characteristics of good evaluation practice from those in the field. Stufflebeam (1999) also developed a program evaluation model metaevaluation checklist based on the standards for evaluation and the CIPP framework. CIPP has been adapted for use in nursing education, program evaluation, and research (Hall, Daly, & Madigan, 2010; Singh, 2004, Suhayda & Miller, 2006). Hall et al. (2010) used the steps and guidelines defined in the model to develop an instrument for their research study. The CIPP framework has also been adapted to evaluate online education (Rovai, 2003). Rovai's (2003) CIPP model adapted for online education will be used to inform this research study. The model will be discussed in detail in Chapter 3.

Although the CIPP model meets the quality standards for evaluation (1994), the approach is management oriented and emphasizes decision making and accountability of managers (Fitzpatrick et al., 2011; Stufflebeam, 1971, Stufflebeam & Shinkfield, 1984). Fitzpatrick et al. (2011) stated that one critique of CIPP is the potential for it to be unfair to stakeholders who may have less power and resources than managers and administrators.

In order to inform research question number four regarding use and how important stakeholder involvement is for utilization of evaluation data, the evaluation

literature was used as a guide. Evaluation use or utilization is one of the few topics in evaluation research on which numerous empirical studies seem to exist. Christie (2007) noted, "Evaluation utilization is arguably the most researched area of evaluation and it also receives substantial attention in the theoretical literature" (p. 8). There were many models, mostly participant-oriented, and theoretical frameworks represented in the literature, but Utilization-Focused Evaluation (UFE) was chosen by the researcher. UFE meets each standard for quality evaluation, with the highest rating for adherence to the utility standard (Patton, 2008, Stufflebeam, 2001). UFE has also been extensively studied in the literature and has been empirically tested and validated (Cousins & Leithwood, 1986; Christie, 2007; Johnson, Greenesid, Toal, King, Lawrenz, & Volkov, 2009).

During the literature review on utilization, one issue noted repeatedly was the variety of terms used to describe and define evaluation utility. It was not a question of how important utility is to evaluation and along with the Joint Committee on Standards for Educational Evaluation (Yarbrough et al., 2011) professional evaluators and researchers agreed that it is a vital component in the evaluation process. However, there is debate (Kirkhart, 2000; Patton, 2008; Weiss, 1980, 1993) among evaluators over which term is most appropriate when discussing evaluation utility; influence, use, and/or utilization. These three terms were either used independently and exclusively and/or interchangeably in the literature. A decision needed to be made regarding which term would be most appropriate in communicating evaluation utility.

It seemed as if the terms use and utilization were used most often in the literature. Although this being true, Kirkhart (2000) felt the term use was too narrow and wanted to get rid of it in order to build a more complete theory of evaluation's consequences using

the concept of “evaluation influence” as a unifying construct. Kirkhart described evaluation influence as “the capacity or power of persons or things to produce effects on others by intangible or indirect means” (p. 7).

Weiss (1980, 1993) expressed a preference for the term use and advocated for a more “fluid and diffuse” definition of utility (Weiss, 1980, p. 18). Weiss felt that utilization failed to capture how evaluation knowledge affect policy and that utilization implied instrumental episodic application. However, Patton (2008) envisioned utilization as a process and use as more instrumentation and episodic in its meaning. Johnson (1998), on the other hand, in describing different models of program evaluation and theoretical frameworks, used both terms interchangeably throughout his writings. In their definition of use, Alkin and Taut (2003) made reference to “evaluation use (or evaluation utilization)... the way in which an evaluation and information from the evaluation impacts the program that is being evaluated” (p. 1).

Other evaluators felt the term utilization sounded more academic and pretentious than use. For this reason, Patton (2008) a strong advocate for the term utilization and originator of the UFE model, agreed. Patton (2008) stated that “I much prefer the verb use instead of utilize, but I make use of both nouns-use and utilization-varying my usage by audience and context” (p. 109). Utilization is the dominate concept used in Patton’s work, but he does use the terms (use and utilization) interchangeably. Since the sample participants are not professional evaluators and have to be able to understand what was being asked in order to collect meaningful data, the decision was made to utilize the term use in the questionnaire and other communications with them. Furthermore, I do not want to appear pompous and pretentious. To ensure accuracy in referring to the literature,

quoting authors, and to avoid misrepresentation in interpretation and meaning all three terms will continue to be used interchangeably throughout the document.

Both the CIPP and UFE models will be discussed in detail below to include how they met criteria for being chosen as conceptual frameworks for this research study.

The CIPP Model

Possibly the best known and most influential evaluation model is the CIPP evaluation model. According to the CIPP model (Stufflebeam, 1971) there are four different types of decisions requiring four different evaluations. The four types of decisions are planning, structuring, implementing, and recycling. The four types of evaluations are context, input, process, and product and correspond to the given order of decision types. The initial letters of the four types of evaluations make up the acronym CIPP, by which the model is generally known and referred to in the literature.

The CIPP model was originally developed by Daniel Stufflebeam in response to concern over the ineffectiveness of the evaluation process in education, particularly evaluation's limited help in making decisions about programs. Stufflebeam (1971) believed that one reason for this ineffectiveness was the lack of "adequate evaluation theory" which was felt to be crucial to obtaining evaluation information for decision making. Later, Stufflebeam provided evidence that the CIPP model with its systematic approach to evaluation also met the information requirements for accountability in addition to decision making. Over the years, further development and improvement of the CIPP model is mostly attributed to Stufflebeam (Stufflebeam, 1971, 2000; Stufflebeam & Shinkfield, 1985).

Context evaluation serves planning decisions. According to Stufflebeam (1971) “the purpose of context evaluation is to systematically provide information that can be used by decision makers to make planning decisions regarding the establishment of new objectives, modification of existing objectives, or confirmation of present objectives” (p. 6). Input evaluation serves structuring decisions. Stufflebeam stated that the purpose of input evaluation is “to identify and assess alternative program strategies for achieving given objectives and to provide information to assist in detailing particular strategies” (p. 9). Process evaluation serves implementing decisions. Process evaluation is “designed to provide information during the implementation stages of a project or program” (Stufflebeam, 1971, p. 10). In process evaluation evaluators are looking for information on plan implementation, barriers that threaten program success, and what revisions are needed (Fitzpatrick et al., 2011). Finally product evaluation serves recycling decisions. Stufflebeam (1971) stated that the “purpose of product evaluation is to relate outcomes to objectives and to assess the overall worth of a procedure in terms of its effects” (p. 12).

In CIPP Stufflebeam (1971) also proposed a systematic approach to evaluation that is consistent with other evaluation models. These steps and guidelines include; (a) focusing the evaluation; (b) collection of information; (c) organization of information; (d) analysis of information; (e) reporting of information; and (f) administration of the evaluation.

There has been a plethora of evaluation research using the CIPP model since its development in 1969 in all levels of formal education, informal education, continuing education, and in performance evaluations. Unlike UFE, CIPP has been adapted for use in both clinical nursing and nursing education. This research project is interested in

CIPP's adaptation for use in nursing education in general and online graduate nursing education specifically. Again much of the literature reports have included practice and descriptive articles. Only one article was found that dealt with online nursing education and was discussed elsewhere in this chapter.

Suhayda and Miller (2006) described the use of the CIPP model to develop an evaluation plan to address all levels of nursing education which included online tracks. The CIPP framework guided data collection, decision-making, continuous quality improvement, and demonstrated educational outcomes. Suhayda and Miller (2006) gave four main reasons that the CIPP model worked for their school of nursing. First, it was an empirically and conceptually proven model by literature review. Second, the model's emphasis on decision making and accountability emphasized both proactive evaluation for program improvement and retroactive evaluation to assess program quality. Third, the authors found the CIPP model to be flexible and comprehensive and allow evaluation of one or more of the CIPP components depending on stakeholder, program needs, and schedules. Finally, the CIPP model adapted well to incorporate various formats and criteria standards developed by the accreditation bodies.

Singh (2004) believed the CIPP model had utility for the nursing profession.

Singh (2004) stated that:

One of the unique aspects of the CIPP model is that it provides a robust foundation for a flexible and innovative framework for nursing education evaluation, as one or more of the accountability components can be used at any one time in accordance with the institution's needs and evaluation budget. That is, each component can stand alone (p. 1).

To demonstrate the utility of the model in nursing education, Singh developed an evaluation matrix to evaluate an undergraduate nursing program. Singh (2004) adapted the model, to include the evaluation types and the steps involved in the CIPP framework, and ensured inclusion of the evaluation standards. Though the model was applied to the program, Singh did not report the research findings.

A descriptive study by Hall et al. (2010) adapted the CIPP model to clinical nursing education at the undergraduate level of nursing education. Stakeholders were identified as nursing faculty and students. The purpose of the study was the use of anecdotal notes by clinical nursing faculty to track student performance and to develop an objective tool for more formal evaluation. The main question addressed in this study was, “If used, how are anecdotal notes used by nursing faculty?” (p. 157). Using the adapted CIPP framework, 14 items were used to question faculty about their note use from the model’s four types of evaluation. The instrument was tested for reliability and validity. The researchers enjoyed a 67% response rate from clinical nursing faculty from six nursing programs in a regional area. Hall et al. (2010) found that 97% of clinical faculty used anecdotal notes during the student evaluation process, and the majority of faculty do so on a weekly basis. Based on faculty feedback and the CIPP evaluation model, a clinical nursing faculty tool was developed after study completion to support clinical faculty in note use (Hall et al., 2010). The researchers found the CIPP model helpful in instrument design, establishing objectivity in the subsequent evaluation tool, and establishing research credibility.

Chapman (2006), using CIPP, developed an evaluation plan for a fully online human resource development graduate degree. Chapman (2006) indicated that the CIPP

model was used as a framework because it placed emphasis on guiding, planning, programming, and implementation efforts. The model also emphasized that the most important purpose for evaluation was improvement; an importance identified by Chapman during a standardized review.

Finally, Rovai (2003) used the CIPP model to adapt a practical framework for evaluation of online distance education programs. Rovai (2003) used an adaptation of Stufflebeam's (2000) CIPP model to develop a systems approach to program evaluation of online education. According to Rovai, a systems approach to evaluation can categorize an evaluation by types; input evaluation, process evaluation, output evaluation, and/or impact evaluation. See Table 2.2 for a comparison and the eventual adaptation between the CIPP model and Rovai's system model for evaluating online programs. This framework will be further described in Chapter 3.

Table 2.2

Comparison and Adaptation of the CIPP and Systems Approach Models

CIPP	Systems Approach Model
Context focuses on planning decisions	Input identifies and evaluates systems capabilities
Input focuses on structuring decisions	Process focuses on implementation; what should be happening and is not
Process on implementation decisions	Output seeks to determine the immediate or indirect effects of the program
Product focuses on outcome attainment	Impact addresses the longer-term results of the program

The UFE Model

UFE is one of the most popular evaluation approaches that is used and taught in the profession today. It is an approach that attempts to make the process and results of evaluation more useful. Utility is the first of five published standards of evaluation, emphasizing the importance of evaluation use. Other standards include feasibility, propriety, and accuracy (Yarbrough, 2011). There has been a great deal of past research on use or utilization of evaluation, but it was not until Michael Patton (Patton, Grimes, Guthrie, Brennan, French, & Blyth, 1975) published his research report on evaluation utilization did the approach begin to evolve into an evaluation model.

To evaluate something means to systematically assess its merit, worth, value, quality, or significance (Scriven, 1991). Patton argued that this definition puts too much emphasis on the value-based judgmental nature of the evaluation process. Patton's (2008) utilization-focused approach emphasized the user and the definition stated:

Utilization-focused program evaluation is evaluation done for and with specific intended primary users for specific, intended uses. Utilization-focused evaluation begins with the premise that evaluations should be judged by their utility and actual use; therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration for how everything that is done, from beginning to end, will affect use. Therefore, the focus in utilization-focused evaluation is on intended use by intended users (p. 37).

Patton (2008) insisted that UFE approach does not advocate any particular content, model, methods, theory, and uses. Patton (2008) stated that UFE "is a process for helping primary intended users select the most appropriate content, model, methods,

theory, and uses for their particular situation” (p. 37). Utilization-focused evaluation is a framework in which established evaluation principles, practices, and high quality standards can be utilized.

Patton (2008) pointed out that UFE is non-linear and without formulated steps, but has devised a rather intricate flowchart to “capture the sometime circular and iterative nature of the process by depicting loops at the points where intended users are identified and again where evaluation questions are focused” (pp. 566-569). Patton’s (2008) 15 premises or prescriptions that constitute the approach do more to form a foundation for use of evaluations. From this flowchart, with its guidelines and steps, items for survey development can be extrapolated. These premises have evolved over time and the most current are listed:

- Commitment to intended use by intended users should be the driving force in an evaluation;
- Concern for utilization is ongoing and continuous from the very beginning of the evaluation;
- The personal factor contribute significantly to use;
- Careful and thoughtful stakeholder analysis should inform identification of primary intended users;
- Evaluations should be focused in some way; focusing on intended use by intended users is the most useful way;
- Focusing on intended use requires making deliberate and thoughtful choices
- Useful evaluations must be designed and adapted situationally;

- Intended users' commitment to use can be nurtured and enhanced by actively involving them in making significant decision about the evaluation;
- High quality participation is the goal not high-quantity participation;
- High-quality involvement of intended users will result in high-quality evaluations;
- Evaluators have a rightful stake in an evaluation in that their credibility and integrity are always at risk, thus the mandate for evaluators to be active-reactive-interactive-adaptive;
- Evaluators committed to enhancing use have both an opportunity and a responsibility to train users;
- Use is different from reporting and dissemination;
- Serious attention to use involves financial and time costs that are far from trivial; and,
- Commitment to improving practice means following up evaluations to find out how they were used (pp. 570-573).

Patton (Donaldson, Patton, Fetterman, & Scriven, 2010; Patton, 2008) believed that UFE is an approach that can respond to any situation and therefore can serve any evaluation purpose; formative, summative, or developmental. UFE is not restrictive in methodology and can collect any kind of data; quantitative, qualitative, or mixed methods. Any type of design may be employed to include naturalistic, experimental, or quasi-experimental. The focus of the approach can be on process, outcomes, impacts, costs, or cost benefits (Donaldson et al. 2010; Patton, 2008). This claim is substantiated in the literature and supported by research.

Patton (2008) noted six primary purposes for evaluation that are based on the diverse reasons why program evaluations are done. These six purposes are for: (a) rendering judgment; (b) facilitating improvements; (c) generating knowledge; (d) accountability; (e) monitoring; and (f) development. These six purposes were used in item development and are further discussed and defined in Chapter 3.

A search of the literature revealed only one article that indicated the use of UFE in online nursing education research (discussed elsewhere in this chapter). However, there are two excellent reviews of the empirical literature on evaluation use conducted over a 34 year span of time. One is by Cousins and Leithwood (1986) who identified 65 empirical studies of evaluation use conducted between 1971 and 1985. Much of this research on utilization was of a retrospective design; identified independent variables that promoted use of utilization of evaluations; and used such theoretical frameworks as communication theory, organizational theory, and decision theory. Cousins and Leithwood (1986) developed a conceptual framework from examining the 65 empirical studies. The framework clustered two categories of factors related to evaluation use; characteristics of evaluation implementation and characteristics of the decision/policy setting. Six characteristics were associated with evaluation implementation and included evaluation quality, findings, timeliness, relevance, communication quality, and credibility. The six characteristics associated with decision or policy setting included political climate, decision characteristics, competing information, personal characteristics, commitment or receptiveness to evaluation, and information needs (Cousins & Leithwood, 1986).

Included in Cousin and Leithwood's (1986) review was Patton's initial landmark study of how evaluations are used that formed the basis for UFE. Patton, Grimes, Guthrie, Brennen, French, and Blyth (1975) used a retrospective field study design to research the utilization of 20 federal health program evaluations. Patton et al. (1975) looked to assess the degree to which the evaluations had been used and to identify the factors that affected varying degrees of utilization. Patton et al. (1975) found that two major themes emerged. First, in order to make an impact on evaluation use evaluators must identify the decision makers who will utilize the evaluation information. This goes back to utilization-focused evaluation's first premise of intended use by the intended user. The evaluator needs to ensure that decision maker's questions are answered.

The second major finding was the importance of the personal factor in getting the evaluation results utilized. The personal factor has to do with the interests and commitments of the key stakeholders involved in the evaluation. Where the key stakeholders are interested, committed, and involved in the evaluation process then the evaluation is likely to be used (Patton, 2008).

More recently Johnson et al. (2009), in a second literature review on evaluation research, identified 41 empirical studies from 1986-2005. Johnson et al. (2009) used Cousins and Leithwood's (1986) framework to organize the research but also added stakeholder involvement to decision/policy setting and evaluation implementation to account for more recent research interest. In general, the same characteristics were also used for the two established categories. For stakeholder involvement, characteristics included involvement with commitment, communication quality, credibility, findings, relevance, personal characteristics, decision characteristics, information needs, and direct

involvement. In two noteworthy studies direct stakeholder involvement was reported as a positive influence on various types of use (Preskill & Caracelli, 1997; Turnbull, 1999). Though nothing was found on the use of UFE in online education, this model is ideal in providing a theoretical framework and guide for this research study.

Summary

This chapter focused on relevant literature related to systematic program evaluation in higher education in general and specifically in schools of nursing. As discussed elsewhere in this chapter, a review of the literature points to a lack of research in systematic program evaluation practice and use in these schools of nursing.

There are many reasons why research is needed in this area. Amid rising educational cost for students and organizational budget cuts in private and public higher educational institutions evaluation research data can help support these huge investments. Stakeholders, such as, students, faculty, administrators, and funders need cost data to make informed decisions regarding online higher education. The challenge facing nurse educators of online education is to gather enough data to perform systematic program evaluation in order to articulate distance education's place in teaching and learning (Rovai, 2003).

Research is also needed in program evaluation at the various levels of online higher education that contributes to evidence based education, theory development, and best practices. According to Lindsay et al., (2009) what is learned from systematic program evaluation data contributes to evidence-based nursing education and to "building a science of nursing education" (p. 181).

On the other hand, given the fact that there is a scarcity of research in systematic program evaluation in online education, the evaluation literature is chalk full of approaches, models, theories, and guidelines for systematic program evaluation. Therefore, the necessary tools to develop research in this area are available for use. I chose two such approaches or models to inform this research study; Rovai's adaptation of the CIPP model for online education evaluation and the UFE model. Utilizing these two models as frameworks, Chapter 3 will describe in detail the steps taken to investigate current practice and use of systematic program evaluation of master's degree level online nursing education.

CHAPTER 3

METHODOLOGY

The purpose of this quantitative study was to investigate to what extent schools of nursing are currently practicing systematic program evaluation of online education and how are they using the results. The following five research questions guided this study:

1. To what extent are schools of nursing systematically evaluating their online education activities at the master's degree level?
2. What are the sources of evaluation data?
3. What are the areas of evaluation?
4. To what extent are the evaluation results utilized by schools of nursing?
5. To what extent do institutional and program characteristics affect evaluation
(a) data source; (b) area; and (c) utilization in schools of nursing online master's degree level programs?

This chapter is organized into six main sections to include (a) conceptual framework, (b) instrument development, (c) study sample, (d) data collection, (e) data analysis, and (f) limitations.

Conceptual Framework

In order to investigate the research problem and questions, the researcher needed to identify an evaluation approach that would provide a framework for evaluation practice and evaluation utilization. It had to encompass both traditional and online education, since many schools of nursing master's degree level programs continue to

utilize both. Therefore, the framework needed to be drawn from nursing, program evaluation, and distance education professional literature in order to situate the study in the defined context and to meet the assigned criteria. In addition, it had to be comprehensive by outlining a methodology and a list of potential evaluation practices and activities that could be used to help inform the study questions and develop survey items. Two evaluation models met my criteria and both are fully described in Chapter 2; Rovai's adaptation of the CIPP model for online program evaluation and Patton's UFE model.

Rovai's Adaptation of the CIPP Model

Rovai (2003) developed a four-phase framework that he advances that meets the outlined criteria for item development. The current study will employ these four categories as its conceptual framework.

Rovai's (2003) framework provides for four major types of program evaluation within the context of online education. These four types are included in a program evaluation framework that synthesizes the systems model of evaluation with the evaluation strategies for evaluation. Rovai's (2003) framework was further adapted and four broad evaluation practice areas were identified; (a) evaluating input, (b) evaluating process, (c) evaluating output, and (d) evaluating impact. Each of these four areas of evaluation practice serves as principle dependent variables for the study.

During the process of conceptualizing Rovai's framework and the task of measuring relevant variables, the researcher quickly realized that each act of evaluation practice activity included three major components; area of evaluation data, source of data (students, faculty, records, alumni, employers), and method of data collection (interview, survey, focus group). The researcher concluded that the study could potentially use three

separate frameworks for understanding program evaluation practice. In one framework, the area of evaluation was an option. In another framework, the source of evaluation data (alumni, teachers, students, records, and employers) was an option. In another framework, methods of evaluation data collection (surveys, interviews, and focus groups) was still another option.

At the start of item pool development, the researcher tried to address all three potential frameworks, but immediately found out that the permutations were overwhelming and the outcome was an instrument that had hundreds of items. For example, for evaluating input area, data could come from different sources to include alumni, teachers, and students. Information could be obtained by varied means or methods to include surveys, interviews, and focus groups. Therefore every single evaluation practice could potentially multiply and present major problems later on and presented disturbing implications for a survey that needed to contain a finite number of items.

It was decided to let go of data collection methods or means as part of the organizing principle in item development. Because of this restriction the study was left with two guiding frameworks that needed to be interlocked; the area of evaluation as discussed above and the source of evaluation data from schools of nursing.

According to Fitzpatrick, Sanders, and Worthen (2011), “within each evaluation study, information sources will be selected to answer the particular questions posed” (p. 346). Fitzpatrick et al. (2011) posed two guidelines for choosing sources for evaluation data. First, identify the concept construct that must be measured in each evaluation. The key concept for each evaluation area construct is systematic program evaluation. Another

guideline is to consider who has knowledge of this concept. Rovai's (2003) identified several primary sources of evaluation data; students, teachers, existing documents or records, alumni, and employers. Therefore for this study the researcher identified for each practice area construct, practice source constructs for the information or data needed.

Ultimately, instrument construction in looking at evaluation practice was guided by this dual framework which was designed to produce scores for each of the constructs presented in Table 3.1. Therefore, a total of four variables are related to evaluation data area (input, process, output, impact), and five variables are related to evaluation data source (students, faculty, records, alumni, employers). Table 3.1 provides definitions of the four practice areas and how they interlock with the five evaluation data sources.

Table 3.1

Evaluating Practice Constructs by Area and Source

Evaluation Practice Area	Definition	Evaluation Practice Source
Evaluating Input	Identifies and evaluates system capabilities used to meet the target audience and satisfy their needs	Students, faculty, alumni
Evaluating Process	Provides information of all components of the program	Students, faculty, program records
Evaluating Output	To determine the immediate or direct effects of the program	Students, faculty, program records
Evaluating Impact	Addresses the longer-term results of the program and the extent to which the program reduced or eliminated student needs and the effects of the program on society at large	Program records, graduates, alumni, employers

Patton's Utilization-Focused Evaluation Model

In order to fully inform the research question regarding responsiveness, use, and how important stakeholder involvement is for utilization of evaluation data, the researcher chose UFE as a second framework. Patton (2008) stated that “UFE begins with the premise that evaluations should be judged by their utility and actual uses: therefore, evaluators should facilitate the evaluation process and design any evaluation with careful consideration for how everything that is done, from beginning to end, will affect use” (p. 37). UFE is an approach to program evaluation that attempts to make the process and results of evaluation more useful. Much of its development and advancement has been attributed to Michael Patton. Patton (2008) defined UFE as:

The systematic collection of information about the activities, characteristics, and results of programs to make judgments about the program, improve or further develop program effectiveness, inform decisions about future programming, and/or increase understanding. Utilization-focused evaluation is evaluation done for and with specific intended primary users for specific, intended uses (p. 39).

Patton (2008) noted that individuals that evaluate programs can conduct evaluations in ways that increase use, especially by being “intentional about the evaluation’s primary purpose” (p. 110). In order to promote use of evaluation data, Patton (2008) advances six primary purposes of an evaluation that promote use. These six purposes are for: (a) rendering judgment; (b) facilitating improvement; (c) generating knowledge; (d) accountability; (e) monitoring; and (f) development. These six purposes are based on the diverse reasons why program evaluations are done, always promoting the outcome of evaluation use. Promoting use has implications for every aspect of

program evaluation – design, measurements, analysis, interpretation, reporting, dissemination, and criteria for judging quality (Patton, 2008). See Table 3.2 for the six purposes outlined in UFE approach that promote the use of program evaluation data.

Table 3.2

Evaluating Use Construct

Evaluation Use	Definition
Promoting evaluation use	Purposes and activities designed to promote the use of evaluation data in order to render judgment, facilitate program improvement, generate knowledge about the program, for stakeholder accountability, to monitor programs, and for development of programs.

Patton’s (2008) UFE framework meets the outlined criteria for item development as it pertains to utilization of the evaluation to include the definition and six primary purposes. Using UFE’s definition, practices, and purposes items were developed that fully captured the variable, evaluating use. See Figure 3.1 for depiction of how evaluating use is depicted in the conceptual framework.

Institutional and Program Characteristics

Background and demographic data were collected on both the master’s level degree program and individual respondents at each school of nursing. See Appendix A, section seven, of the instrument for these items and questions. This study intentionally used the terms institutional and program characteristics to identify program and respondent background and demographic data.

Institutional and program characteristics of master's degree level programs and individual respondents will be used in the analysis of question five, "To what extent do institutional and program characteristics affect evaluation (a) data source; (b) area; and (c) use in schools of nursing online master's degree level programs?". Institutional and program characteristics data will also be used for the purpose of study sample description. Items identified for institutional and program characteristics are those attributes which are in place to facilitate the goals and objectives of a program. With input from the literature, major professor, methodologist, and personal experience, items were identified. This study attempted to determine if there is significantly statistical support for institutional and program characteristics influencing program evaluation practice and use of online nursing education at the master's degree level.

After a careful review of the literature, Rovai's (2003) adaptation of the CIPP model for program evaluation practice, and Patton's (2008) UFE framework, a conceptual framework for this study was constructed. The researcher predicted that selected institutional and program characteristics influenced systematic program evaluation practice and use of online nursing education at the master's degree level. The conceptual framework which defined the key variables as well as illustrated the relationships tested appears as Figure 3.1.

Instrument Development

This study required development of an original instrument. Instrument development was a meticulous process that took 16 months from start to finish. A number of events took place during that time to include a pilot study. Ultimately,

development resulted in the instrument that is included in Appendix A. The instrument consisted of 84 items. Support documents are in Appendix B.

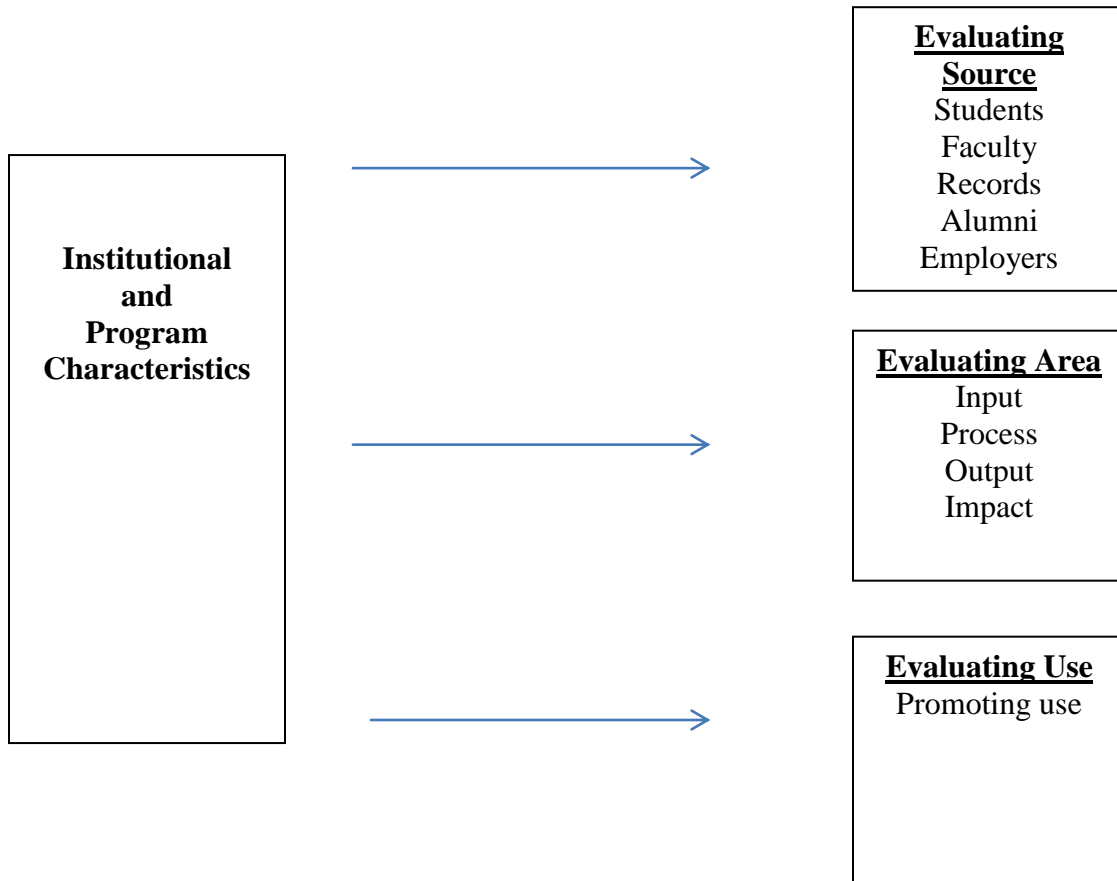


Figure 3.1. Conceptual Framework.

Survey items were drawn from a series of places to include the already identified conceptual frameworks used for this study, existing instruments and checklists, study construct definitions, the literature, from critique and interviews with expert nurse educators, methodologist, and other committee members. The steps in the item and eventual survey development process are summarized in Table 3.3. Throughout the instrument construction phase there was ongoing dialogue with the methodologist and major professor regarding item development.

Table 3.3

Steps to Instrument Development

Steps
Construct clarification
Identify study construct activities
Development and refinement of item pool
Expert review
Refinement of online survey and survey items post expert review
Response scale construction
Refinement of items and survey post-prospectus
Description of pilot instrument
Pilot study

Construct Clarification

The first step in instrument development, construct clarification, involved defining the constructs for the study. Rovai's (2003) four evaluation practice areas, five practice sources of evaluation data, and evaluating use are fully defined in the previous section of this chapter and in Chapter 2.

Identify Study Construct Activities

The second step was to identify study construct activities for the dependent variables from which to develop prototype survey items. Activities for the four evaluation practice areas and five practice sources were drawn from Rovai's (2003) framework for a systems approach to online program evaluation. Activities for the construct use were drawn from the UFE theoretical framework to include the definition for evaluation utilization, purposes used to promote use, and UFE checklist of premises. This step

produced a total of 98 activities for the dependent variables alone. See Appendix C for a list of construct activities and their sources and areas for the dependent variables.

Development and Refinement of Item Pool

The third step was to identify from the construct activities items for the initial item pool. This task proved the most difficult. Item development began by identifying activities that reflected the broad evaluation practices and activities for the dependent variables, evaluating practice and use. Initial work began with identifying activities for practice; evaluating input, evaluating process, evaluating output, and evaluating impact. As mentioned in the previous section, as work began to build the item pool the researcher realized that each evaluation activity was in and of itself a complex of three distinct features. These included:

- the focus of the evaluation or area, i.e., collecting data to evaluate satisfaction, teaching effectiveness, technology use in the classroom, etc.;
- the means or method by which data is being collected, i.e., interviews, focus groups, etc.; and,
- the source of that data, i.e., students, alumni, teachers, employers, etc.

This had disturbing implications for a survey that contained a finite number of items. For example, for the topic “technology use in the classroom”, information could be obtained from different sources to include alumni, teachers, and students. Data could be obtained by varied means, i.e., surveys, interviews, focus groups. Therefore every single area, source, and topic activity could potentially multiply out into six or eight items. This presented a serious challenge since a goal was set for a maximum number of items on a “doable” survey of approximately 60 items. Consequently, some decisions had to be

made that meant full theoretical coverage of all three areas would be forfeited. Therefore, ultimately for the sake of economy, the “how” or method or means by which data are gathered was removed and items were arranged by source and area of evaluation.

Even after arranging the activities by area and source the pool generated 81 items for evaluation practice alone. The items were repeatedly and strategically reviewed for redundancy and relevancy. Over several months and after numerous iterations the practice area and source item pool was refined and reduced from 81 to 48.

The next task in this step was to develop prototype items that would measure use in the systematic program evaluation process. This step was not as complex given that general evaluation activities and premises had already been identified in the UFE framework and evaluation literature. An initial pool of 10 items was constructed to represent evaluation purpose activities and UFE’s 15 premises.

As the final task in this step, the researcher developed prototype items for the demographic items; institutional and program characteristics. In the beginning the demographic variables only included program characteristics items. These items were extrapolated from the literature pertaining to nursing education organizations, from personal experience as a nurse educator, and from the methodologist and major professor. Fifteen items were developed for program characteristics. The initial item pool now contained 72 items that was ready for critique and review by an expert panel. See Appendix D for the complete list of items from the initial item pool.

Expert Review Panel

The fourth step was a review of the items by an expert panel. The researcher identified ten nurse educators to comprise the expert panel. They were chosen for their

many years as nurse educators, for their experience as researchers, and their administrative acumen. All ten are employed at the same school of nursing that is part of a major urban research university. Eight experts were contacted in person and two were contacted by email due to their lack of on-site availability. Eight out of the ten identified nurse educators chose to participate in the expert review.

The eight expert reviewers were all women. There was one professor (PhD) and five associate professors (PhDs). These five experts were actively involved in research and are known nationally for their research agendas. The panel further included two assistant professors (PhDs), both of which were directors of either undergraduate programs or graduate programs. They often were given the responsibility of program evaluation. The eighth expert was a clinical assistant professor who taught research at the undergraduate and graduate levels, was in a PhD program, and had collaborated on numerous research projects and published widely.

The expert review was done electronically. Using this method, the expert panel could also critique the online survey. Reviewers were also asked to participate in a 15 minute face-to-face or telephone conversation with the researcher in order to walk through the survey item by item to tell what was right or wrong with each item and if there were major omissions.

A hardcopy of the survey was formatted and laid out in draft form for review by the methodologist and major professor before entering items into SurveyMonkey. After the items were reviewed and refined they were formatted for hardcopy for delivery via email and online delivery via SurveyMonkey for expert critique.

Each expert was sent a letter of participation with instructions, a link to the online survey, and a hardcopy of the survey. The experts were asked to rate in their opinion the importance of each item to online program evaluation using a 5-point Likert scale ranging from not important to extremely important. They were also asked to make comments on the hardcopy regarding the items. See Appendix E for expert reviewer's survey instrument.

The researcher met with each of the eight experts individually in a follow-up conversation to discuss and solicit further item critique. This also served the purpose of clarifying any questions regarding comments made in the online version of the survey. Specific questions of concern regarding each items was not given to the experts. Interest was in the open and ad-lib comments made regarding the survey items and online survey. As a result, the researcher obtained additional data that would help with item refinement.

The expert critique panel results were compiled for review and discussion with the methodologist and major professor. This led to the next step in instrument development, refinement of survey items. A summary of the expert panel critique is presented as Appendix F.

Refinement of Survey Items Post Expert Review

In general, the experts agreed there was an appropriate number of questions on the survey and that it took 10-15 minutes to complete the online survey as stated in the introductory letter. Overall the online survey was visually appealing and received comments from good to great. No items were deleted for redundancy or irrelevancy. This was positive feedback for the process of repeated reviews of the survey items before the expert critique.

Many of the suggestions pointed the way for improvement and were adopted.

Principle changes included:

- Adding questions to determine eligibility for the study or “a weed out question”.
- Clarifying terminology that related to technology, i.e., examples, definitions, different terminology.
- Adding participant demographic items, i.e., job title, faculty and evaluator experience.
- Changing the response scale in order to capture frequency of evaluation practice reflecting actual practice.
- Using headers that would highlight each category, i.e., Section 1, Section 2, etc.

Certain changes were not implemented. Six of the eight experts made a recommendation to remove the recurring stem before each item. A decision was made to continue to use the stem, wanting to ensure that participants were continually reminded of the topic and source of measurement. There was a possibility that in not including the recurring stem could pose a potential threat to validity and interfere with the accuracy of the intended measurement of the item.

Another recommendation was to categorize each section differently. The experts suggested different headings, such as, technology, students, faculty, curriculum, student services, retention, and evaluation. After discussion with the methodologist the categories remained the same.

Going into the expert review, the researcher did have some trepidation about the utilization items drawn from the UFE framework. This trepidation proved to be fruitful because seven of the eight experts raised questions and voiced confusion about what was being measured; planning for use of the evaluation versus the actual use of the evaluation versus planning to evaluate the program. Ultimately this confusion challenged the ability to measure utilization in a meaningful fashion and a major validity issue could be whether they understood the information that was desired. After all, the experts were not trained in program evaluation. After extensive discussion with the methodologist and major professor it was decided to leave these items in the survey. However, there was a change in wording, hoping to obtain this vital utilization information. Instead of asking “how often does your school of nursing promote or participate in each of the following practices...” it was asked “to what extent is your school of nursing engaging in evaluation utilization practices in order to enhance use”. The items reflected these changes in preparation for the pilot. The researcher awaited any additional input the pilot would yield regarding these items.

Another major critique was of a previously included research question pertaining to barrier and facilitator factors as predictors of online program evaluation. Experts could not identify items that would measure these factors in the survey. This was evidence that the researcher had not adequately defined nor operationalized this construct in a way in which the experts could identify. This was also a validity issue. Ultimately, after many attempts to further operationalize, this led to a decision to abandon this part of the survey and not to include barrier and facilitator factor items as part of the pilot or the current research study.

Response Scale Construction

The final step in instrument development was construction of a response scale that would appropriately measure the items. One expert reviewer raised a question about information regarding actual or model evaluation practices. The researcher quickly determined from this feedback that actual practices were the target measures and that an agreement scale would render unattainable any data that would satisfy the purpose of the study. Therefore, a frequency scale would be most appropriate to ask participants how often or how many times an evaluation activity had taken place.

The researcher vacillated about using a numeric frequency scale, noting that such a scale would be neither consistent nor appropriate when moving from item to item in the order in which they were formatted and laid out in the instrument. Several verbal frequency scales were tested and evaluated before tentatively deciding to use a four-point Likert scale such as; “never” (1), “rarely” (2), “occasionally” (3), and “always” (4). The methodologist quickly pointed out that such a scale seemed dis-embodied and that a more concrete scale was needed. It was decided to still use a four-point Likert scale but with more defined choices: 1 = Never; 2 = 1 or 2; 3 = 3 or 4; and 4= 5 or more. Subsequently survey participants would be given the following directions in order to focus their answers:

Below you will find a list of activities and practices used to evaluate online education programs. We are interested in how many times in the past five (5) years did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master’s degree level nursing programs.

See Table 3.4 for an example of the response scale used to measure the four practice constructs using two items from the survey.

Table 3.4

Survey Instrument Response Scale to Measure Evaluating Practice

Items	Never ↔ 5 or more			
1. Collecting data from students about satisfaction with online learning	1	2	3	4
2. Collecting data from students about perceived value of course content	1	2	3	4

Initially, I used the same four point Likert response scale to measure the variable evaluating use that was used to measure evaluating practice. See Table 3.5 for an example.

Table 3.5

Survey Instrument Response Scale to Measure Evaluating Utilization

Items	Never ↔ 5 or more			
49. Engaging in utilization practices by identifying stakeholders in the evaluation process	1	2	3	4
50. Engaging in utilization practices by assessing program readiness for evaluation	1	2	3	4

The prototype survey was now ready for prospectus defense. See Appendix G for pre-prospectus defense prototype instrument to include revisions after the expert review.

Refinement of Survey Post-Prospectus Defense

At prospectus defense questions were raised from committee members about three main areas of the prototype survey. First section six and the items that pertained to evaluation utilization. Discussion focused on the term “utilization” used in the items and specifically if study participants would understand the items pertaining to this term. It was suggested not to assume understanding from participants who were not professional evaluators and to use the term “use” instead. Second the wording of the utilization items were confusing, i.e., “engaging in utilization practices by using evaluation data to improve online programs”. It was felt that the term engaging was not what was actually happening with determining use and that the items should be revised. Finally, the response scale was not appropriate to measure evaluating use and the extent of utilization of evaluation data in schools of nursing.

Subsequently, the researcher met with the methodologist and major professor to discuss these issues. The major effort was to revise the 14 items in section six of the survey to reflect the change in terminology and brainstorm different ways to word the items to measure what we were asking. Ultimately, the items were revised and a different response scale determined in order to insure understandability among study participants of the utilization items. Also items were reworded to reflect change of terminology to “use” from “utilization”.

Another response scale was also developed for the use items. This also involved rewording the stem question and instructions. The instructions, question, and response scale are as follow:

No matter what data you get, the evaluation process will require that your school of nursing engage in a number of specific activities to enhance evaluation

utilization in order to make improvements to or decisions about your program. To what extent is your school of nursing engaging in the following program evaluation utilization practices?

The 4-point Likert scale was: 1 = Not at all; 2 = Somewhat; 3 = Greatly; and 4 = Totally.

Sections one – five of the prototype survey were unchanged.

As suggested by the committee a pilot study was conducted. The pilot study had two major objectives. First was to test the data collection procedures and to determine if an adequate response rate would be achieved. The second objective was to determine how well the instrument itself functioned and this was largely accomplished by examining item distribution and reliability.

Description of Pilot Instrument

The pilot instrument had 78 items. The items measured constructs for the four major evaluation areas grouped in five sections by source, the variable evaluating use in section six, and the institutional and program characteristics in section seven. Twenty items were used to measure evaluating input construct; eleven items were used to measure evaluating process construct; thirteen items were used to measure evaluating output construct; four items are used to measure evaluating impact construct; ten items are used to measure evaluating use construct.

In section seven, the institutional and program characteristics items had also been refined and items added to obtain more useful and accurate data about the schools of nursing and their master's level degree programs. The researcher took recommendations from the expert review, the literature, and consultation with the methodologist and major professor. In addition, three general items pertaining to faculty participants were included that would more completely inform the study. After refinement and checking for

relevancy and redundancy 20 items remained for organizational and program characteristics; 13 pertaining to institutional and program characteristics and seven pertaining to faculty.

Furthermore, two open ended questions asking how the instrument performed were added at the end of the survey. The two questions comprised a section eight and are as follows:

- Did you have any trouble completing this survey? If so, please specify.
- Were there any problem items that needed to be improved?

These two questions would not of course be on the final research study instrument. See Appendix H for the pilot version of the survey with supporting documents after all revisions were made.

Pilot Study

After obtaining Institutional Review Board or IRB approval for a pilot along with the main study, the pilot survey was sent to 90 randomly selected schools of nursing and their nurse administrators. The random sample was taken from the population sample of 473. Utilizing the revised eight-section instrument, the survey was sent using the online survey software, SurveyMonkey. See Table 3.6 for approach to pilot data collection.

A cover letter was sent to introduce the study and provide information on accessing the survey link. The questionnaire included the informed consent as the first page of the survey. The second page contained an eligibility question with the option to check “no” if the school of nursing did not have online nursing education. If the “no” option was checked, SurveyMonkey would advance the respondent to the last page of the survey that stated my thanks and appreciation for participation.

Table 3.6

Approach to Data Collection for Pilot

Date	Nature of Contact
May 22	Initial contact: cover email letter with survey link sent to participants
June 5	First reminder email letter sent to participants not yet responded
June 21	Second reminder email letter sent to participant not yet responded
July 3	Survey closed

The cover letter with a survey link was emailed to each of the 90 randomly selected nurse administrators/schools of nursing. A second reminder was emailed two weeks after the initial contact. A third reminder was emailed two weeks after the second contact. In each contact email letter, participants were given the option of a PDF copy of the instrument and with directions on where to mail it once completed. Each participant was promised a copy of the doctoral study's executive summary to be emailed upon completion of the study.

Overall the instrument performed extremely well. See Appendix I for a distribution and reliability results for key measures. However, the pilot study did not yield the response rate (23%) that we had hoped for. See Table 3.7 for a summary of the pilot response rate.

The response rate of 23%, 19 respondents out of 82, was poor. An optimum response rate according to the literature is 60%. Some reasons for the poor response rate were speculated.

- The busyness of nurse administrators and their finding the time for even a 20 minute survey.
- Inconvenient time of year at the beginning of a new academic year.
- The survey was not visually appealing or too long.
- The lack of an incentive offering.
- Schools of nursing did not have online nursing education offerings.

Table 3.7

Summary of Participation and Response Rate for Pilot

Sample	Number
Total random sample	90
Total participants opted out of SurveyMonkey	6
Total ineligible respondents (no online nursing education)	2
Total eligible participants	82 (90-6-2)
Total eligible respondents	23
Total incomplete or unusable surveys	4
Total usable surveys /respondents	19 (23-4)

A memorandum was sent to the members of the committee to update them on the results of the pilot study. To address the non-response issue the committee and researcher brainstormed many solutions and proposed the following changes to the survey and

the methodology. The committee agreed to the following contingent upon IRB amendment approval:

- intentionally and strategically send out contact emails; pre-notice and follow-on first and second contacts;
- decrease the timing between email contacts from two weeks to one week;
- make wording of the contact letters more personable and inviting to participants;
- cue respondents to how many sections are remaining in the survey, i.e., Section 1 of 7;
- use Qualtrics, another web-based survey software platform, since participants may have opted out of SurveyMonkey due to multiple uses; and
- if the response rate is less than 45% after the first reminder, implement a one-time mail and paper survey to non-respondents only as the second participant reminder.

IRB approval was obtained for all the above amendments and revisions to include a mail and paper survey as a last reminder to non-respondents.

At this stage it is necessary that the researcher acknowledge a clerical error that was made in the production of the final questionnaire. In section six with the items that deal with utilization there is a minor disjuncture between the focus question or the directions and the response scale. The researcher inadvertently put slightly wrong directions, referring not to how much they agreed with the items but how frequently they did the items. However, and somewhat surprisingly, the impact was relatively minor for this clerical error. Evidence of this is observed in terms of reliability; the utilization scale

demonstrated a reliability or Alpha of .77. Apparently the wording of the items and the 4-point Likert response scale used, strongly disagree to strongly agree, were intact enough to override any confusion from the directions or focus question. See Appendix A for the main study survey.

Study Sample

This study was concerned with investigation of systematic program evaluation of online master's degree level nursing programs. The researcher anticipated that most master's degree level nursing programs have transitioned partially or completely to an online format. The survey sample was drawn from all schools of nursing master's degree level programs that are accredited by the Commission on Colleges in Nursing Education (CCNE) or the National League for Nursing Accrediting Commission (NLNAC). The CCNE and the NLNAC are the two national professional nursing accrediting agencies that accredit all public and private nursing programs in the U.S.

The CCNE and NLNAC are membership-based organizations and mandate ongoing program evaluation. Evaluation among and within schools of nursing in the U.S. tend to be homogeneous. The CCNE and NLNAC websites contain and maintain a listing of all the nursing programs they accredit. In addition to a list of accredited programs the websites also contain schools of nursing names, names of deans and/or directors, and contact information to include postal addresses, email addresses, and phone numbers. Initially, several options for participants were considered to include chairs/deans/directors, associates/assistant directors, and faculty from each school of nursing with master's level degree programs. This could present a major problem with overall data collection management. Ultimately, it was decided that what deans and/or

directors (or their designee) had to say about how their online programs are evaluated was of the greater interest, since they are the ultimate decision makers for the particular college or school of nursing programs. Therefore, it was this group that comprised the sample.

At the time of the main study, the population sample totaled 383 schools of nursing or population participants. This included the most recent accredited schools of nursing and those that were no longer accredited listed on the CCNE and NLNAC websites since the pilot study. Excluded from this total were the 90 schools of nursing participants used for the pilot. Schools of nursing with dual accreditation were counted as one participant for the sample.

Data Collection

The study used mixed-modes as a means of data collection; Internet survey and mail and paper survey. I closely followed the tailored design method developed by Dillman, Smyth, and Christian (2009) for Internet, mail, and mixed-mode surveys. This included guidelines for the layout and design of the questionnaire, methods of assembling, and multiple distribution and contacts using both methods. See Table 3.8 for a summary to the approach to data collection.

Internet Surveys

The initial plan for data collection was to use Internet survey only. The use of the Internet is a useful mode for conducting surveys targeted at very specific populations such as master's degree level nursing programs (Dillman et al., 2009). The use of an online survey has been made easier by the availability of numerous web survey software packages and services. As suggested by the methodologist, for the main study the

researcher used Qualtrics, another web-based survey software platform. The main reason being that many participants may have opted out of SurveyMonkey due to multiple

Table 3.8

Approach to Data Collection

Date	Nature of Contact
August 19	First contact – Email pre-notice letter announcing the survey and asking for correct contact information
August 30	Second contact – Email cover letter with survey link, informed consent, instrument
September 13	Third contact – Email first reminder to non-respondents
October 16	Forth contact - Mail and paper survey to non-respondents
November 16	Data collection period closed

contacts and the burden of too many online surveys. Qualtrics is a newer Internet survey tool and participants may not have had time to opt out of receiving surveys.

There are many advantages to Internet-based surveys. The Internet can potentially reduce some types of survey errors. First, use of the Internet can reduce coverage error. Coverage error is “when all members of the population have a known, nonzero chance of being included in the sample for the survey and when those who are excluded are different from those who are included on measures of interest” (Dillman et al., 2009, p. 17). The CCNE and NLNAC lists of accredited nursing programs contain all nursing programs in the U.S. along with contact information such as director’s name, phone number, and email address. Obtaining information from these organizations increased

the likelihood that all nursing programs that meet criteria will be in the survey population or sample.

Second, use of the Internet can reduce sampling error. Sampling error “results from surveying only some rather than all members of the population and exists as a part of all sample surveys” (Dillman et al., 2009, p. 17). The study sampled all online master’s degree level nursing programs.

Third the administrative support and tracking of participants increased convenience and accuracy in data record keeping and data collection. Online survey software is available to assist the researcher in constructing surveys and there are services for data collection, storage, and analysis (Dillman et al., 2009; Ritter & Sue, 2007).

In using Qualtrics, multiple participants were reached within the same and short space of time using a panel distribution. Qualtrics assigned an identification number to each participant and tracked their response or non-response. This was invaluable when it came time to send reminders to non-respondents. Qualtrics also made data analysis less time consuming. The online survey tool featured a means to download data to analysis software such as SPSS.

Fourth, online surveys are also economical. There are no interviewers or data entry personnel, other than the researcher, to train and pay (Ritter & Sue, 2007). Travel expenses are kept to a minimal. There is no need for postage. Expenses include the cost of survey software or the services of a web-based survey host. Access to Qualtrics was through the university and obtained without cost to the researcher.

Finally, online surveys can save time. The use of online surveys has the potential to collect a lot of data much quicker than by other modes. Response time is cut short and

responses can be transmitted to the researcher immediately or posted to a database. The researcher paid close attention to potential limitations of online surveys. First, steps were made to ensure emails were not flagged as spam. Second, the participants were provided clear instructions for how to access the survey. And third, according to Qualtrics there were no bounced emails to deal with.

Mail and Paper Surveys

According to Dillman et al. (2009), mail and paper surveys continue to be widely used. In fact, mail survey response rates can be significantly higher than those obtained by Internet and telephone. Mail survey is used more and more in combination with other modes such as the Internet. This is very applicable when response rates are low.

Strengths of a mail survey include:

- the availability and accessibility of technologies and software paper surveys are easy to design and administer;
- enhanced coverage;
- the ability to deliver incentives more easily;
- access for participants that prefer paper surveys; and
- can be used effectively with mixed-mode design (Dillman et al., 2009).

Limitations include cost of design and implementation, a lengthy field period required to conduct quality mail surveys, and participants who may not possess the ability to read. An inability to read or lack of education was not an issue with participants for this study.

Mixed-Mode Design

This study used a mix-mode design; Internet and mail and paper survey. The researcher did not achieve the desired results in response rate with the first two Internet participant contacts alone. Mixed-mode survey design was recommended by one committee member. There are two main advantages to this approach in the literature; reduce the coverage error associated with the use of the Internet survey and to improve response rates in an effort to reduce nonresponse error with the Internet. Since all participants had an assigned email address, coverage error was not an issue for this study and Qualtrics did not report incorrect or bounced addresses. On the other hand, there could have been errors in program contact information and emails could not be forwarded since the participant email address was associated with an identification number within Qualtrics. Therefore, mixed-modes were used in hopes of improving the response rate.

For this study, both modes relied on similar forms of visual communication. The same survey structure and wording were used to ensure that a participant interpreted the items and questions the same across modes. When using mixed-modes these two issues could have presented barriers and limitations to the study.

Approach to Data Collection

Keeping in mind the low response rate from the earlier pilot study, every opportunity was made to maximize a high response rate. Closely following the guidelines developed by Dillman, et al. (2009), on August 30, 2012, an email pre-notice was sent to all 383 participants using the name and email address of the individual contact listed on the CCNE and NLNAC websites. The pre-notice and all following contact email letters were personalized, i.e., “Dear Dr. Doe”; this is a function within Qualtrics. The body of

the pre-notice introduced the study, gave information about the study, appealed for the participant's help, requested participation, and indicated what would be the value of such research to the nursing profession. The researcher also asked the participants to confirm if the addressee was or was not responsible for program evaluation at the master's degree level within their school of nursing. If they were not, participants were asked to supply the correct contact information. The participant panel within Qualtrics was updated to delete schools of nursing that had responded "no" to having online master's level degree programs. The participant panel was also updated to reflect correct or alternative contact information. Participants were informed to look for the survey within one week. See Appendix A for the main study survey and Appendix B for supporting documents to include the pre-notice email letter.

On September 6th, 2012, the second email contact was sent via Qualtrics to 376 participants. From the 383 original participants five had been excluded. Two respondents after receiving the pre-notice, wanted IRB approval from their own particular organization and five other respondents who reported no online education via email communication.

The second Internet contact included a cover letter with the link to the survey, the informed consent, instructions on how to take the survey, the instrument, and a thank you page. The cover letter included access to a paper copy of the survey via a link to a PDF version along with my contact information to include mailing address. The informed consent gave participants the option of being taken to the end of the survey and a thank you page with a click "no" response or continuing to the instructions page with a click "yes" response. The survey instructions page included a detailed definition of online

education used for this study and an eligibility question. If the answer to the eligibility question was “no”, respondents were taken to a thank you page at the end of the survey.

The survey was formatted to peak and keep the respondent’s interest, for visual appeal, and ease of navigation. Once respondents had completed and submitted the survey, they were immediately taken to a thank you page within Qualtrics. Respondent’s data was recorded and stored by Qualtrics for later retrieval and analysis. See Appendix B for a copy of the first contact email to include letter, informed consent, instructions, survey, and a thank you page.

The first contact yielded 33 responses. The researcher informed the methodologist and major professor of this dismal response. With their approval the third contact and first reminder email letter was sent out. Simultaneously, it was highly recommended that an amendment be submitted to IRB to seek approval for the fourth contact and second reminder be a mail and paper survey. As noted earlier, mail and paper had been suggested by one of the committee members after seeing the response rate from the pilot.

On September 13th, 2012, the third email contact/first reminder was sent to all non-respondents. The participant list for the first reminder had been updated by Qualtrics to remove all respondents and two participants that had opted out of Qualtrics surveys. Therefore, the first reminder was sent to 341 non-respondents in the same format as before; cover letter with access to the link or PDF version, informed consent, instructions, the instrument, and a thank you page. The first reminder yielded 35 additional respondents. This included one respondent that used the PDF option for response and mailed the survey to the address provided. This respondent’s survey data was manually entered into Qualtrics for continuity of data collection, tracking, and later analysis.

IRB gave approval for the mail and paper survey on October 4th, 2012. Immediate preparation began for the mail and paper survey. The documents included a cover letter, the informed consent, the instrument, and a stamped self-address return envelope. To lend legitimacy to the packet The University of Georgia logo was added to the addressee label and cover letter. There were no major changes to the research design. The participant list used for Qualtrics was correlated with a hard copy maintained by the researcher. Individual identification (ID) numbers were manually assigned to non-respondents in order to correctly match each mail respondent to their name, email address, and electronically assigned ID given in Qualtrics. This later on allowed for ease and accuracy of data transfer from the mailed survey into Qualtrics.

The instrument was downloaded into Word format and saved in PDF format for mail out. Due to the expense of mailing 301 packets, the researcher was able to request and receive funds from an advanced education degree grant to assist with postage. Using guidelines from Dillman et al. (2009) tailored design method 301 surveys were formatted and distributed to non-respondents via mail and paper survey on October 15th 2012. At the close of data collection on Nov. 16th, 84 participants had responded via mail and paper survey. See Table 3.9 for summary of responses by method.

Response Rate

A mixed modes approach was used to maximize participant response to the survey. Originally, there were 473 participants in the population sample. Ninety participants were used for the pilot study leaving an adjusted population sample of 383 participants. Of the 383 participants 39 (10%) respondents were ineligible for the study due to not utilizing online nursing education at their school of nursing. Respondents

indicated ineligibility via email correspondence (11) after receiving email contact and by Internet and mail and paper response (28). This left a final population sample of 344.

Table 3.9

Summary of Responses by Method of Data Collection

Survey Contact	Mailed	Received
1 st contact/Internet	376	33
1 st reminder/Internet	341	35
2 nd reminder/Mail and Paper	301	84
Total		152

Of the 344 population sample, nine participants (3%) refused to respond or opted out of Qualtrics. Seventeen participants (5%) actually opened the questionnaire in Qualtrics and never responded to the survey. Usable data were collected from 107 respondents resulting in a research study response rate of 31%. See Table 3.10 for a summary view of participation and survey response rate calculation.

Personal Characteristics of Study Respondents

Respondents ranged in age from 33 to 78 years with a mean age of 57.84. The majority of respondents were self-reported as White/Caucasian (82.2%). The remainder was African American (3.7%); Euro-American (.9%); Hispanic (.9%); White/Hispanic (.9%); White/South African (.9%); and other (.9%). The overwhelming majority was female (85%). Respondents providing data for this study reported less than 1 year to 50 years ($M = 22.11$; $SD = 9.72$) as nurse educators. The top three job titles of

the respondents included “Dean” (14%), “Associate Dean” (7.5%), and “Chair” ((4%).

See Appendix J for a complete list of respondent job titles. Respondents reported being in their job titles from less than 1 year to 18 years ($M = 3.82$; $SD = 3.76$) years. See Table 3.11 for a summary of personal characteristics of study respondents.

Table 3.10

Participation and Response Rate Calculations

Sample	Number	Value
Population	473	
Random sample used for pilot	90	
Adjusted population	383 (473-90)	
Total ineligible respondents	39	10% (39/383)
Final survey population sample	344 (383-39)	
Participants refused	9	3% (9/344)
Previewed/never responded	17	5% (17/344)
Total respondents	107	31% (107/344)
Total non-respondents	211 [344 - (9 + 17 + 107)]	

Institutional Characteristics of Study Respondents

The majority of the schools of nursing were public institutions (58%) versus private (43%). The overwhelming majority were accredited by CCNE (75.7%).

Organizations at the state, regional, and profession level were indicated by four schools of nursing as “other” for accreditation. The size of respondents’ online programs was reported as number of students which ranged from two schools of nursing with 18 to one

school with 5500 with a median of 132.50. One other school also reported a large number of students (4000). These numbers were questioned and checked by the researcher for accuracy and found to be correct as recorded. Small programs were those with less than 49 students (18.7%), medium sized programs contained between 50 – 149 students (29%), large programs had 149-299 students (20.6%), and very large programs reported numbers of student greater than 300 (35.2%).

Respondents reported a mean percent of 36.49% (SD = 32.62) of students attending online programs full time. Online programs are taught by a mean percent of 24.65% (SD = 25.16) of part time faculty. Schools of nursing reported a minimum of two years and maximum of 60 years that their master's degree program had been in existence. The number of years using online course work varied from 1 to 16 (M = 8.38; SD = 3.90) years. The percentage of schools of nursing using both (42%) outnumbered web-enhanced (28%) and web-based (25%) as methods of delivery. Approximately 9% utilized online education in the program core curriculum courses, 4.7% utilized online education in the program specialty track courses only, and 80.4% reported utilizing online education in both. Study respondents reported the most recent year of school of nursing evaluation of their online master's degree level program as within the past 5 years; which would correspond with CCNE and NLNAC requirements. Overwhelmingly tuition had increased (86.9%) within the past two years in educational organizations. See table 3.12 for a summary of institutional characteristics of study respondents.

Table 3.11

Personal Characteristics of Study Respondents (N=107)

Characteristic	n	Values
Age		M = 57.84; SD = 7.58
Missing data	13	
Race		
White/Caucasian	88	82.2%
African American	4	3.7%
Euro-American	1	.9%
Hispanic	1	.9%
White/Hispanic	1	.9%
White/South African	1	.9%
Other	1	.9%
Gender		
Female	91	85.0%
Male	10	9.3%
Missing data	6	5.6%
Years as Nurse Educator		M = 22.11; SD = 9.72
Missing data	4	
Job Title		
Dean	15	14.0%
Associate Dean	8	7.5%
Chair	4	3.7%
Other	80	74.8%
Years with Job Title		M = 3.82; SD = 3.76
Missing data	7	

Table 3.12

Institutional Characteristics of Study Respondents (N = 107)

Characteristics	n	Value
Type of organization		
Public	58	54.2%
Private	43	40.2%
Missing data	6	5.6%
Accreditation		
NLNAC	17	15.9%
CCNE	81	75.7%
Both	4	3.7%
Missing data	5	4.7%
Size of program		
Less than 49	20	18.7%
50-149	31	29.0%
149-299	22	20.6%
Greater than 300	27	25.2%
Missing data	7	6.5%
Percent FT students		Mean% = 36.49; SD = 32.62
Missing data	6	
Percent PT faculty		Mean% = 24.65; SD = 25.16
Missing data	9	
Years in existence		M = 20.58; SD = 12.69
Missing data	6	
Years online		M = 8.38; SD = 3.90
Missing data	6	
Method of online		
Web-enhanced	30	28.0%
Web-based	27	25.2%
Both	45	44.1%
Missing data	5	4.7%

Use of online		
Core only	10	9.3%
Specialty only	5	4.7%
Both	86	80.4%
Missing data	6	5.6%
Tuition status		
Decreased	1	.9%
Stayed the same	9	8.4%
Increased	93	86.9%
Missing data	4	3.7%
Recent year of evaluation		
0 - 2005	4	3.7%
2007-2012	92	96.0%
Missing data	11	10.3%

Data Preparation

Data was prepared using a SPSS data set. Initially a panel of 376 participants was entered into Qualtrics. Qualtrics assigned respondent identifications numbers that correlated with participant's email addresses and names. This process allowed for detailed tracking of respondents, accuracy, and ease of data entry. Using a function within Qualtrics, the majority of mail and paper survey data were entered into Qualtrics by the researcher. Raw data was downloaded from Qualtrics into SPSS 19.0. Mail and paper survey data not entered into Qualtrics was manually entered into the already established SPSS data set.

Coding for items 1-67 was automatically done by Qualtrics at the time of download into SPSS. A 4-point Likert scale had been used to score survey items. For items 1-53, "never" = 1; "1 or 2 times" = 2; "3 or 4 times" = 3; and "5 or more times" =

4. For items 54-67, “strongly disagree” = 1; “disagree” = 2; “agree” = 3; and “strongly agree” = 4. Background information did not utilize a response scale. Where multi-choice responses were required, Qualtrics simply coded 1, 2, or 3 as the responses appeared numerically.

This study used two separate frameworks for understanding the practices of online education program evaluation; Rovai’s adaptation of the CIPP model and UFE. As described in the conceptual framework there are three main dependent variables for this study; source of evaluation data, area of evaluation, and use of evaluation results. There are 10 key variables; five related to source, four related to area, and one related to use. Evaluation utilization had only one sub-scale which was used. In preparation for data analysis scales and sub-scales were created using the same identifiers as the main and key variables. Items were identified for each scale and sub-scale and were drawn as exactly identified from the two frameworks for this study. Each survey items 1-53 had dual assignment; to either one of the sub-scales for source of evaluation data or one of the sub-scales for evaluation area. Survey items 54-67 were used solely for the use sub-scale. See Appendix C for a more detailed description of the process of item identification for scales and sub-scales. The distribution and reliability of these scales and sub-scales are presented in Table 3.13.

Recodes

A number of recodes were performed. First, data analysis for some research questions utilized additive scores of key sub-scales. This required creating new variables using SPSS. As noted, the study’s main variables were (a) source of evaluation data with

sub-scales students, faculty, records, alumni, and employers, (b) area of evaluation with sub-scales input, process, output, impact, and (c) evaluating use which held the

Table 3.13

Distribution and Reliability of Scales and Sub-Scales

Scales/ Sub-Scales	Total Number Of Items	M	SD	Mean	Alpha Item Means
<u>Source</u>	53	130.51	36.00	2.46	.97
Students	13	36.39	9.18	2.80	.89
Faculty	9	21.88	8.26	2.43	.93
Records	20	46.85	15.13	2.34	.94
Alumni	6	14.62	4.93	2.43	.87
Employers	5	11.44	5.18	2.29	.93
<u>Area</u>	53	130.51	36.00	2.46	.97
Input	20	49.20	14.97	2.46	.93
Process	11	29.00	7.84	2.64	.86
Output	15	35.49	10.06	2.36	.89
Impact	7	17.12	6.78	2.45	.93
<u>Utilization</u>	14	47.25	9.49	3.38	.77
Use	14	47.25	9.49	3.38	.77

distinction of main and key variable/scale and sub-scale. A new variable was created for each sub-scale that represented additive scores of items belonging to each sub-scale. A new variable was also created for each sub-scale that represented additive scores divided

by the number of items within the sub-scale to obtain mean item means. This recoding now allowed use of these calculations when answering applicable research questions. See Table 3.14 for recodes of sub-scales for use in appropriate data analysis.

Table 3.14

Sub-Scales for Key Study Sub-Scales

Scales/ Sub-scales	Recoded Sub-scales (denoting additive items)	Recoded Sub-scales Item Means (additive items/ number items)
<u>Source</u>		
Students	Total_stu1	Total_stu2
Faculty	Total_facu1	Total_facu2
Records	Total_records1	Total_records2
Alumni	Total_alumni1	Total_alumni2
Employers	Total_employ1	Total_employ2
<u>Area</u>		
Input	Total_input1	Total_input2
Process	Total_process1	Total_process2
Output	Total_output1	Total_output2
Impact	Total_impact1	Total_impact2
<u>Utilization</u>		
Use	Total_use1	Total_use2

There were two demographic items that needed recoding. Item number 70 asked respondents about the size of their master's degree level nursing program. There were two very extreme values when compared to the other values reported; 5500 and 4000

students. As mentioned, the researcher did recheck survey entries for accuracy and found these numbers to be correct as recorded. It was decided to allow the values to remain as part of the data set. Because of these extreme values reported by very large programs the mean was not a very meaningful indicator of central tendency. Consequently, it will be better to look at the median value. For analytical purposes however, when observing program size, the non-parametric shape undercut most of the analysis that was desired. Programs were re-conceptualized as small (less than 49), medium (50-149), large (150-299), and very large (greater than 300). The variable was recoded as quartiles and renamed “quartsize”.

Item number 82 “in what year were you born?” was also recoded. Respondents were asked for a year. Using SPSS this item was recoded to the variable “age” for easier analysis by subtracting the reported year from the current year of 2012.

Survey data was further carefully and meticulously scrutinized for possible negative items, non-responses, incomplete surveys, missing data, and inaccurate entries. At the end of data cleaning a total of 107 surveys were determined eligible for use and data analysis.

Data Analysis

Research data was analyzed using SPSS 19 statistical software. In collaboration with the methodologist, statistical analysis was identified that best addressed the following research questions:

1. To what extent are schools of nursing systematically evaluating their online education activities at the master’s degree program level?
2. What are the sources of evaluation data?

3. What evaluation areas are being used?
4. To what extent are evaluation results utilized in schools of nursing?
5. To what extent do institutional and program characteristics affect evaluation
(a) data source; (b) area; and (c) utilization in schools of nursing online master's
degree level programs?

The first research question was addressed by constructing a frequency table. Means were calculated for each evaluation practice variable (source and area of evaluation) items. The item means were then rank ordered and compared in order to assess how schools are performing systematic program evaluation, for example, most common practices and least common practices.

Research question two was addressed by first calculating distributions and reliabilities of key variables (or scales and sub-scales) for source of evaluation data. This was done to assess internal validity of items. Next, using the new variables created in SPSS, additive item scores were calculated for the five source sub-scales. From the additive scores items means, standard deviations, and mean of means were calculated. Item means were assessed to determine commonality between sub-scale item means of the practice variable source of evaluation data. Mean of means were then rank ordered to determine which sources of evaluation data were used most and least often by schools of nursing in the respondent sample.

Research question three used the same analysis as question two for the key variable area of evaluation. Additive scores were calculated for the four area sub-scales.

Research question four was addressed by also calculating item frequencies, means, and percentages as in question one. Item means were rank ordered and compared

for the most common use practices and least common use practices of evaluation data. Similar analysis used in question two and three were also used. Distribution and reliability were calculated for the use items. Evaluating use scale did not contain sub-scales, but the same type of analysis of computing additive scale scores for use items was determined.

Research question five was addressed by examining the possible impact of program characteristics on the various evaluation practices; source and area of evaluation, and use of evaluation. Three predictor variables were identified from among program characteristics that could potentially impact evaluation practice and use; type of organization (public/private); size of the master's level degree program, and length of operation of online program in years. The objective was to examine potential relationships or correlation between these three organizational characteristics and sub-scale scores obtained during the analysis of question two, three, and four. This required a series of bivariate analyses to include correlation and t-test to look at relationships between variables. T-test was used to analyze the relationship between the sub-scale scores and the categorical variable, type of organization. The relationship between sub-scale scores and continuous variable years in operation was analyzed using Pearson's correlation analysis. The relationship between sub-scale scores and continuous variable size of online program was analyzed using Spearman's correlation.

Limitations

This study, as is the case with many current studies, suffered from a lower than desirable response rate. Merriam and Simpson (2000) stated that a low response rate affects the degree to which researchers are assured of a representative sample. At the

onset of this study the researcher had hoped for a high enough response rate of 60% or better that would allow generalizations from the data about the population of schools of nursing as a whole. In actuality, the study did not receive such a response rate. Despite repeated attempts and multiple methods the ultimate response rate was 31%. Inferences regarding statistical generalizability beyond this population must be made with caution.

The resulting 107 responses did allow the researcher to conduct the planned analysis and to learn from the findings. However, as stated it requires special caution not to generalize beyond the sample as it currently stands. Therefore, for those analyses that present means or compare item means, such as question one, the reader should use caution interpreting this. Furthermore, be aware no claims of generalization beyond the sample of respondents can be made. Consequently, in reporting the findings the researcher will be careful to refer only to the population.

However, data was collected regarding type of institution in order to compare respondents to non-respondents, just to see if there was a match in some way. Is the responding sample like the non-respondents in any way? As previously reported 54.2% of respondents were public organizations and 40.2% were private. When investigating this same variable for non-respondents (n=220), it was found that 57% were public organizations and 43% were private. Speculation can be made that when making this comparison, and as far as type of institution is concerned, respondents and non-respondents are closely matched. See Table 3.15 for a summary of this data.

Table 3.15

Types of Schools of Nursing in the US with Master's Level Education

Participants	Public	Private
Respondents	54.2% (n=58)	40.2% (n=43)
Non-respondents	57.0% (n=126)	43.0% (n=94)

However, with data available for only one institutional variable for both sample respondents and non-respondents, strict generalizations of the data results cannot be applied to the population. Also, the researcher had no knowledge of participant eligibility for the study.

A comparison of size of master's program and length of years in operation of online programs would have being highly desirable from among study respondents versus non-respondent. However, current and accurate data was not readily available to the researcher without making some form of contact with the school of nursing.

A second limitation may have been in the decision to use mixed-modes surveys. Though this study wanted to enhance data quality by improving response rate with the mail and paper survey, using an additional mode may have introduced measurement error. Hopefully, using the same survey as downloaded from the Internet and similar wording nullified this happening.

CHAPTER 4

FINDINGS

The purpose of this chapter is to report findings as related to the five research questions. The purpose of the study was to investigate to what extent schools of nursing are currently practicing systematic program evaluation of online education and how are they utilizing the results. Five research questions guided this study. There will be a separate section below for each of the research questions.

1. To what extent are schools of nursing systematically evaluating their online education activities at the master's degree level?
2. What are the sources of evaluation data?
3. What areas of evaluation are being used?
4. To what extent are evaluation results utilized by schools of nursing?
5. To what extent do institutional and program characteristics affect evaluation
(a) data source; (b) area; and (c) utilization in schools of nursing online master's degree level programs?

Findings Related to Research Question 1

Research question one reads "To what extent are schools of nursing systematically evaluating their online education activities at the master's degree level?" Research question one focused on the frequency of evaluating practices for sources of evaluation data and areas of evaluation. Frequencies, item means and standard deviations, and rank order means were calculated and reported as findings. The directions for the

response scale asked for a time interval of within the past five years on a 1 (Never) to 4 (5 or more times) point Likert response scale. A detailed frequency table of all practice items can be found in Appendix K.

Analysis results revealed that all practices were being done, that is, no respondents reported zero for any response. For the response “never” practice items 8 and 7 tied at the lowest frequency of 3 or 2.8%. Practice item 41 was reported as having the highest frequency (68/63.6%) for the “never” response.

Only six respondents (5.6%) reported practicing item 8 “1 or 2 times” within the past five years, which represented the least reported in this category. The majority in this response category (41/38.3%) reported item 18 “collecting data from faculty about control over online course content” as most frequently performed “1 or 2 times” within the past five years.

Respondents reported most frequently performing “3 or 4 times” within the past five years practice items 9 and 10 (25/23.4%). The least reported (3/2.8%) in this category was item 47, item 27, and item 48.

Finally, 79 (73.8%) schools of nursing reported practice item 8 as being performed most frequently “5 or more times” within the past five years. Practice item 27 was reported least by nine (8.4%) respondents for “5 or more times” within the past 5 years.

Overall, evaluating practice items had a mean range of 3.64 – 1.59. The most commonly occurring practice, “collecting data from students about perceptions of teaching effectiveness”, had the highest mean ($M = 3.64$; $SD = .72$) and the highest

frequency and percentage of five or more times (79/73.8%). Only three (2.8%) respondents reported never performing this practice.

“Compiling and analyzing data from program records about policies hindering transferability of online course work” had the lowest mean ($M = 1.59$; $SD = .98$). Sixty-eight respondents or 63.6% reported never performing this practice of evaluation. Only 10 reported five or more times or 9.3%.

Table 4.1 outlines the *top ten practices* rank ordered by item means, reported by schools of nursing. The top-ten list represents the most frequently occurring evaluation practices for source of evaluation data and evaluation area. “Collecting data from students about perceptions of teacher effectiveness” holds the top position on the list as the most frequently used evaluation practice ($M = 3.64$; $SD = .72$). “Collecting data from alumni about skills obtained” ($M = 2.97$; $SD = 1.04$) and “collecting data from alumni about employment rates” ($M = 2.97$; $SD = .98$) means are tied for positions nine and ten. Furthermore, of the top ten practices, schools of nursing reported most often collecting data from students (5/10) as the primary source of evaluation data. Of the remaining top evaluation practices, schools of nursing reported collecting data from records (3/10) and collecting data from alumni (2/10) as primary sources of evaluation data. Collecting data from faculty and collecting data from employers were not represented in the top ten practices as frequently used evaluation sources.

Evaluating process is the most frequently occurring evaluation area practice reported by schools of nursing (items 8 and 7). “Collecting data from students about perceptions of teaching effectiveness” held the top position ($M = 3.64$; $SD = .72$). “Collecting data from students about satisfaction with online learning” was at number

two ($M = 3.35$; $SD = .87$). All evaluating area practices were represented in the top ten.

In assessing the area of evaluation most often practiced, schools of nursing reported evaluating output (4/10). Schools of nursing frequently obtained output data from records and alumni (items 32, 34, 45, 46). The remainder of the evaluating area practices is listed in the frequency in which they were reported: evaluating process (3/10; item 8, 7, and 9); evaluating input (2/10; item 3, 5), and least evaluating impact (1/10, item 43).

Table 4.1

Top Ten Rank Ordered Item Means for Evaluating Practice

Rank	Item Number/Item	Mean	SD
1	8. Collecting data from students about perceptions of teaching effectiveness	3.64	.72
2	7. Collecting data from students about satisfaction with online learning	3.35	.87
3	3. Collecting data from students about satisfaction with technology support for online learning	3.33	.94
4	5. Collecting data from students about adequacy of the technology learning system (Blackboard, Web-ct, eLive, etc.) used for online learning	3.27	.95
5	32. Compiling and analyzing data from program records about program enrollment rates	3.10	1.06
6	34. Compiling and analyzing data from program records about program graduation rates	3.09	1.04
7.5	45. Collecting data from alumni about satisfaction with the degree obtained	2.99	1.06
7.5	9. Collecting data from students about satisfaction with online library support and access	2.99	1.03
9.5	46. Collecting data from alumni about satisfaction with skills obtained	2.97	1.04

The *bottom ten items* represent evaluation practices occurring least frequently in schools of nursing by source of evaluation data and area of evaluation. These items had a mean range of 1.94 – 1.59. See table 4.2 for the bottom ten reported evaluation practices by their rank ordered means. For evaluation data source schools of nursing reported less frequently “collecting data from employers about job retention rates of graduates from online education program” (M = 1.94; SD = 1.19). Reported least frequently was “compiling and analyzing data from program records about policies hindering transferability of online coursework” (M = 1.59; SD = .97). From the bottom evaluation practices by source of data, schools of nursing reported least frequently obtaining evaluation data from records (6/10), alumni (2/10), students (1/10), and finally employers (1/10). Collecting data from faculty was not represented in the bottom ten practices.

All evaluating area practices were represented in the bottom ten. In assessing the area of evaluation less often practiced, schools of nursing reported evaluating output (5/10, 37, 38, 47, 48, and 41). The remainder of the evaluating area practices is listed in the frequency in which they were reported: evaluating input (3/10; item 42, 12, and 27); tied were items evaluating process (1/10; item 28) and impact (1/10, item 53).

Table 4.2

Bottom Ten Rank Ordered Item Means for Evaluating Practice

Rank	Item Number/Item	Mean	SD
44	53.Collecting data from employers about job retention rates of graduates from our online education program	1.94	1.19
45	37.Compiling and analyzing data from program records about faculty retention rates	1.84	1.04
46.5	28.Compiling and analyzing data from program records about comparisons between online versus traditional end-of-course evaluations	1.71	.94
46.5	42.Compiling and analyzing data from program records about faculty selection process to teach online	1.71	1.03
48.4	38.Compiling and analyzing data from program records about staff retention rates	1.67	.99
48.5	12.Collecting data from students about cost-benefit of online learning	1.67	1.03
50	47.Collecting data from alumni about reduction of learning needs	1.65	1.07
51	27.Compiling and analyzing data from program records about comparative cost effectiveness of online versus face-to-face instruction	1.62	.92
52	48.Collecting data from alumni about elimination of learning needs	1.60	1.01
53	41.Compiling and analyzing data from program records about policies hindering transferability of online coursework	1.59	.97

Findings Related to Question 2

“What are the sources of evaluation data?” is research question two. This question focused on what sources do schools of nursing report that they utilize most often for collecting evaluation data. The five source sub-scales included collecting data from (a) students, (b) faculty, (c) records, (d) alumni, and (e) employers. First, distributions and reliability were calculated for each sub-scale to evaluate item internal validity. Source scale consisted of 53 items which were rated on a 4-point Likert scale, from 1, “never”, to 4, “5 or more.” Cronbach’s alpha was consistently greater than .70 for the scale and sub-scales.

Next, additive mean scores were calculated for the five source sub-scales. From the additive mean scores standard deviations and mean of means were calculated. See Table 4.3 for distributions for sub-scales by source.

Table 4.3

Source Sub-Scales Mean Scores, SD, Mean of Means

Sources Sub-Scales	Number of Items	Mean Score	SD	Mean of Means
Students	96	36.38	9.18	2.80
Faculty	100	21.88	8.26	2.43
Records	86	46.84	15.13	2.34
Alumni	94	14.62	4.93	2.43
Employers	99	11.44	5.18	2.28

The last step in the analysis was to rank order the means of sub-scales. The most common source of evaluation data is from students with the number one ranked mean ($M = 2.80$). It was also common to collect evaluation data from faculty and alumni, means tied for second place ($M = 2.43$). Schools of nursing reported using employers for evaluation data least. See Table 4.4 for rank ordered means for sub-scales by source.

Table 4.4

Rank Ordered Means for Sub-Scales by Source

Source Sub-Scale	Mean of Means	Mean Rank
Students	2.80	1
Faculty	2.43	2.5
Alumni	2.43	2.5
Records	2.34	4
Employers	2.28	5

Findings Related to Question 3

Question three states “What are the areas of evaluation?” Research question three focuses on what areas of evaluation do schools of nursing report performing most often. The four area sub-scales included (a) evaluating input, (b) evaluating process, (c) evaluating output, (d) evaluating impact. First, distributions and reliability were calculated for each sub-scale to evaluate item internal validity. Area sub-scales consisted of the same 53 items which were rated on a 4-point Likert scale, from 1, “never”, to 4, “5 or more.” Cronbach’s alpha was consistently greater than .80 for each sub-scale.

Next, additive item scores were calculated for the four area sub-scales. From the additive mean scores, standard deviations and mean of means were calculated. See Table 4.5 for distributions for sub-scales by area.

Table 4.5

Area Sub-Scales Mean Scores, SD, Mean of Means

Area Sub-Scale	Number of Items	Mean Score	SD	Mean of Means
Input	90	49.20	14.97	2.46
Process	93	29.00	7.83	2.63
Output	83	35.49	10.06	2.36
Impact	99	17.12	6.78	2.45

The last step in the analysis was to rank order the mean of means sub-scales. There were no commonalities among evaluating areas. Respondents reported evaluation practices focused on process most frequently than any other area (M=2.63). This was followed by evaluating input (M=2.46), evaluating impact (M=2.45), and last by evaluating output (M=2.36). See Table 4.6 for rank ordered means for sub-scales by area.

Findings Related to Question 4

Question four reads “To what extent are evaluation results utilized in schools of nursing?” Research question four focused on evaluation data use and how schools of nursing report this use. Frequencies, item means and standard deviations, and rank order

Table 4.6

Rank Ordered Means for Sub-Scales by Area

Sub-Scale	Mean of Means	Mean Rank
Process	2.63	1
Input	2.46	2
Impact	2.45	3
Output	2.36	4

means were calculated and reported as findings. All findings from the 14 use items are reported and are representative of schools of nursing reporting agreement with use of evaluation data on a 1(strongly disagree) to 4 (strongly agree) point Likert response scale.

Frequency results indicated that all utilization activities were being done to some extent of agreement, that is, no respondents reported zero for any response. For the response “strongly disagree” item 65 (uses evaluation data for accountability to someone or something), item 66 (uses data to generate knowledge about their program), and item 56 (identifies questions to be answered by the evaluation process) tied at the lowest frequency of 1 (.9%). Item 61(reviews evaluation results with staff to identify best practices in online education) utilization activity was reported by schools of nursing as having the highest frequency (15/14.0%)) for the “strongly disagree” response.

Only two respondents (1.9%) for each item (65 and 56) reported “disagree”, which represented the least reported in this category. The majority in this response category (20/18.7%) reported performing use activity item 61 “my school of nursing reviews evaluation results with staff to identify best practices in online education”.

Respondents reported that they “agree” most often with use activity item 58, which was “my school of nursing keeps stakeholders informed of interim evaluation findings” (50/46.7%). The least reported in this category was item 65 which was “uses evaluation data for accountability to someone or something” (30/28.0%).

Finally, sixty-seven (63.6%) schools of nursing reported “strongly agree” to use activity item 65 (uses evaluation data for accountability to someone or something). Evaluation use activity item 58 (keeps stakeholders informed of interim evaluation findings) was reported as “strongly agree” by 32 (29.9%), making this the least occurring use activity with this response.

Item means for program evaluation utilization activities ranged from 3.63 – 2.76. When assessing item means schools of nursing reported strong agreement that their school of nursing uses evaluation data for accountability to someone or something ($M = 3.63$; $SD = .58$). They reported least that their school of nursing reviews evaluation results with staff to identify best practices in online education ($M = 2.76$; $SD = 1.01$). Commonality was seen with three items that schools of nursing reported as performing frequently. Those were item 66 ($M = 3.48$; $SD = .67$), item 62 ($M = 3.48$; $SD = .69$), and item 67 ($M = 3.48$; $SD = .66$).

In summary all ratings on utilization are high with little variation. See Table 4.7 for a summary of evaluating use activities ranked ordered by items means.

Table 4.7

Frequency Table – Evaluating Use Practices

Item	n	Mean	SD	Rank of means	Frequencies (n/%)			
					Strongly Disagree	Dis- agree	Agree	Strongly Agree
65.My school of nursing uses evaluation data for accountability to someone or something (e.g., an accrediting or certifying body, funding organizations)	98	3.63	.58	1	1/ .9%	2/ 1.9%	30/ 28.0%	67/ 62.6%
60.My school of nursing reviews evaluation results with faculty to identify best practices in online education	101	3.55	4.19	2	6/ 5.6%	14/ 13.1%	40/ 37.4%	42/ 39.3%
66.My school of nursing uses evaluation data to generate knowledge about our program (e.g., to identify patterns of effectiveness, trends in participation, general problems with online learning)	100	3.48	.67	3	1/ .9%	7/ 6.5%	36/ 33.6%	58/ 54.2%
62.My school of nursing uses evaluation data to monitor program performance at regular intervals	100	3.48	.69	3	2/ 1.9%	5/ 4.7%	37/ 34.6%	58/ 54.2%
67.My school of nursing uses evaluation data to make developmental changes to our program (e.g., alternative course sequencing, innovative delivery)	99	3.48	.66	3	2/ 1.9%	3/ 2.8%	41/ 38.3%	55/ 51.4%
56.My school of nursing identifies questions to be answered by the evaluation process	98	3.47	.59	6	1/ .9%	2/ 1.9%	46/ 43.0%	51/ 47.7%
63.My school of nursing uses evaluation data to improve (make better) our online program	100	3.43	.76	7	3/ 2.8%	8/ 7.5%	33/ 30.8%	58/ 54.2%

54.My school of nursing identifies stakeholders in the evaluation process	102	3.41	.68	8	2/ 1.9%	5/ 4.7%	45/ 42.1%	52/ 48.6%
57.My school of nursing uses evaluation methods that are appropriate for the questions identified (e.g., qualitative, quantitative, mixed methods)	101	3.36	.70	9	2/ 1.9%	7/ 6.5%	46/ 43.0%	48/ 44.9%
64.My school of nursing uses evaluation data to render judgment about our program (e.g., determine the overall merit, worth, significance, or value)	101	3.33	.80	10	4/ 3.7%	9/ 8.4%	39/ 36.4%	51/ 47.7%
59.My school of nursing decides how evaluation results are to be disseminated to the stakeholders	100	3.32	.73	11	3/ 2.8%	7/ 6.5%	46/ 43.0%	46/ 43.0%
55.My school of nursing assesses program readiness for evaluation	98	3.20	.74	12	2/ 1.9%	13/ 12.1%	48/ 44.9%	37/ 34.6%
58.My school of nursing keeps stakeholders informed of interim evaluation findings	101	3.07	.80	13	4/ 3.7%	17/ 15.9%	50/ 46.7%	32/ 29.9%
61.My school of nursing reviews evaluation results with staff to identify best practices in online education	98	2.76	1.01	14	15/ 14.0%	20/ 18.7%	39/ 36.4%	26/ 24.3%

Findings Related to Question 5

Research question five reads, “To what extent do institutional and program characteristics affect evaluation (a) data source; (b) area; and (c) utilization in schools of nursing online master’s degree level programs?” Question five examined the possible impact of three identified program characteristics on the various variables focusing on source of evaluation data, area of evaluation, and utilization. This series of bivariate correlation tests and t-test in which the recoded (additive scores) evaluating use scale and source and area sub-scales are the outcome variables and the predictor variable are the following:

- Type of organization; public or private
- Size of online master’s degree level program
- Years in operation of online education

Predictor Variable 1

In order to determine the impact of type of organization (a categorical variable), public/private, a series of independent samples t-test were conducted in order to compare means. The statistical analysis revealed not one of the 11 variables compared demonstrated significance as evidenced by $p > .05$. Public schools of nursing were no more apt to look at sources of evaluation data, utilize evaluation results, or focus on one area of evaluation compared to private schools of nursing.

Predictor Variable 2

In order to determine the impact of size of the master’s degree level program (continuous variable) simple regression analysis, Spearman’s correlation analysis was conducted. Statistical analysis revealed:

- Larger programs are more apt to seek student data ($r_s = .39$; $p < .05$)
- Larger programs are more apt to seek evaluation data from records ($r_s = .31$; $P = .04$).
- Larger programs are more apt to focus on evaluation input ($r_s = .26$; $P = .02$)
- Larger programs are more apt to focus on evaluation process ($r_s = .33$; $p < .05$)
- Larger programs are more apt to focus on evaluation output ($r_s = .31$; $P = .01$)

There was no distinction in larger programs seeking data from faculty ($P = .65$) and employer ($P = .11$). There was also no distinction in larger programs evaluation data utilization ($P = .63$). Finally, there was no distinction in larger programs focusing on evaluating impact ($P = .13$).

Predictor Variable 3

In order to determine the impact of years in operation (ordinal variable) simple regression analysis or Pearson correlation coefficient was conducted. Programs that have been longer in existence have a tendency to seek:

- Faculty as a source of evaluation data ($r = .21$; $P = .04$)
- Records as a source of evaluation data ($r = .30$; $P = .01$)
- Alumni as a source of evaluation data ($r = .24$; $P = .02$)
- Employers as a source of evaluation data ($r = .21$; $P = .04$)

There was no distinction in older programs seeking data from students, but in general the more mature programs are more apt to seek a variety of sources of evaluation data.

Programs that have been in existence longer also have a tendency to focus on evaluation area:

- Input ($r = .25$; $P = .02$)
- Output ($r = .34$; $P = .02$)
- Impact ($r = .24$; $P = .02$)

There was no distinction in older programs focusing on evaluation area process ($P = .05$), but in general older programs are more apt to focus on a variety of evaluation area practices. Also there was no significance in older programs being more apt to utilize evaluation results ($P = .09$).

CHAPTER 5

SUMMARY, CONCLUSIONS, AND IMPLICATIONS

Among their many responsibilities program administrators know that evaluation and assessment ought to be done, but the purpose of this study was to investigate the extent to which schools of nursing are currently practicing systematic program evaluation of online education and how are they utilizing the results. Five research questions guided this study.

1. To what extent are schools of nursing systematically evaluating their online education activities at the master's degree level?
2. What are the sources of evaluation data?
3. What are the areas of evaluation?
4. To what extent are evaluation results utilized by schools of nursing?
5. To what extent do institutional and program characteristics affect evaluation
(a) data source; (b) area; and (c) utilization in schools of nursing online master's degree level programs?

This final chapter will present (a) a summary of the study including findings; (b) conclusions with discussion; (c) implications for practice; and (d) suggestions and recommendations for future research.

Summary of the Study

Data for this study was gathered from schools of nursing with master's level online nursing education that were either accredited by the National League for Nursing

Accrediting Commission, the Collegiate Commission on Nursing Education, or both. The online nursing education program was either web-enhanced, web-based, or both. This study concentrated on identifying what systematic program evaluation practices and utilization activities were being conducted by schools of nursing.

The conceptual framework for this study was drawn from two program evaluation models; Rovai's (2003) adaptation of the CIPP (context, input, process, product) program evaluation model adapted for online education and Utilization-Focused Evaluation (UTE). This study required development of an original instrument in order to address the five research questions. The item pool was extracted from these two program evaluation models as well as from the evaluation literature. The item refinement process took place over several months again relying on the literature, but also on an expert panel of nurse educators and results from the pilot study. The final instrument for the main study contained 84 items; 53 practice items pertaining to evaluation data sources and areas of evaluation; 14 activity items pertaining to utilization of evaluation data; and 17 items pertaining to program characteristics. The survey was designed to gather evaluation data from nurse educators, but can also be used by any profession or discipline with online degree programs.

The sample used for this study was a population sample of 383 schools of nursing with master's level degree online education. Data collection utilized mixed-modes; Internet and mail and paper surveys. Survey design and distribution closely followed the tailored design method developed by Dillman, Smyth, and Christian (2009) for Internet, mail, and mixed-mode surveys. One hundred-fifty two) surveys were returned; 68 by Internet and 84 by mail and paper. Data was downloaded from the

Internet survey platform into SPSS 19 for statistical analysis. After data cleaning 107 (31%) surveys were determined eligible for use and analysis.

Statistical analysis of the research questions included (a) simple frequencies, (b) mean and standard deviation, (c) t-test, and (d) bivariate analysis using correlation. A summary of the research findings is presented in the next section.

Summary of Findings

Findings from the research data focused on five main areas; (a) practices of systematic program evaluation; (b) which sources were being used most often for evaluation data; (c) what area of evaluation was being used most often; (d) how are schools of nursing utilizing the data; and (e) characteristics of programs practicing and utilizing program evaluation.

There were 53 survey items that measured evaluation practice. Data analysis revealed that all evaluation practices were being performed, that is, not one respondent reported never performing a practice. Overall, among the participants, evaluating practice items had a mean range of 3.64-1.59. The most commonly reported evaluation practice was “collecting data from students about perceptions of teaching effectiveness” ($M = 3.64$; $SD = .72$) with the highest frequency and percentage (79/73.8%).

The least commonly reported evaluation practice was “compiling and analyzing data from program records about policies hindering transferability of online course work” ($M = 1.59$; $SD = .98$). Some schools of nursing may question the quality of online education or perceive online programs as having too lenient teaching and learning standards or a less rigorous curriculum. Schools of nursing may institute policies that hinder a seamless transfer of credit hours from online programs. Sixty-eight respondents

(63%) reported never performing this practice. However, written comments on mail and paper surveys regarding this practice indicated confusion evidenced by question marks and “not sure”.

In assessing the *top ten ranked* item means and frequencies, evaluating process was the most frequently occurring evaluation area practice reported by respondents. The item “collecting data from students about perceptions of teaching effectiveness” was ranked number one ($M = 3.64$; $SD = .72$). Ranked item mean number two also dealt with process, “collecting data from students about satisfaction with online learning” ($M = 3.35$; $SD = .87$). For source of evaluation data, schools of nursing reported collecting data from students most often (5/10).

In assessing the *bottom ten ranked* item means, evaluating output was reported as least occurring in frequency from the bottom ten ranked items. There were five items out of the bottom ten that measured evaluating output. For evaluation data source respondents reported less frequently collecting data from records (6/10).

The second major finding revealed that the most common source of evaluation data was from students ($M = 2.80$). According to Rovai’s (2003) adaptation of the CIPP program evaluation model for online education programs there are five possible sources of evaluation data; (a) students, (b) faculty, (c) institutional records, (d) alumni, and (e) employers. Additionally, study respondents reported using employers for evaluation data the least. Faculty and alumni ranked means tied for second place and records rank mean was fourth.

In a third set of findings respondents reported evaluating process most frequently from among the four evaluation areas ($M = 2.45$). The four evaluation areas included (a) evaluating input, (b) evaluating process, (c) evaluating output, and (d) evaluating impact.

The fourth finding revealed little variation in frequency, means, and rank ordered means among evaluation utilization activities. There were 14 items used to measure the process of program evaluation utilization. These items were developed using the UFE model developed by Patton (2008). More specifically, UTE's definition, premises, and purposes were used. The item means ranged from the highest at 3.63 to the lowest at 2.76. Frequency results showed that all utilization activities were being performed to some extent.

The use item with the highest rank ordered mean was, "my school of nursing uses evaluation data for accountability to someone or something" ($M = 3.63$; $SD = .58$). Frequencies for strongly agree were high (67/62.6%). The use item with the lowest rank ordered mean was "my school of nursing reviews evaluation results with staff to identify best practices in online education" ($M = 2.76$; $SD = 1.01$).

In the fifth major finding, it was found that in general older online nursing education programs at the master's degree level are more apt to seek a variety of sources of evaluation data and focus on a variety of areas for evaluation. Three identified online program characteristics were chosen to examine possible impacts on the various variables focusing on source of evaluation data, area of evaluation, and utilization of evaluation.

These three online program characteristics were:

- Type of organization; public or private
- Size of online master's degree level program

- Years in operation

Public schools of nursing were no more apt to look at sources of evaluation data, utilize evaluation results, or focus on one area of evaluation than private schools of nursing.

Larger online master's degree level programs were more apt to use students and records as sources of data. Larger programs are also more apt to focus on evaluation input and process.

Conclusions and Discussion

Based on the findings, four conclusions can be extracted from this study. Each conclusion will be discussed in a separate section below.

Conclusion 1: Systematic program evaluation is reported to be a widespread practice in schools of nursing.

To *some degree of frequency* all program evaluation practices and use activities are being performed in schools of nursing, as reported by participants of this study.

Taking a comprehensive look at evaluation practices for online nursing education at the graduate level, this study showed a wide range of evaluation practices being done, particularly by schools of nursing with more experience in online education.

This conclusion is somewhat surprising since this is not the state currently reflected by the literature, which has, to date, taken a more piecemeal approach to systematic program evaluation. For instance, Little, Passmore, and Schullo (2006) researched nursing students' satisfaction with the introduction of new technology into an existing online course. Billings, Connors, and Skiba (2001, 2005) looked at practices that focused on outcome evaluation to include access, convenience, connectedness, proficiency with technology, and technology using students as a source. In the same

study, the researchers also looked at educational practices of online learners and the use of technology.

More specifically, in literature that met criteria for systematic program evaluation, measurement of evaluation practices surrounding sources of data, focus of evaluation, and stakeholder involvement were not comprehensive. Avery, Cohen, and Walker (2008) collected evaluation data from faculty on course objectives being met, the need for technical support, support for diverse learning styles, and faculty interaction online. Ali, Hodson-Carlton, and Ryan (2002) collected data from students and faculty on course outcomes being met, student satisfaction with online course ware and design, and student satisfaction with faculty participation and feedback.

In addition, it is noteworthy to compare the evaluation practices used in research by Lindsay, Jeffrey, and Singh (2009) who also used the original CIPP model. Lindsay et al. (2009) reported collecting data from faculty and practices related to nurse educator's needs in order to implement online education. These practices included workload, initial technology training, and ongoing technology support. However, Lindsay et al. (2009) did not provide for full theoretical coverage of the CIPP framework and therefore did not fully capture what was being done in program evaluation. Avery, Cohen, and Walker (2009) used UFE to guide utilization of evaluation data, but only shared the data with faculty.

In addition to the more comprehensive nature of the practices queried in this study, there are several other explanations for the difference in this study and those reported in the literature. With new programming, such as online teaching and learning, there is greater need for determining the formative and summative program features.

Further, perhaps the current social context that calls for greater accountability is a keen motivator. Additionally, what schools of nursing are doing in terms of program evaluation activities and outcomes, was better captured by items in this survey that measured practices of evaluation by data source, evaluation area, and evaluation utilization.

It may be conjectured that schools of nursing doing program evaluation well participated in this study and schools of nursing not doing so well did not participate. This conjecture introduces the notion of nonresponse error defined by Dillman, Smyth, and Christian (2009) as when participants selected for a survey, in this case all schools of nursing with online master's level degree programs, who do not respond are somehow different in a way that is important to the study from those who do respond. Furthermore, given the low response rate for this study one might suspect and hypothesize that schools of nursing that were doing the best job of evaluating their program were more apt to respond than those that were doing little or none. This could have made for an overstatement of the level of program evaluation activity that is really going on in schools of nursing. However, using the findings from this study, schools of nursing who wish to establish or refine their evaluation systems will be able to see what is being done by schools doing program evaluation well.

Finally, there is always the possibility of a lack of evaluator knowledge or competency to address the survey items pertaining to practices and use of evaluation data. The evaluation practices measured by this study, although comprehensive of systematic program evaluation, may have been responded to by participants with limited knowledge of the process. Therefore, this study and the literature reports schools of nursing

performing evaluation practices; in actuality they may not be doing so in a systematic way as defined by Patton (2008) for this study.

Conclusion 2: There is a discrepancy between the espoused utilization of evaluation data and reported practices related to the predominant source of evaluation data and primary focus of evaluation.

This study provides deeper, more nuanced insight into utilization. A noteworthy discrepancy was discovered when reviewing the findings in that the majority (63%) of the respondents reporting the use of evaluation data for accountability to someone or something (accrediting/certifying body, funding agency) and with the highest agreement that this was being done as part of utilization of evaluation ($M = 3.63$; $SD = .58$)).

However, respondents reported the most common source of evaluation data was from students and the primary emphasis was placed on evaluating process.

Patton (2008) stated that the purpose of evaluation findings for accountability “demonstrates that resources are well-managed and efficiently attain desired results” (p. 140). Schools of nursing are very familiar with having to demonstrate performance through outcome measurement, monitoring, maintaining accreditation and certification, and response to auditors. Primary stakeholders are funding authorities, certifying bodies, administrators of programs, data managers, employers, and faculty (Patton, 2008, Rovai, 2003). Student achievement records and faculty would be primary sources of evaluation data to determine performance as far as goals and targets being met, indicators showing improvement, resource allocation, staff qualifications, eligible participants to programs, and the use of quality improvement mechanisms.

While components of process evaluation examine accountability, this is not the primary emphasis. An evaluation that focuses on process examines more of what is happening within the program and what should be happening, for example, teacher effectiveness in education. Rovai (2003) agreed that part of process evaluation also addresses the efficiency of the teaching and learning process.

If the purpose of evaluation was for accountability, more output and impact evaluation would have been expected. Output evaluation seeks to determine the immediate or direct effects of the program, to include assessing graduation rates and making sure program objectives are met. Impact or outcome evaluation is more concerned with longer-term results and program effectiveness at the societal level; more in line with accountability of organizations to something or someone.

In Lindsay, Jeffery, and Singh's (2009) study that examined a new online master's level degree nursing program with an emphasis on accountability to stakeholders, it was reported that faculty were the major stakeholders. Lindsay et al. (2009) used accountability indications from the CIPP program evaluation model for process related to program implementation and product related to program outcomes to focus the evaluation. This study demonstrated the need to not only focus on evaluating process but to also look at evaluating impact or outcomes when looking at accountability. Also in this study the researchers identified the need to include other stakeholders as sources of evaluation data, to include faculty, students, administrative partners, and employers; thus verifying the importance of including all key stakeholders.

There may be pragmatic reasons for using process evaluation to suffice as accountability measures. It takes greater human and financial resources to do outcome

evaluation. Also, evaluator competencies are not usually part of a nurse educator's skill set and more sophisticated evaluation experience and skill sets are needed to conduct longer-term outcome and impact assessment.

Conclusion 3: Two major stakeholder groups, employers who hire program graduates and staff members who help administer and implement the program, are typically not included in program evaluation practices and utilization activities.

Patton, Grimes, Guthrie, Brennen, French, and Blyth (1975) in earlier research found that a major impact on evaluation utilization was getting decision makers and stakeholders interested, committed, and involved in the evaluation process. The first premise of UFE is intended use by the intended user (Patton, 2008). Employers and staff have an important role in evaluation especially as it relates to evaluating input and program outcomes or impact.

This study revealed that respondents reported using employers least as a source for evaluation data. This is not surprising, since the literature review did not support the use of employers as sources of data or as major stakeholders in program evaluation. Accessing employer data is difficult. However, employers are a major source of information regarding the impact of the program at a local and societal level and for program outcomes (Rovai, 2003). Data to be obtained from employers could include:

- Employer satisfaction with graduates from online programs.
- Benefits of the online program to the employer.
- Changes in learner and employee job performance.
- Learner and employee performance on the job in terms of evaluation.
- Return on investment when financing employee education.

Also, when employers are not included as major stakeholders in program evaluation, input evaluation is affected. Decisions regarding student needs, employer needs, organizational needs, and even societal needs remain unknown and do not factor into program development and improvement.

Reviewing evaluation results with staff was noted least among respondents for utilization activities. This has important implications for program development and improvement. Staff functions as administrators of programs, data managers and analyzers, and as trainers and subject matter experts, particularly as it relates to institutional technology. In a study by Wilson (2001) it was found that in many instances faculty are willing to use online education technologies but needed training and the expertise of technology staff. Therefore, excluding part of the development and implementation team from knowing about and utilizing evaluation feedback renders the evaluation process incomplete and ultimately devalued.

Additionally, staff members themselves could be sources of helpful evaluation data if their perception of program outcomes were garnered. Further, training programs and staff in technology should be periodically evaluated to determine needs for improvement. Also, evaluation data should seek to determine how staff attitudes, behaviors, and goals changed because of a program.

Conclusion 4: The use of two program evaluation approaches to frame and conceptualize the study resulted in a comprehensive and coherent measurement of evaluation practice and evaluation utilization being performed in schools of nursing.

The new conceptualization model combining CIPP and UFE developed for this study provides a major contribution to the field of evaluation in general and specifically

to online education. The combination model allows schools of nursing to achieve the ability to assess evaluation practices comprehensively.

An initial search of the program evaluation literature quickly revealed that in order to fully measure evaluation practices and utilization activities there would be a need for two program evaluation models to guide this study. These two models individually emphasize two different approaches to program evaluation. The CIPP model emphasizes decision making and accountability of managers. Since this research study was interested in evaluation practices of online education, Rovai's (2003) adaptation of the CIPP model for program evaluation of online education was essential. UFE emphasizes how important stakeholder involvement is for utilization of evaluation data (Fitzpatrick, Sanders, & Worthen, 2011). CIPP and UFE models have been individually empirically tested and validated in research and meets professional evaluation standards set by the Joint Committee on Standards for Educational Evaluation (Yarbrough, Shulha, Hopson, & Caruthers, 2011). Conceptually combining the two models facilitated more careful consideration of systematic program evaluation practices and use activities.

Literature supports evaluation research based on models from professional program evaluation literature that is comprehensive and uses a systematic approach to data collection and analysis (Ruhe & Zumbo, 2009). All five research studies that met criteria for systematic program evaluation and that were included in the literature review utilized program evaluation models to frame their research. Two were traditional models, CIPP and UFE. Two were distance learning models; the EDUCAUSE framework and Chickering and Gamson's *Seven Principles of Good Education*. Limitations of distance learning evaluation models include their lack of rigor and the tendency to rely on self-

reports and qualitative data (Rovai, 2003). Also, as already noted, they are not comprehensive enough in their approach to evaluating online education and utilization of the data.

Avery, Cohen, and Walker (2008) also used a combination of evaluation models to develop an instrument and guide their evaluation study. The researchers used Chickering and Ehrmann's (1996) education principles adapted to technology, Billings's (2000) framework for evaluation quality of online education, and UFE (Patton, 2008) to guide data analysis. Though the researchers did not share their instrument they did describe it. The tool was developed to be used by faculty and others to evaluate online courses. However, from the researcher's description, the instrument did not make full use of many of the evaluation practices and utilization activities outlined by the evaluation approaches or those used in this research study. Therefore, the comprehensiveness of the tool is questionable in its theoretical coverage. The instrument developed for this study is comprehensive and the full potential of using multiple evaluation approaches is depicted in the combination of the CIPP and UFE models.

Implications for Practice

The aim of this study was to contribute to the program evaluation literature by examining the extent of systematic program evaluation of online nursing education in schools of nursing. This study plays a part in advancing a science of nursing education (Lindsay et al. 2009). Schools of nursing assessing evaluation practices and use activities from various areas and from multiple sources can use the information to inform curriculum, pedagogy, and outcomes of online nursing education.

Specifically, this study has important practical implications that extend to overall program administration, program development, capacity development, and development of professional online education. These are developed below.

Strengthening Evaluation as Part of Overall Program Administration and Program Development

A major implication of this study has to do with overall program development and administration. This study offers reassurance and confirmation that state of the art program evaluation practices and use activities are being done in schools of nursing that responded to this survey. Program developers and administrators can determine if the design and delivery of the program was effective and whether the outcomes were met (Caffarella, 2002). Therefore, schools of nursing who wish to establish or refine their evaluation process, now have models from schools that are doing extensive program evaluation. Accountability partners and other stakeholders can be assured that evaluation data is available for decision making and quality improvement. Faculty and students are made aware of teaching and learning effectiveness. Administrators will know if learning resources lead to positive or negative educational outcomes and resultant effects on retention and graduation rates. Finally, evaluation experts outside of nursing, who may interact with nursing programs, will have a sense of what matters to the profession. This makes visible the best comprehensive practices currently being done.

This study also suggests a need to aggressively seek broader stakeholder input, particularly alumni and employers, both in the planning of the evaluation to heighten their commitment to utilization as well as their being sources of evaluation data. While it may be more difficult to obtain evaluation data from alumni and employers than from

students, this should not excuse or deter schools of nursing from aggressively seeking evaluation data from these important stakeholder groups.

Furthermore, findings from this study revealed the lack of staff member involvement in evaluation data use. Program evaluation is strengthened when evaluation results are shared with all stakeholders. Can this be alleviated by involving higher education accrediting bodies? How much of evaluation utilization could be or should be driven by mandating this as part of the accrediting process? If so, one recommendation could be to actively involve accrediting bodies in promoting evaluation practice and data use. By including criteria and requirements that would promote evaluation utilization, more use activities could be assured.

Developing Evaluation and Online Instructional Capacity

Another important area to consider as an implication for practice has to do with capacity development within schools of nursing or other educational organizations without online programs. A significant incidental finding from this study is that of the schools of nursing responding to the survey, 39 (25%) did not provide online education at the master's degree level. This was somewhat surprising given the current popularity of online programs and the advances in educational technology. This anecdotal finding raises several questions around building capacity for an online program. Are technology training resources needed? Is cost a barrier? Is there resistance from stakeholders such as faculty and staff? Are students satisfied with online learning and willing to pay extra? Is there positive feedback obtainable for stakeholder's use? From this study, schools of nursing will be able to see what is being done and what matters most to respondents.

Furthermore, schools of nursing operating with limited evaluation and online instructional capacity may wish to build upon what they have. Schools of nursing can broaden their capacity in program evaluation design and execution. This includes broader stakeholder involvement beyond easily assessable evaluation data from students to alumni and employers. Schools of nursing can emphasize appropriate areas of evaluation that can easily become marginalized due to accessibility, such as outcome and impact evaluation.

Developing Professional Online Education

There are also implications for development of professional online education. One implication can be seen with the use of the survey instrument developed for this study. The survey instrument provides the adult education profession as well as the field of program evaluation with a valid and reliable tool to perform systematic program evaluation of online education. The instrument can be used in its entirety by administrators and other stakeholders as a guide for what to consider in developing new online programs.

It could also be used in a formative sense for program improvement, as well as a summative evaluation for decision making regarding quality in online education that may impact program outcome and impact. Furthermore, the instrument has the potential to be used for evaluation of specific content areas by source of data and focus of evaluation. Examples may include the use of the instrument in support of faculty peer review or to revise student course evaluations. Finally, evaluation data utilization can also be promoted by using this instrument. Program strengths and weaknesses can readily be identified in order to improve stakeholder involvement and use of evaluation data.

Implications for Future Research

This study raises areas and topics for future investigations into systematic program evaluation of online education. A foundation has been laid for the advancement of a conceptual framework using a combination program evaluation model; a model that was also utilized to develop the instrument used for this study. Four main areas are highlighted for future research.

The first implication has to do with the continued testing of the framework used in this study. The subsequent conceptual framework developed for this study has its' foundation in a combination model based on CIPP and UFE models of program evaluation. This conceptualization offers future researchers a visual to structuring or framing their research on program evaluation. Furthermore, future researchers may want to test or modify the conceptual framework based on their perceptions of CIPP and UFE. This will help future researches determine which models are most useful and valuable in program planning and evaluation of online education. This not only contributes to evidence-based educational practices, but theory development and best practices in online education evaluation (Horne & Sandmann, 2012).

Additionally, this study empirically tested the combination of two program evaluation models used to conceptualize the study; a combination model. The components, premises, and guidelines of the combination model revealed its' usefulness and value in survey development, guiding data analysis, and interpreting data findings. The developed survey instrument performed well. Researchers can be assured of a quality instrument, with measurement precision and accuracy, that has been tested for its' validity and reliability.

The second implication is about data collection. Although the survey instrument had validity and was carefully developed and procedures for data collection were meticulously executed, the study is still left with a less than desirable response rate that does not allow for meaningful generalization. Another research study (Hudson-Gallogly, 2012), also targeting schools of nursing, produced similar response rate results when using the Internet. Researchers must face the reality that despite the convenience of Internet survey, if the response rate suffers to the point that valid knowledge claims cannot be made, relying on this paradigm alone must be reconsidered.

Future researchers, especially with this population, should consider using mail and paper survey or mixed-modes. On assessing the response rate after the first email contact, one of the committee members suggested mail and paper for the final reminder. Approval from the Institutional Review Board (IRB) was sought and obtained for adding a mail and paper survey as another method of data collection. The results were dramatically better. A final recommendation for a less than optimal response rate would be to consider extending this same study but opting to use a smaller sample for a telephone survey.

A third set of implications addresses further study of certain variables. Key variables used in this study were carefully considered and fastidiously defined, however the items measured rather broad concepts, such as, teaching effectiveness and employer satisfaction with graduates. From a utilization of evaluation data standpoint, it would be useful to further investigate those items that scored a low rank ordered mean and low frequency of being done. For example, a quantitative approach to why schools of nursing

are not sharing evaluation data with staff can provide meaningful insight into attitudes and values of administrators and faculty.

Finally, this study also raises questions for future study through qualitative inquiry. Qualitative research studies can provide for more focused and in-depth insight into program evaluation. A case study approach can answer questions regarding stakeholder's perspectives and perceptions of evaluation utilization at their particular institution. Case study can also study those institutions which optimize employer and faculty input into utilization of evaluation data. Furthermore, using case study can explore how institutions do outcome and impact evaluations of online education.

Program evaluation research using mixed qualitative methodologies could result in a comprehensive, holistic study. For example, case study and interviews can better ascertain stakeholder values that may underline a program. In addition, using interviews and focus groups future researchers can investigate an online program's reputation or assess the value added of auxiliary support such as library or one-stop-registration.

Summary

Planning any educational program without planning for the evaluation of that program is to devalue the human experience of teaching and learning. By not evaluating the worth, value, and merit of online education program administrators and stakeholders are in essence deciding not to make decisions about revision and improvement of a product that could potentially impact society as a whole. Online teaching and learning has over time evolved, matured, and is no longer considered marginalized. Furthermore, online education is here and is only going to expand, as evidenced by the movement to

massive open online courses or MOOCs, which is the latest uproar in free online education (Carey, 2012).

That is why this study is so important. This study took a comprehensive look at evaluation practices and evaluation data use activities for online education. Also, this study developed a new conceptualization model combining CIPP and UFE to find out to what extent these practices and activities are being done in online education at the master's degree level. Finally, and the greatest value of this study, is to provide online education programs with the necessary tools (conceptualized combination model and instrument) to achieve the ability to assess evaluation practices comprehensively.

Online teaching and learning is constantly changing with the development of new and innovative technologies and course delivery structures (Ruhe & Zumbo, 2009). There is also increasing stakeholder expectations and increasing competition as more education institutions and organizations enter the online education market. Ruhe and Zumbo (2009) stated that “this context creates an ongoing need for continuous improvement, course updates, and blending of new technologies and pedagogies “(p, vi). Systematic program evaluation finds its' importance in this context.

REFERENCES

- Ali, N.S., Hodson-Carlton, K., & Ryan, M. (2002). Web-based professional education for advanced practice nursing: A consumer guide for program selection. *The Journal of Continuing Education in Nursing*, 33(1), 33-8. doi: 103153511.
- Alkin, M., & Taut, S. (2003). Unbundling evaluation use. *Studies in Educational Evaluation*, 29, 1-12.
- American Association of Colleges of Nursing. (1999, July). *Distance technology in nursing education*. Retrieved from <http://apps.aacn.nche.edu/Publications/WhitePapers/whitepaper.htm>
- Avery, M., Cohen, B., & Walker, J.D. (2008). Evaluation of an online graduate nursing curriculum: Examining standards of quality. *International Journal of Nursing Education Scholarship*, 5(1), 1-17. Retrieved from <http://ezproxy.gsu.edu:2048/login?url=http://ezproxy.gsu.edu:4518/login.aspx?direct=true&db=rzh&AN=2010184724&site=ehost-live>
- Bangert, A. (2006). The development of an instrument for assessing online teaching effectiveness. *Journal of Educational Computing Research*, 35(3), 227-244. Retrieved from <http://ezproxy.gsu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=EJ757189&site=ehost-live>
- Bangert, A., & Easterby, L. (2008). Designing and delivering effective online nursing

- courses with the evolve electronic classroom. *CIN: Computers, Informatics, Nursing*, 26(2), 99-105. Retrieved from <http://ezproxy.gsu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=2010189414&site=ehost-live>
- Beta-Jones, B., & Avery, M. (2004). Teaching pharmacology to graduate nursing students: Evaluation and comparison of web-based and face-to-face methods. *Journal of Nursing Education*, 43(4), 185-189.
- Bevil, C. (1991). Program evaluation in nursing education: Creating a meaningful plan. In M. Garbin (Ed.), *Assessing educational outcomes* (pp. 53-67). New York, NY: National League for Nursing.
- Billings, D. (2000). A framework for assessing outcomes and practices in web-based courses in nursing. *Journal of Nursing Education*, 39(2), 60-67. Retrieved from <http://ezproxy.gsu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=2000025314&site=ehost-live>
- Billings, D., Connors, H., & Skiba, D. (2001). Benchmarking best practices in web-based nursing courses. *Advanced Nursing Science*, 23(3), 41-52.
- Billings, D., Skiba, D., & Connors, H. (2005). Best practices in web-based courses: Generational differences across undergraduate and graduate nursing students. *Journal of Professional Nursing*, 21(2), 126-133.
- Caffarella, R. S. (2002). *Planning programs for adult learners: A practical guide for educators, trainers, and staff developers* (2nd ed.). San Francisco: Jossey-Bass.
- Carey, K. (2012). Into the future with MOOC's. *The Chronicle of Higher Education*, ,

n/a. Retrieved from

<http://ezproxy.gsu.edu/login?url=http://search.proquest.com/docview/1038136902?accountid=11226>

Carmichael, D. E. (2001, June). An educational evaluation of WebCT: A case study using the conversational framework. *Proceedings from the ED-MEDIA 2001 World Conference on Educational Multimedia, Hypermedia, & Telecommunications*.
<http://ezproxy.gsu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED466144&site=ehost-live>

Cervero, R. M., & Wilson, A. L. (2006). *Working the planning table: Negotiating democratically or adult, continuing, and workplace education*. San Francisco: Jossey Bass.

Chapman, D. D. (2006). Developing an evaluation plan for a fully online degree. *HRDOnline Submission*. Retrieved from
<http://www.westga.edu/~distance/ojdla/spring91/chapman91.pdf>

Chickering, A., & Ehrmann, S. (1996). Implementing the seven principles: Technology as lever. *AAHE Bulletin*. Retrieved from
<http://www.tltgroup.org/programs/seven.html>

Christensen, C. M., & Eyring, H. J. (2011). *The innovative university: Changing the DNA of higher education from the inside out*. San Francisco, CA: Jossey-Bass.

Christie, C. (2007). Reported influence of evaluation data on decision maker's actions: An empirical examination. *Journal of Evaluation*, 28(1). 8-25. doi: 10.1177/1098214006298065.

Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation

- utilization. *Review of Educational Research*, 56(3), 331-364. Retrieved from ERIC database.
- Cragg, B., Dunning, J., & Ellis, J. (2008). Teacher and student behaviors in face-to-face and on-line courses: Dealing with complex concepts. *Journal of Distance Education*, 22(3), 115-128.
- Daroszewski, E., Kinser, A., & Lloyd, S. (2004). Online, directed journaling in community health advanced practice nursing clinical education. *Journal of Nursing Education*, 43(4), 175-180.
- Dillman, D., A, Smyth, J. D., & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys: The tailored design method*. Hoboken: Wiley.
- Donaldson, S. I., Patton, M. Q., Fetterman, D. M., & Scriven, M. (2010). The 2009 Claremont debates: The promise and pitfalls of utilization-focused and empowerment evaluation. *Journal of Multidisciplinary Evaluation*, 6(13), 15-57. Retrieved from http://survey.ate.wmich.edu/jmde/index.php/jmde_1/article/view/260/250
- Edwards, P. (2005). Impact of technology on content and nature of teaching and learning. *Nursing Education Perspectives*, 26(6), 344-347.
- Fearing, A., & Riley, M. (2005). Graduate students' perceptions of online teaching and relationship to preferred learning styles. *MEDSURG Nursing*, 14(6), 383-389.
- Fitzpatrick, J., Sanders, J., & Worthen, B. (2011). *Program evaluation: Alternative approaches and practical guidelines* (4th ed.). Boston, MA: Pearson.
- Forester, J. (1989). *Planning in the face of power*. Berkeley: University of California Press.

- Graham, C., Cagiltay, K., Lim, B., Craner, J., & Duffy, T. M. (2001). Seven principles of effective teaching: A practical lens for evaluating online courses. *Technology Source*. Retrieved from http://technologysource.org/article/seven_principles_of_effective_teaching
- Green, K. (2009). *The campus computing project*. Retrieved from <http://www.campuscomputing.net/2009-campus-computing-survey>
- Hall, M., Daly, B., & Madigan, E. (2010). Use of anecdotal notes by clinical nursing faculty: A descriptive study. *Journal of Nursing Education*, 49(3), 156-159.
- Hoban, G., Neu, B., & Castle, S. R. (2002, April). *Assessment of student learning in an educational administration online program*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.
- Horne, E., & Sandmann, L. (2012). Current trends in systematic program evaluation of online graduate nursing education: An integrative literature review. *The Journal of Nursing Education*, 1(10), 570-576.
- Huckstadt, A., & Hayes, K. (2005). Evaluation of interactive online courses for advanced practice nurses. *Journal of the American Academy of Nurse Practitioners*, 17(3), 85-89.
- Hudson-Gallooly, K. (2012). *Exploring abortion education in undergraduate nursing programs in Georgia*. (Unpublished doctoral dissertation). University of Georgia, Athens.
- Jackson, L. A. (2000). Using the Internet for training and development. In E. E. Kossek

- & R. N. Block (Eds), *Managing human resources in the 21st century: From case concepts to strategic choice* (pp. 20.7 – 20.27). South-Western College Publishing.
- Johnson, R. B. (1998). Toward a theoretical model of evaluation utilization. *Evaluation and Program Planning*, 21, 93-110. doi: 10.1016/S0149-7189(97)00048-7.
- Johnson, K., Greenesid, L., Toal, S., A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on evaluation use: A review of the empirical literature from 1986 to 2005. *American Journal of Evaluation*, 30, 377-410. doi: 10.1177/1098214009341660.
- Joint Committee on Standards for Educational Evaluation. (1994). *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- King, J. (2003). The challenge of studying evaluation theory. In C. S. Christie (Eds.), *New directions for evaluation: No. 97. The practice theory relationship in evaluation* (pp.57-58). San Francisco: Jossey-Bass.
- Kirkhart, K. (2000). Reconceptualizing evaluation use: An integrated theory of influence. *New Directions for Evaluation*, 88, 5-22.
- Kirkpatrick, D., & Kirkpatrick, J. (2006). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler.
- Leasure, A., Davis, L., & Thievon, S. (2000). Comparison of student outcomes and preferences in a traditional vs. World Wide Web-based nursing research course. *Journal of Nursing Education*, 39(4), 149-154.
- Leners, D., Wilson, V., & Sitzman, K., (2007). Twenty-first century doctoral education:

- Online with a focus on nursing education. *Nursing Education Perspectives*, 28(6), 332-336.
- Lindsay, G., Jeffery, J., & Singh, M. (2009). Paradox of a graduate human science curriculum experienced online: A faculty perspective. *The Journal of Continuing Education in Nursing*, 40(4), 181-186.
- Little, B. B., Passmore, D., & Schullo, S. (2006). Using synchronous software in web-based nursing courses. *CIN: Computers, Informatics, Nursing*, 24(6), 317-325. Retrieved from <http://ezproxy.gsu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=cin20&AN=2009341674&site=ehost-live>
- Madaus, G. F., & Kellaghan, T. (2000). Models, metaphors and definitions in evaluation. In D. L. Stufflebeam, G. F. Madaus, & T. Kellaghan (Eds.), *Evaluation models: Viewpoints on education and human services evaluation* (2nd ed., pp. 19-32). Boston: Kluwer Academic.
- Merriam, S. B., & Simpson, E. L. (2000). *A guide to research for educators and trainers of adults*. Malabar, FA: Krieger.
- Mills, A. (2007). Evaluation of online and on-site options for master's degree and post-master's certificate programs. *Nurse Educator*, 32(2), 73-77.
- Moore, M., & Kearsley, G. (1996). *Distance education: A systems approach*. Belmont, CA: Wadsworth/ITP.
- National League for Nursing Accrediting Commission, Inc. (2008). *NLNAC 2008 standards and criteria*. Retrieved from <http://www.nhac.org/manuals/SC2008.htm>
- Nauman, L. H. (2006). Creating new futures in nursing education: Envisioning the

- evolution of e-nursing education. *Nursing Education Perspectives*, 27(1), 12-15.
- Newman, A. (2003). Measuring success in web-based distance learning. *EDUCAUSE Center for Applied Research Bulletin*, 4, 1-11.
- O'Neil, C. A., & Fisher, C. A. (2008). Should I take this course online? *Journal of Nursing Education*, 47(2), 53-58.
- O'Neil, C. A., Fisher, C. A., & Newbold, S. K. (2004). *Developing an online course: Best practices for nurse education*. Springer Publishing Company: NY.
- Patton, M. Q. (2008). *Utilization-focused evaluation*. Los Angeles: Sage.
- Patton, M. Q., Grimes, P. S., Guthrie, K. M., Brennan, N. J., French, B. D., & Blyth, D. A. (1975). *In search of impact: An analysis of the utilization of federal health evaluation research*. Retrieved from ERIC database.
- Potempa, K., Stanley, J., Davis, B., Miller, K., Hassett, M., & Pepicello, S. (2001). Survey of distance technology use in AACN member schools. *Journal of Professional Nursing*, 17(1), 7-13.
- Preskill, H., & Caracelli, V. (1997). Current and developing conceptions of use: Evaluation use TIG survey results. *Evaluation Practice*, 18, 209-225. Retrieved from ERIC database.
- Ritter, L. A., & Sue, V. M. (2007). Introduction to using online surveys. *New Directions for Evaluation*, 115, 14. doi: 10.1002/ev.230.
- Roblyer, M. D., & Ekhaml, L. (2000). How interactive are YOUR distance courses? A rubric for assessing interaction in distance learning. *Online Journal of Distance Learning Administration*, 3(2).
- Rossi, P., Lipsey, M. W., & Freeman, H. E. (2004). *Evaluation: A systematic approach*.

Thousand Oaks, CA: Sage.

- Rovai, A. (2003). A practical framework for evaluating online distance education programs. *The Internet and Higher Education*, 6, 109-124.
- Ruhe, V., & Zumbo, B. (2009). *Evaluation in distance education and e-learning: the unfolding model*. New York: Guilford Press.
- Russell, T. L. (1999). *The no significant difference phenomenon*. Raleigh: North Carolina State University.
- Scriven, M. (1991). *Evaluation thesaurus*. Newberry Park, CA: Sage.
- Seiler, K., & Billings, D. (2004). Student experiences in web-based nursing courses: Benchmarking best practices. *International Journal of Nursing Education Scholarship*, 1(1), 1-12.
- Singh, M. (2004). Evaluation framework for nursing education programs: Application of the CIPP model. *International Journal of Nursing Education Scholarship*, 1(1), 1-4.
- Singh, M., Jeffrey, J., & Lindsay, G. (2008). Isolated learning for caring professionals: Advantages and challenges. *The International Journal of Learning*, 15(6), 179-186.
- Spatariu, A., Hartley, K., & Bendixen, L. (2004). Defining and measuring quality in online discussions. *The Journal of Interactive Online Learning*, 2(4). Retrieved from <http://www.ncolr.org>
- Story, L., Butts, J., Bishop, S., Green, L., Johnson, K., & Mattison, H. (2010). Innovative strategies for nursing education program evaluation. *Journal of Nursing Education*, 49(6), 351 – 354.

- Stufflebeam, D. L. (1971). The relevance of the CIPP evaluation model for educational accountability. *Journal of Research and Development in Education*, 5, 19-15.
- Stufflebeam, D. L. (1999). Program evaluation models metaevaluation checklist. *Foundational Models for 21st Century Program Evaluation*. Kalamazoo, MI: The Evaluation Center as Part of Western Michigan University.
- Stufflebeam, D. L. (2000). The CIPP model. In Stufflebeam, Madaus, & Kellaghan (Eds.), *Evaluation models: Viewpoints on education and human services evaluation* (2nd ed., pp. 19-32). Boston: Kluwer Academic.
- Stufflebeam, D. L. (2001). Evaluation models. In D. L. Stufflebeam, *New directions for evaluation: No. 89* (pp. 7-98). San Francisco: Jossey-Bass.
- Stufflebeam, D. L., & Shinkfield, A. J. (1984). *Systematic evaluation: a self-instructional guide to theory and practice*. Boston: Kluwer-Nkjhoff.
- Suhayda, R., & Miller, J. (2006). Optimizing evaluation of nursing education programs. *NurseEducator*, 31(5), 200-206.
- Thompson, M., & Irele, M. (2007). Evaluating distance education programs. In M. Moore (Ed.), *Handbook of distance education* (pp. 419-450). Mahwah, New Jersey: Lawrence Erlbaum Associates Publishers
- Thurmond, V. (2002). Considering theory in assessing quality of web-based courses. *Nurse Educator*, 27(1), 20-24.
- Turnbull, B. (1999). The mediating effect of participation efficacy on evaluation use. *Evaluation and Program Planning*, 22, 131-140. doi: 10.1016/S0149-7189(99)00012-9
- Tyler, R. W. (1949). *Basic principles of curriculum and instruction*. Chicago:

- University of Chicago Press.
- U.S. Department of Education. (2008). *National center for education statistics*. Retrieved from <http://nces.ed.gov/fastfacts/display.asp?id=80>
- Weiss, C. H. (1980). Knowledge creep and decision accretion. *Science communication*, 1(3), 382-404.
- Weiss, C. H. (1993). Where politics and evaluation research meet. *Evaluation Practice*, 14(1), 93-106.
- Wentling, T. L., & Johnson, S. D. (1999, October). The design and development of an evaluation system for online instruction, In P. Kuchinke (Ed.), *Proceedings of the 1999 Academy Human Resource Development Annual Conference* (pp. 548-553). Washington, DC. Academy of HRD.
- <http://ezproxy.gsu.edu:2048/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED448743&site=ehost-live>
- Wills, C., & Stommel, M. (2002). Graduate nursing students' precourse and postcourse perceptions and preferences concerning completely web-based courses. *Journal of Nursing Education*, 41(5), 193-201.
- Wilson, C. (2001). Faculty attitudes about distance learning. *Educause Quarterly*, 2, 70-71.
- Woo, M. A., & Kimmick, J. V. (2000). Comparison of Internet versus lecture instructional methods for teaching nursing research. *Journal of Professional Nursing*, 16(3), 132-139.
- Yarbrough, D. B., Shulha, L. M., Hopson, R. K., and Caruthers, F. A. (2011). *The*

program evaluation standards: A guide for evaluators and evaluation users (3rd ed.). Thousand Oaks, CA: Sage

APPENDICES

APPENDIX A

MAIN STUDY SURVEY INSTRUMENT

Survey Information and Eligibility

In this questionnaire we have identified six main variables that provide guidelines for systematic program evaluation of online education. Online education for this study is described as web-enhanced (partially online) and/or web-based (totally online) teaching/learning. Web-enhanced education is defined as teaching/learning events that combine aspects of online and face-to-face education or is partially online. Web-based education is defined as no face-to-face meeting with the instructor and the teaching/learning experience is totally online.

Does your school of nursing provide online (web-enhanced and/or web-based) educational activities at the master's degree level? If yes, please proceed with the survey. If no, using the enclosed self-addressed envelope, please return this page along with the survey.

- ☐ Yes
- ☐ No

Survey Instructions

As the number of online programs grows many of us are confronted with significant questions about how to evaluate such programs. Different programs have approached evaluation in various ways and some have not yet begun the process of systematic evaluation of online education. Below are a list of possible activities and processes of systematic program evaluation of online education. A typical nursing program will vary considerably whether they will do some or all of these activities. Your frank assessment is very important to us because we are trying to develop benchmarks for the field and what is really being done out there that goes beyond textbook prescriptions that say "you must". We are interested in how many times within the past 5 years did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs. As you complete the survey please base responses on program evaluation in your current school of nursing. For sections 1-6 choose only one answer for each response.

Systematic Program Evaluation of Online Nursing Education at the Master's Degree Level Survey

Section 1: Collecting data from students: (Section 1 of 7)

We are interested in how many times within the past 5 years did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs.

	Never	1 or 2 times	3 or 4 times	5 or more times
1. Collecting data from students about online learning needs (e.g., needs assessment)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2. Collecting data from students about technological knowledge in preparation for learning online	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3. Collecting data from students about satisfaction with technology support for online learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4. Collecting data from students about satisfaction with the admission process to our online program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5. Collecting data from students about adequacy of the technology learning system (e.g., Blackboard, eLive, etc.) used for online learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6. Collecting data from students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

about perceptions of the online learning social environment				
7. Collecting data from students about satisfaction with online learning	○	○	○	○
8. Collecting data from students about perceptions of teaching effectiveness	○	○	○	○
9. Collecting data from students about satisfaction with online library support and access	○	○	○	○
10. Collecting data from students about perceived quality of academic advisement	○	○	○	○
11. Collecting data from students about perceived quality of financial aid support	○	○	○	○
12. Collecting data from students about cost-benefit of online learning	○	○	○	○
13. Collecting data from students about perceived value of online course content	○	○	○	○

Section 2: Collecting data from faculty: (Section 2 of 7)

We are interested in how many times within the past 5 years did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs.

	Never	1 or 2 times	3 or 4 times	5 or more times
14. Collecting data from faculty about online teaching materials being up to date	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
15. Collecting data from faculty about technological knowledge in preparation to teach online	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
16. Collecting data from faculty about student technological preparation for online learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
17. Collecting data from faculty about responsiveness to students' needs in the teaching/learning environment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
18. Collecting data from faculty about control over online course content	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
19. Collecting data from faculty about satisfaction with the course-ware used for online education (e.g., Blackboard, etc.)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
20. Collecting data from faculty about satisfaction with university	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

technology support for online education				
21. Collecting data from faculty about reasonableness of workload when teaching online	○	○	○	○
22. Collecting data from faculty about instructor satisfaction with student learning	○	○	○	○

Section 3: Compiling and analyzing data from program records: (Section 3 of 7)

We are interested in how many times within the past 5 years did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs.

	Never	1 or 2 times	3 or 4 times	5 or more times
23. Compiling and analyzing data from program records about applicants' standardized test score results (e.g., GRE, MAT)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
24. Compiling and analyzing data from program records about applicants' prior academic achievement results (e.g., GPA)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
25. Compiling and analyzing data from program records about demographic data on students applying to our online program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
26. Compiling and analyzing data from program records about compliance of our online program with requirements from certifying bodies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
27. Compiling and analyzing data from program records about comparative cost effectiveness of online versus face-to-face instruction	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
28. Compiling and analyzing data from program records about comparisons between online versus traditional end-of-course evaluations	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

29. Compiling and analyzing data from program records about trends in online course enrollment	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
30. Compiling and analyzing data from program records about online course failure rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
31. Compiling and analyzing data from program records about online course completion/withdrawal rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
32. Compiling and analyzing data from program records about program enrollment rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
33. Compiling and analyzing data from program records about program retention rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
34. Compiling and analyzing data from program records about program graduation rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
35. Compiling and analyzing data from program records about national certification pass rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
36. Compiling and analyzing data from program records about faculty qualifications to teach online	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
37. Compiling and analyzing data from program records about faculty retention rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

38. Compiling and analyzing data from program records about staff retention rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
39. Compiling and analyzing data from program records about structures hindering effectiveness of our online program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
40. Compiling and analyzing data from program records about policies hindering effectiveness of our online program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
41. Compiling and analyzing data from program records about policies hindering transferability of online coursework	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
42. Compiling and analyzing data from program records about faculty selection process to teach online	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Section 4: Collecting data from alumni: (Section 4 of 7)

We are interested in how many times within the past 5 years did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs.

	Never	1 or 2 times	3 or 4 times	5 or more times
43. Collecting data from alumni about employment rates	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
44. Collecting data from alumni about satisfaction with online teaching/learning	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
45. Collecting data from alumni about satisfaction with the degree obtained	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
46. Collecting data from alumni about satisfaction with skills obtained	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
47. Collecting data from alumni about reduction of learning needs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
48. Collecting data from alumni about elimination of learning needs	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Section 5: Collecting data from employers: (Section 5 of 7)

We are interested in how many times within the past 5 years did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs.

	Never	1 or 2 times	3 or 4 times	5 or more times
49. Collecting data from employers about satisfaction with overall performance of graduates from our online education program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
50. Collecting data from employers about common problems of graduates from our online education program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
51. Collecting data from employers about the professionalism of graduates from our online education program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
52. Collecting data from employers about essential educational requirements needed by graduates from our online education program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
53. Collecting data from	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

employers about job retention rates of graduates from our online education program				
---	--	--	--	--

Section 6: My school of nursing: (Section 6 of 7)

In order to get the data mentioned in the above five sections schools of nursing could possibly be involved in a number of specific activities in planning and implementing the evaluation. We are interested in to what extent does your school of nursing engage in the following program evaluation practices.

	Strongly disagree	Disagree	Agree	Strongly agree
54. My school of nursing identifies stakeholders in the evaluation process	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
55. My school of nursing assesses program readiness for evaluation	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
56. My school of nursing identifies questions to be answered by the evaluation process	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
57. My school of nursing uses evaluation methods that are appropriate for the questions identified (e.g., qualitative, quantitative, mixed methods)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
58. My school of nursing keeps stakeholders informed of interim evaluation findings	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
59. My school of nursing decides how evaluation results are to be disseminated to	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

the stakeholders				
60. My school of nursing reviews evaluation results with faculty to identify best practices in online education	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
61. My school of nursing reviews evaluation results with staff to identify best practices in online education	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
62. My school of nursing uses evaluation data to monitor program performance at regular intervals	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
63. My school of nursing uses evaluation data to improve (make better) our online program	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
64. My school of nursing uses evaluation data to render judgment about our program (e.g., determine the overall merit, worth, significance, or value)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
65. My school of nursing uses evaluation data for accountability to someone or something (e.g.,	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

an accrediting or certifying body, funding organizations)				
66. My school of nursing uses evaluation data to generate knowledge about our program (e.g., to identify patterns of effectiveness, trends in participation, general problems with online learning)	○	○	○	○
67. My school of nursing uses evaluation data to make developmental changes to our program (e.g., alternative course sequencing, innovative delivery)	○	○	○	○

Section 7: Background information: (Section 7 of 7)

The final section of the survey is designed to collect information about you and your school of nursing.

68. Which of the following best describes your school of nursing? (Check one)
- ☐ Public
 - ☐ Private
69. From which of the following does your master's degree level program hold accreditation? (Check all that apply)
- ☐ National League for Nursing Accrediting Commission (NLNAC)
 - ☐ Commission on Colleges in Nursing Education (CCNE)
 - ☐ Other _____
70. Approximately how many students are in your master's degree level nursing program?
71. Of all the students in your master's degree level program approximately what percentage are full time students?
72. Of all the courses taught in your master's degree level program approximately what percentage are taught by part time faculty?
73. Approximately how many years have your master's degree level program been in existence?
74. Approximately how many years have your master's degree level program used online course ware? (e.g., Blackboard, eLive, etc.)
75. Which of the following best describes the method of online education delivery provided at the master's degree level? (Web-enhanced education is defined as teaching/learning events that combine aspects of online and face-to-face education or is partially online. Web-based education is defined as no face-to-face meeting with the instructor and the teaching/learning experience is totally online)
- ☐ Web-enhanced
 - ☐ Web-based
 - ☐ Both
76. Which of the following best describes where online education is used in the master's degree level curriculum?
- ☐ Program core curriculum courses only
 - ☐ Program specialty track curriculum courses only
 - ☐ Both

77. During the past 2 years has tuition for your organization:

- ☐ Decreased
- ☐ Stayed the same
- ☐ Increased

78. Indicate the most recent year in which your school of nursing evaluated their online master's degree level program.

79. Approximately how long have you been a nurse educator? (Months or years)

80. What is your current administrative job title?

81. Approximately how long have you held this current job title? (Months or years)

82. In what year were you born?

83. What is your gender?

- ☐ Female
- ☐ Male

84. What is your race/ethnicity?

*We appreciate you taking the time to answer our survey.
Many thanks!*

APPENDIX B

MAIN STUDY SUPPORTING DOCUMENTS

FIRST EMAIL CONTACT - PRENOTICE

Dear [Nurse Administrator]:

Online nursing education programs are booming and many of us are not sure how to approach the evaluation of these programs. I am a registered nurse and doctoral candidate student in Adult Education at the University of Georgia in Athens, GA under the supervision of Dr. Lorilee Sandman. We are conducting a research study on systematic program evaluation of online nursing education and how its practice and use is occurring in schools of nursing across the nation. Your institution is vital to discovering what is being done and to providing insight for this study. Therefore we are contacting you and other schools of nursing administrators, with a master's level degree program listed as being accredited by the National League for Nursing Accrediting Commission or the Commission on Colleges in Nursing Education.

In approximately one week you or your designee will receive an email with an introduction to the study and the survey link. The survey will take approximately 20 minutes of your valuable time. We will ask you or your designee to share what has been or currently being done in the way of systematic program evaluation of online education at the master's degree level.

We are contacting you in advance to ensure we have the appropriate contact or designated faculty at your school of nursing who is responsible for program evaluation at the master's degree level. This communication precedes the actual study introduction and the survey link.

_____No action is needed on your part if you are the appropriate individual and the email used for this correspondence is correct.

_____If you would like for this survey to be delivered to you at a different address, please email the preferred address to emh56@uga.edu before August 27, 2012.

_____If there is a more appropriate faculty or person within your school of nursing to participate in this survey, please reply to emh@56uga.edu before August 27, 2012 and include the following information (please feel free to forward this introductory email to your designee).

Name:
Title:
Institution:
Email:
Phone:

Many thanks in advance for your consideration. We are enthusiastically looking forward to the results that will be produced as a result of the input from your school of nursing!

Respectfully,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602
Phone: 770-761-8310 Email: emh56@uga.edu

Lorilee R. Sandmann, Ph.D.
Professor, Adult Education
Lifelong Education, Administration, and Policy
University of Georgia
850 College Station Rd., 413 River's Crossing
Athens, GA 30602
Phone 706.542.4014 W, 706.340.3746 C
sandmann@uga.edu
<http://www.coe.uga.edu/leap/academic-programs/adult-education/>

SECOND EMAIL CONTACT – FIRST LETTER WITH LINK

Dear [Nurse Administrator]:

Online nursing education is booming and many of us are not sure how to approach the evaluation of these programs. Different programs have various ways in which to evaluate and some have not yet begun the process of systematic evaluation of online education. We are conducting a study in the field of systematic program evaluation of online nursing education at the master's degree level. Online education for this study is described as web-enhanced (partially online) **and/or** web-based (totally online) teaching/learning. Web-enhanced education is defined as teaching/learning events that combine aspects of online and face-to-face education or is partially online. Web-based education is defined as no face-to-face meeting with the instructor and the teaching/learning experience is totally online.

Your school of nursing is instrumental in advancing state of the art evaluation practice and use.

We are turning to you in order to better understand to what extent is program evaluation being done in schools of nursing and what are the predictors for good systematic evaluation of online education. This is an area that is not well researched, and the results of this study can significantly impact the way schools of nursing perform program evaluation of online education. Schools of nursing who wish to establish or refine their evaluation system will be able to see what is being done at a majority of schools and what is being done at particularly excellent schools.

This online survey is designed to take approximately 20 minutes of your valuable time to complete. All individually identifiable responses will be treated with confidentiality, and only summarized data will be published. An executive summary of the research findings will be provided to participants upon completion of this study. This executive summary will be sent to the address used for this correspondence.

To complete the survey, please follow the link at the end of this message. On the first page of the survey (via Qualtrics) is the consent form. By clicking on the “next” option at the end of the informed consent you are agreeing to participate in the research project. This link is uniquely tied to this survey and your email address; please do not forward the message for other individuals to complete.

As an alternative or supplement to the online survey, a PDF file has been created. The PDF file is located at [location]. This file will allow you to complete or simply view the PDF version of the survey. If you decide to complete the survey in the paper form please send your completed survey the address below and include your contact information as a separate document in order to allow me to update either database.

3850 WILLOW BEND DRIVE
STOCKBRIDGE, GA 30281

Thank you for taking the time to offer your perspective and expertise to our study. Your response is vital to us as we strive to examine online nursing education program evaluation practice and use at the master's degree level!

Respectfully,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602
Phone: 770-761-8310 Email: emh56@uga.edu

Lorilee R. Sandmann Ph.D.
Professor, Adult Education
Lifelong Education, Administration, and Policy
University of Georgia
850 College Station Rd., 413 River's Crossing
Athens, GA 30602
Phone 706.542.4014 W, 706.340.3746 C
sandmann@uga.edu
<http://www.coe.uga.edu/leap/academic-programs/adult-education/>

INFORMED CONSENT

Dear Nurse Administrator:

Online nursing education is booming and many are not sure how to approach the evaluation of these programs. I am a registered nurse and doctoral student under the direction of Dr. Lorilee Sandmann in the Department of Lifelong Education, Administration, & Policy at The University of Georgia. We invite you to participate in a research study entitled Systematic Program Evaluation of Online Nursing Education at the Master's Degree Level: Current Practices. The purpose of this study is to investigate to what extent schools of nursing are currently performing systematic program evaluation of online education and how they are utilizing the results. Using a quantitative approach the following questions will guide this study; (a) to what extent are schools of nursing systematically evaluating their online education activities at the master's degree program level; (b) what aspects of the program are being evaluated; (c) to what extent are evaluation results used; (d) and to what extent do institutional and program characteristics affect evaluation data source, area, and use of schools of nursing online graduate programs.

Please know that you have been invited to participate because of your role in your school of nursing; actually responsible for or best able to answer questions regarding program evaluation in your master's level degree nursing program. Participation will involve responding to 84 items on a survey and should only take approximately 20 minutes of your time. Your involvement in the study is voluntary, and you may choose not to participate or to stop at any time without penalty or loss of benefits to which you are otherwise entitled. Your name and other records will be kept confidential to the fullest extent allowed. Only the research team will have access to the information you provide. Information may also be shared with those who make sure the study is done correctly (UGA IRB). Data will be labeled with a code rather than your name. The key to the code will be in a password protected file. The coded data file will be maintained on a separate computer in a locked cabinet in the researcher's office. The results of the research study may be published, but your name will not be used. In fact, the published results will be presented in summary form only. Your identity will not be associated with your responses in any published format.

The findings from this project may provide information on state of the art online program evaluation. Schools of nursing who wish to establish or refine their evaluation system will be able to see what is being done at a majority of schools and what is being done at particularly excellent schools. There are no known risks or discomforts associated with this research.

If you have any questions about this research project, please feel free to call me at 770-761-8310 or Dr. Lorilee Sandmann at 706-542-4014 or send an e-mail to emh56@uga.edu. Questions or concerns about your rights as a research participant should be directed to The Chairperson, University of Georgia Institutional Review Board, 629 Boyd GSRC, Athens, Georgia 30602; telephone (706) 542-3199; email address irb@uga.edu. By completing and returning this questionnaire in the envelope provided, you are agreeing to participate in the above described research project.

Thank you for your consideration! Please keep this letter for your records.

Sincerely,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602

Lorilee R. Sandmann, Ph.D.
Professor, Adult Education Lifelong Education, Administration, and Policy
University of Georgia
850 College Station Rd., 413 River's Crossing
Athens, GA 30602
Phone 706.542.4014 W, 706.340.3746 C
sandmann@uga.edu

THIRD EMAIL CONTACT – FIRST REMINDER

Dear [Nurse Administrator]:

We recently sent you a survey seeking your perspectives regarding program evaluation practice and use of online nursing education at the master's degree level. Online education for this study is described as web-enhanced (partially online) **and/or** web-based (totally online) teaching/learning. Web-enhanced education is defined as teaching/learning events that combine aspects of online and face-to-face education or is partially online. Web-based education is defined as no face-to-face meeting with the instructor and the teaching/learning experience is totally online.

We are hoping you will take 20 minutes to give us your valuable input into investigating the extent of systematic program evaluation of online nursing education at the master's degree level. If you have not yet responded, we look forward to your participation. Please see the survey link at the end of this message. This link is uniquely tied to this survey and your email address; please do not forward the message for other individuals to complete.

As an alternative or supplement to the online survey, a PDF file has been created. The PDF file is located at [location]. This file will allow you to complete or simply view the PDF version of the survey. If you decide to complete the survey in the paper form please send your response to the address below and include your contact information as a separate document in order to allow me to update the database.

3850 Willow Bend Drive
Stockbridge, GA 30281

Thank you for contributing your time and insight to our study regarding practice and use of systematic program evaluation of online nursing education at the master's degree level. Although many of us are swamped by emails this time of year if I don't hear from you I will send a paper copy out as a third and final reminder. We look forward to your response!

Respectfully,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602
Phone: 770-761-8310 Email: emh56@uga.edu

Lorilee R. Sandmann Ph.D.

Professor, Adult Education

Lifelong Education, Administration, and Policy

University of Georgia

850 College Station Rd., 413 River's Crossing

Athens, GA 30602

Phone 706.542.4014 W, 706.340.3746 C

sandmann@uga.edu

<http://www.coe.uga.edu/leap/academic-programs/adult-education/>

FOURTH CONTACT – MAIL AND PAPER SURVEY REMINDER

Dear [Nurse Administrator]:

Many of us are swamped by requests to complete survey this time of year and we understand how valuable your time is during the semester. We are hoping you will take 20 minutes to give us your valuable input into investigating the extent of systematic program evaluation of online nursing education at the master's degree level. To the best of our knowledge, the survey has not yet been completed. If you feel our records are incorrect please respond directly to emh56@uga.edu.

We are continuing to collect data for our study on this very important subject of program evaluation. Therefore, we are contacting you again for your response in order to collect meaningful data and provide you an opportunity to contribute to the success of this study. The survey is enclosed with this letter, along with a stamped self-addressed envelope. The survey is to be completed based on your knowledge of program evaluation practices within your school of nursing. If you are not the person to answer this survey, please pass it on the appropriate faculty member for completion.

Thank you for contributing your time and insight to our study regarding practice and use of systematic program evaluation of online nursing education at the master's degree level. We look forward to your response!

Respectfully,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602
Phone: 770-761-8310 Email: emh56@uga.edu

Lorilee R. Sandmann Ph.D.
Professor, Adult Education
Lifelong Education, Administration, and Policy
University of Georgia
850 College Station Rd., 413 River's Crossing
Athens, GA 30602
Phone 706.542.4014 W, 706.340.3746 C
sandmann@uga.edu

APPENDIX C

INITIAL ITEM POOL BY SOURCE and AREA

Models reviewed and used from literature:

Patton –Utilization-Focused Evaluation model (2009)
Rovai – A Practical Framework for Evaluating Online Distance Education Programs (2003)
Chapman – Evaluation Plan for Fully Online Degree Programs (2005)
Lockee, etc- Measuring Success: Evaluation Strategies for DE (2002)
Educause/Newman – Measuring Success in Web-Based DL (2003)

Quality indicators (benchmarks) used:

Institute of Higher Education – Quality on the Line: Benchmarks for Success in Internet-Based Distance Education
Billings – Benchmarking Best Practices in Web-Bases Nursing Courses (2001)

PE = Personal Experience

* = fall into several categories/area; realistic expectations of evaluation; measurement issues

Initial Item Pool by Source and Area

#	Item	Author/source/citation	Area
1 .	Student demographics	Rovai	Input
2.	Standardized test scores (GRE. MAT)	Rovai/PE	
3.	Prior academic achievement (GPA)	Rovai/PE/Chapman	
4.	Employment status (work/clinical experience)	PE	
5.	Instructor qualifications	Rovai	
6.	Entrance requirements (GRE, MAT, GPA, clinical experience)	Rovai/PE/Chapman	
7.	Educational needs targeted by the program	Rovai	
8.	Program objectives responsive to targeted student's needs	Rovai/Chapman	
9.	Course objectives responsive to targeted student's needs	Rovai	
10 .	Currency of course materials	Rovai	

11	Efficiency of course development process	Rovai/Chapman/Billings/PE
12	Technical knowledge of potential students	Rovai/Billings/Chapman
13	Technical knowledge of faculty/potential faculty	Rovai/PE/Chapman
14	Technical equipment used in the program	Rovai/Billings
15	Response to student curriculum needs	Rovai
16	Student course evaluations	Rovai/Billings/
17	Selection of faculty to teach online	Rovai
18	Faculty qualifications to teach online	Rovai/Chapman
19	Faculty qualification to teach content of their courses	Rovai
20	Faculty selection to teach at a distance	Rovai
21	Faculty control over course content	Rovai
22	Faculty technologic needs	Rovai/Billings/Chapman/Lockee/Educause
23	Faculty satisfaction with technology	Rovai/Billings/Chapman/Lockee/Educause
24	Faculty satisfaction with workload	Rovai/Chapman/Lockee
25	Faculty satisfaction with training	Rovai/Chapman/Lockee
26	Faculty satisfaction with teaching load	Rovai/Chapman/Lockee
27	eLearning system meet program requirements	Rovai/Chapman/Lockee
28	Adequacy of eLearning system	Rovai/Chapman/Lockee
29	eLearning system ease of use	Rovai/Chapman/Lockee/Billings
30	eLearning integration issues	Rovai/Chapman/Lockee/Billings
31	eLearning system technical support	Rovai/Chapman/Lockee/Billings
32	eLearning system	Rovai/Chapman/Educause

.	suitable in terms of cost		
33	Effect of the eLearning environment on student progress	Rovai/Chapman/Lockee/Billings/ Educause	
34	Effect of eLearning environment on student performance	Rovai/Chapman/Lockee/Billings/ Educause	
35	Courses offered	Rovai	Process
36	Course enrollments	Rovai/Billings	
37	Course enrollment trends	Rovai/Chapman/Lockee	
38	Overall course persistence rates	Rovai/Chapman/Lockee/Billings/ Educause	
39	Student reasons for asking for refunds	Rovai	
40	Student reasons for enrollment termination	Rovai/Chapman	
41	Percent of student who take other distance education courses	Rovai/Chapman	
42	Course evaluations	Rovai/Chapman/Lockee/Billings	
43	Student support needs	Rovai/Chapman/Lockee/Billings Educause	
44	Faculty support needs	Rovai/Chapman/Lockee/Billings/ Educause	
45	Adequacy of support services (e.g., online library, student advisement, delivery of course materials, counseling, program administrators, placement, registrar's office, financial aid)	Rovai	
46	Program implemented as intended	Rovai/Lockee/Billings/Chapman	
47	Interaction between students/instructors in	Rovai/PE	

	the learning environment		
48	Timely instructor feedback	Rovai/Billings	
49	Course evaluation (student)	Rovai/Chapman/Lockee/Billings	
50	Instructor persistence rates	Rovai/PE/Educause	
51	Student time on task (dealing with the content of the course)	Rovai/PE/Billings	
52	Integrity of students work assured	Rovai/Lockee	
53	Cost effectiveness	Rovai/Chapman/Lockee/Billings/ Educause	
54	Interaction rates	Rovai/Billings	
55	Transferability of coursework	Rovai	Output
56	Recognition of degrees/certificates by profession	Rovai/Chapman/Lockee	
57	Identify students who could not enroll in F-to-F programs; access	Rovai	
58	Program completion rates	Rovai/Chapman/Lockee/Billings	
59	Course completion rates	Rovai/Chapman/Lockee/	
60	Structures hindering program outcomes	Rovai	
61	Policies hindering program effectiveness	Rovai	
62	Structures hindering program effectiveness	Rovai	
63	Policies hindering program outcomes	Rovai	
64	Course evaluation by students	Rovai/Chapman/Lockee/Billings Educause	
65	Program relevance to student needs	Rovai/Chapman	
66	Student gratification	Rovai/Billings	
67	Skill	Rovai/Chapman/Lockee	

.	development/educational gains		
68	Program grade trends	Rovai/PE/Chapman/Billings	
.			
69	Course grade trends	PE	
.			
70	Graduation rates	Rovai/Chapman/Educause/Lockee	
.			
71	Extent to which program reduced/eliminated student needs	Rovai/Chapman	Impact
.			
72	Attitudes of graduates concerning the program	Rovai/Chapman	
.			
73	Changes in learner educational expectations	Rovai	
.			
74	Employer satisfaction with graduates	Rovai/PE	
.			
75	Benefits of graduates learning to employer	Rovai	
.			
76	Changes in graduates job performance	Rovai/PE	
.			
77	Return on investment in terms of cost to student (increased pay, promotion, better job)	Rovai/Chapman/Educause/Lockee	
.			
78	Return on investment in terms of organizational impact	Rovai/Chapman/Educause/Lockee	
.			
79	National certification pass rate	Rovai/PE	
.			
80	Program accreditation information	Rovai/Lockee	Input/ program characteristic
.			
81	Student demographics	Rovai/Educause/Lockee	Input/ program characteristic
.			
82	Input from program coordinators	Chapman	Input/Impact
.			
83	Use program evaluation data to establish other online programs	Chapman	Use
.			
84	Keep stakeholders informed with evaluation interim findings	Patton	Use
.			
85	Strengths and weaknesses of program from faculty	Chapman/PE	Output/ impact
.			

86 .	Curriculum effectiveness	Chapman	Process
87 .	Employment rates of graduates	Chapman/PE	Impact
88 .	Employer satisfaction with graduate preparation	Chapman	Impact
89 .	Stakeholder identification	Patton	Utilization
90 .	Program ready for evaluation	Patton	Utilization
91 .	Evaluation questions identified	Patton	Utilization
92 .	Evaluation methods appropriated to question used	Patton	Utilization
93 .	Decide on dissemination methods to staff	Patton	Utilization
94 .	Review of evaluation results for best practices	Chapman/Lockee	Utilization
95 .	Comparison of online program's goals, objectives, focus, and course offering with identified top-tiered nursing online programs	Chapman	*
96 .	Faculty awards, incentives, recognition	Educause	Impact
97 .	Market reach/new market opportunities	Educause	Impact
98 .	Positive change in learner attitude	Lockee	Impact

APPENDIX D

INITIAL PROTOTYPE ITEMS

Collects demographic data on students applying to our online program

Collects standardized test score results of students applying to our online program (i.e. GRE, MAT)

Collects prior academic achievement data on students applying to our online program (GPA)

Collects data from students to evaluate online educational needs

Periodically evaluates program objectives for responsiveness to targeted student's needs

Periodically evaluates course objectives for responsiveness to targeted student's needs

Periodically evaluates program courses for currency of materials used in online courses

Collects data from potential students to evaluate technological knowledge in preparation for learning online

Collects data from potential faculty to evaluate technological knowledge in preparation to teach online

Collects data to evaluate response to student curriculum needs

Collects data from students to evaluate satisfaction with technology support for online education

Collects data from students to evaluate satisfaction with technology used in online education

Collects data to assess faculty qualifications to teach online

Periodically evaluates faculty selection process to teach online

Periodically evaluates faculty control over course content

Collects data from faculty to evaluate faculty technologically needs for online teaching

Collects data from faculty to evaluate satisfaction with technology used in online education

Collects data from faculty to evaluate satisfaction with technology support for online education

Collects data from faculty to evaluate online course development training needs

Collects data from faculty to evaluate online teaching load

Collects data from faculty to evaluate workload

Collects data from students to evaluate satisfaction with the admissions' process to our program

Collects data to evaluate that program requirements are met by the eLearning

Collects data to evaluate eLearning system's ease of use

Collects data to evaluate the adequacy of the eLearning system used for online education

Collects data to evaluate eLearning system technical support

Collects data to evaluate the ease of eLearning system integration

Collects data from students and faculty to evaluate the effect of the eLearning environment on student progress

Collects data from students and faculty to evaluate the effect of the eLearning environment on student performance

Periodically collects data to evaluate SON cost benefit of online education

Periodically collects data to evaluate student cost benefit of online education

Collects data from student end-of-course evaluations

Periodically evaluates student perceptions of online learning

Periodically evaluates student satisfaction with online learning

Periodically evaluates student perceptions of teaching effectiveness

Collects data to evaluate student satisfaction with the eLearning system

Periodically compares online versus traditional end-of-course evaluations

Periodically collects data that compares online course enrollment, retention, and completion rates to traditional courses

Collects data from students to evaluate online library support

Collects data from students to evaluate the quality of academic advisement

Collects data from students to evaluate the quality of financial aid support

Collects data from students to evaluate the quality of scholarship offerings

Collects data from faculty to evaluate teaching effectiveness in online courses

Collects data to compare student achievement in face-to-face versus traditional course

Collects data on trends in online course enrollment

Collects data on online course failure rates

Collects data on online course completion rates

Collects data on reasons for withdrawals from online courses

Collects data on instructor persistence rates

Collects data to evaluate cost benefit of online education

Collects data to evaluate student cost benefit of online education

Collects data on transferability of online coursework

Collects data on recognition of degrees/certificates by the nursing profession

Collects data that identifies students who could not enroll in face to face programs

(Collects data that)... Tracks student course enrollment data (rates)

Tracks student program enrollment rates (data)

Tracks student course retention (completion) rates

Tracks student program retention (completion) rates

Tracks student program graduation rates

Tracks student pass rates for national certification

Collects data on student satisfaction with degree obtained

Collects data on student satisfaction with skills obtained

Collects data on student perceptions of met learning needs

Collects data on faculty retention rates

Collects data on staff retention rates

Periodically evaluates structures hindering program outcomes and effectiveness

Periodically evaluates policies hindering program effectiveness and outcomes

Collects data on program grade trends

Collects data on course grade trends

Collects data on student gratification with course

Collects data on student gratification with program

Collects data on student satisfaction with degree obtained

Collects data from graduates to evaluate reduced/eliminated student learning needs

Collects data on student employment rates

Periodically evaluates alumni satisfaction with online program

Periodically evaluates employer satisfaction with graduates from online program

Tracks national certification pass/fail rates

Tracks graduation rates

Collects data to evaluate cost benefit of online education

Collects data to evaluate student cost benefit of online education

Collects data to evaluate return on investment of online education

Identify stakeholders in the evaluation process

Assess program readiness for evaluation

Identify questions to be answered by the evaluation process

Use evaluation methods that are appropriate for the questions identified

Keep stakeholders informed of interim evaluation findings

Decide how evaluation results are to be disseminated to the staff

Review evaluation results with faculty and staff for best practices

Use evaluation data to improve online courses

Use evaluation data to improve online programs

Use evaluation data to establish other online programs

What type of organization is your SON?

_____Public

_____Private

Indicate the considered setting or location of your SON?

_____Urban

_____Suburban

_____Rural

Which nursing accrediting organization is your SON a member?

_____NLN

_____CCNE

How long has your SON used eLearning systems for online masters; education?

What type of online education does your SON provide at the masters' level?

_____Web enhanced

_____Web based

_____Combination of both

In which area (s) of the masters' curriculum does your SON use online education?

_____graduate core courses only

_____graduate specific program courses only

_____both

Indicate the most recent year in which your SON's evaluated your online masters' program. _____

APPENDIX E

REDUCED ITEM POOL FOR EXPERT REVIEW

[UGA LETTERHEAD]

Directions: Below you will find a list of activities and practices used to evaluate online education programs. Imagine that you are the director of a nursing program and you are going to evaluate your web-based and/or web-enhanced master's degree level program. In reviewing activities and practices that can be evaluated you recognize that some seem important to the evaluation process and some less so. Which of the following would you in fact use?

How important is each of the following practices for the evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs?

1= not important	2 = somewhat important	3 = quite important
4= very important	5 = extremely important	

Collecting data from students:

Collecting data from students about online educational needs (e.g., needs assessment)

Collecting data from students about technological knowledge in preparation for learning online

Collecting data from students about satisfaction with technology support for online education

Collecting data from students about satisfaction with the admission process to online programs

Collecting data from students about adequacy of the technology learning system used for online education

Collecting data from students about perceptions of the social environment of online courses

Collecting data from students about satisfaction with online learning

Collecting data from students about perceptions of teaching effectiveness

Collecting data from students about satisfaction with online library support and access

Collecting data from students about perceived quality of academic advisement

Collecting data from students about perceived quality of financial aid support

Collecting data from students about cost-benefit of online education

Collecting data from students about perceived value of course content

Collecting data from faculty:

Collecting data from faculty about currency of online teaching materials

Collecting data from faculty about technological knowledge in preparation to teach online

Collecting data from faculty about student technological preparation for online learning

Collecting data from faculty about responsiveness to students' needs

Collecting data from faculty about control over online course content

Collecting data from faculty about satisfaction with the course ware used for online education (e.g., blackboard, Web-ct, etc.)

Collecting data from faculty about satisfaction with university technology support for online education

Collecting data from faculty about reasonableness of workload

Collecting data from faculty about instructor satisfaction with student learning

Compiling and analyzing data from program records:

Compiling and analyzing data from program records about applicants' standardized test score results (e.g., GRE, MAT)

Compiling and analyzing data from program records about applicants' prior academic achievement results (e.g., GPA)

Compiling and analyzing data from program records about demographic data on students applying to online programs

Compiling and analyzing data from program records about compliance of online programs with requirements from certifying bodies

Compiling and analyzing data from program records about comparative cost effectiveness of online versus face-to-face instruction

Compiling and analyzing data from program records about comparing schools of nursing between online versus traditional end-of-course evaluations

Compiling and analyzing data from program records about trends in online course enrollment

Compiling and analyzing data from program records about online course failure rates

Compiling and analyzing data from program records about online course completion/withdrawal rates

Compiling and analyzing data from program records about program enrollment rates

Compiling and analyzing data from program records about program retention rates

Compiling and analyzing data from program records about program graduation rates

Compiling and analyzing data from program records about national certification pass rates

Compiling and analyzing data from program records about faculty qualifications to teach online

Compiling and analyzing data from program records about faculty retention rates

Compiling and analyzing data from program records about staff retention rates

Compiling and analyzing data from program records about structures hindering effectiveness of online programs

Compiling and analyzing data from program records about policies hindering effectiveness of online programs

Compiling and analyzing data from program records about policies hindering transferability of online coursework

Compiling and analyzing data from program records about faculty selection process to teach online

Collecting data from alumni:

Collecting data from alumni about employment rates

Collecting data from alumni about satisfaction with online teaching

Collecting data from alumni about satisfaction with the degree obtained

Collecting data from alumni about satisfaction with skills obtained

Collecting data from employers:

Collecting data from employers about satisfaction with overall performance of graduates from online programs

Collecting data from employers about recurrent problems of graduates from our online programs

No matter what data you get, the evaluation process will involve a number of specific actions to enhance use of the results. How important is planning for use of a well-crafted evaluation plan of online (web-enhanced and/or web-based) master's degree level nursing programs?

Planning for use of the evaluation:

Planning for use of the evaluation by identifying stakeholders in the evaluation process

Planning for use of the evaluation by assessing program readiness for evaluation

Planning for use of the evaluation by identifying questions to be answered by the evaluation process

Planning for use of the evaluation by using evaluation methods that are appropriate for the questions identified (e.g., qualitative, quantitative, mixed methods)

Planning for use of the evaluation by keeping stakeholders informed of interim evaluation findings

Planning for use of the evaluation by deciding how evaluation results are to be disseminated to the stakeholders

Planning for use of the evaluation by reviewing evaluation results with faculty to identify best practices in online education

Planning for use of the evaluation by reviewing evaluation results with staff to identify best practices in online education

Planning for use of the evaluation by using evaluation data to improve online programs

Planning for use of the evaluation by using evaluation data to establish new online programs

Answer the following demographic questions regarding your school of nursing (SON) master's level degree program.

What type of organization is your SON?

Public
Private

Indicate the considered setting or location of your SON?

Urban
Suburban
Rural

Which nursing accrediting organization is your SON a member?

NLN
CCNE
Both of the above

What is the size of your master's degree level in nursing program?

1 – 25 students
25 – 50 students
50 – 75 students
75 – 100 students
Greater than 100 students

How many students are full time in your SON master's degree level program?

Less than 25%

25 – 50%

50% - 75%

Greater than 75%

How many students are part time in your SON master's degree level program?

Less than 25%

25 – 50%

50% - 75%

Greater than 75%

How many of your faculty are full time in your master's degree level program?

Less than 25%

25 – 50%

50% - 75%

Greater than 75%

How many of your faculty are part time in your master's degree level program?

Less than 25%

25 – 50%

50% - 75%

Greater than 75%

How long has your master's level program been in existence?

1-4 years

5-10 years

10 or more years

How long has your SON used online course ware (e.g., Web-ct, blackboard) for master's degree level education?

1 – 2 years

3 – 4 years

5 years

Longer than 5 years

What type of online education does your SON provide at the master's degree level?

Web enhanced

Web based

Combination of both

None

In which area (s) of the master's curriculum does your SON use online education?

Graduate core courses only

Graduate specific program courses only

Both

None

During the past 2 years has the tuition for your organization:

Increased

Decreased

Stayed the same

Indicate the most recent year in which your SON evaluated their master's degree level program. _____

APPENDIX F

EXPERT REVIEW CRITIQUE SUMMARY

n = 8

In General

Good number of questions. Took from 10-15 minutes to complete as stated in the letter.

Look at ranges in demographic items for overlap

Be consistent in wording of items

Many terms will need defining

Visual appeal of overall survey

General consensus good/great

Layout of survey

Would add more color

Add design

Break up question 3 “collecting data from program records” items to fit two pages instead of one.

Thank you letter at the end of the survey

Letter

Need weed out question before survey begins

Change term “research” to “examine”

Which items refer to second question, barriers and facilitators?

Only one questionnaire?

Change wording of third research question

Directions

Actual practice or model practice? How will you determine that in your directions?

Include in directions a summary of what each question will ask, e.g., “section I will be student data”, etc.

Add after “how important...” please select the response that best reflex...”

Question “How important”

This question confused me since it was asked indirectly; the activities were rated but I wasn’t asked which I would drop

Scale

Hard to tell the difference between quiet and very important

How about adding “N/A”. Every item may not apply

Items heading or source identification

Panel had different ideas on how to categorize each section, e.g., by technology, students, faculty, curriculum, student services, retention, enrollment, evaluation. Liked the term “section” to introduce the source topic, e.g., Section I. One panelist suggested a timeline order from admission to graduation

From one expert “I think you could reduce the wordage or reading with your survey by not repeating the main question with each item (note experts were consistent in this critique 6 out of 8). I also think that the actual question should be at the end of the

question, not the beginning. For example, for question 2. (which I think should be question 1) I would revise it to read:

When evaluating an online (web-enhanced and/or web-based) master's degree level nursing program HOW IMPORTANT IS COLLECTING DATA FROM FACULTY ABOUT THE:

Currency of online teaching materials?

Technological knowledge of faculty preparing to teach an online course?

Students' technological preparation for online learning?

Individual item comments:

#2 – need example of “technological knowledge”

#4 – confusion, not sure what this meant

#5 – what is “technology leaning systems?”

#6 – what is meant by “social environment?” Peer-to-peer

#11 – is this a directors responsibility; in many Schools of nursing cost is external to the running of the program; meaning we don't see it.

#14 – use “up to date” instead

#18 – use another word than “control”

#36 – are there qualifications to teaching online?

#41 – define transferability

#48 – delete “our” just for consistency

#49-58 - Change to “planning to evaluate the program”. “This section quite confusing”

53, 55, 58- seems to me that (referred to #) are actually use of the evaluation, not planning for the evaluation

#63 – “what percentile” and look at range for overlap. Be more specific in demographics – look at ranges for over lap

#64 – “what percentage?”. Is this question redundant to #63?

#68 – Do you want them to use the year? Stated? Or this range of years? Why not ask them the number of years? Then can do a mean (SD) instead of just frequency???

#69 – If you don't define web enhanced and web based...will others know the definition?

Do you care about synchronous or non synchronous programs? (consistent critique)

#70 – understand this...unclear about this terminology (core vs. specific program courses; may need to put example)

#72 – Are you interested in whether they evaluate their online separate from tradition? Or whether they have instituted changes as a result of a program evaluation? Just some thoughts.

APPENDIX G

PROTOTYPE INSTRUMENT FOR PROSPECTUS

[UGA LETTERHEAD]

Introduction: In this questionnaire we have identified five main constructs that provide guidelines for good systematic program evaluation of online education. Online education for this study is described as web-enhanced **and/or** web-based teaching/learning. Web-enhanced education is defined as teaching/learning events that combine aspects of online and face-to-face education. Web-based education is defined as no face-to-face meeting with the instructor and the teaching/learning experience is totally online.

Does your school of nursing provide online (web-enhanced and/or web-based) educational activities at the master's degree level?

Yes

No

Instructions: One of your responsibilities as chief nurse administrator or master's degree level program director/coordinator is to systematically evaluate your school of nursing (SON) online education program at regularly designated intervals.

Below you will find a list of activities and practices used to evaluate online education programs. *We are interested in **how many times** within the **past 5 years** did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs. As you complete the survey please base your responses on program evaluation in your current school of nursing. For sections 1-6 choose only one answer for each response.*

1= Never

2 = 1 or 2

3 = 3 or 4

4 = 5 or more

Section 1: Collecting data from students:

Collecting data from students about online learning needs (e.g., needs assessment)

Collecting data from students about technological knowledge in preparation for learning online

Collecting data from students about satisfaction with technology support for online learning

Collecting data from students about satisfaction with the admission process to online programs

Collecting data from students about adequacy of the technology learning system

(Blackboard, Web-ct, eLive, etc) used for online learning

Collecting data from students about perceptions of the social environment of online learning

Collecting data from students about satisfaction with online learning

Collecting data from students about perceptions of teaching effectiveness

Collecting data from students about satisfaction with online library support and access

Collecting data from students about perceived quality of academic advisement

Collecting data from students about perceived quality of financial aid support

Collecting data from students about cost-benefit of online learning

Collecting data from students about perceived value of online course content

Section 2: Collecting data from faculty:

Collecting data from faculty about online teaching materials being up to date

Collecting data from faculty about technological knowledge in preparation to teach online

Collecting data from faculty about student technological preparation for online learning

Collecting data from faculty about responsiveness to students' needs in the teaching/learning environment

Collecting data from faculty about control over online course content

Collecting data from faculty about satisfaction with the courseware used for online education (e.g., blackboard, Web-ct, etc.)

Collecting data from faculty about satisfaction with university technology support for online education

Collecting data from faculty about reasonableness of workload when teaching online

Collecting data from faculty about instructor satisfaction with student learning

Section 3: Compiling and analyzing data from program records:

Compiling and analyzing data from program records about applicants' standardized test score results (e.g., GRE, MAT)

Compiling and analyzing data from program records about applicants' prior academic achievement results (e.g., GPA)

Compiling and analyzing data from program records about demographic data on students applying to online programs

Compiling and analyzing data from program records about compliance of online programs with requirements from certifying bodies

Compiling and analyzing data from program records about comparative cost effectiveness of online versus face-to-face instruction

Compiling and analyzing data from program records about comparing schools of nursing between online versus traditional end-of-course evaluations

Compiling and analyzing data from program records about trends in online course enrollment

Compiling and analyzing data from program records about online course failure rates

Compiling and analyzing data from program records about online course completion/withdrawal rates

Compiling and analyzing data from program records about program enrollment rates

Compiling and analyzing data from program records about program retention rates

Compiling and analyzing data from program records about program graduation rates

Compiling and analyzing data from program records about national certification pass rates

Compiling and analyzing data from program records about faculty qualifications to teach online

Compiling and analyzing data from program records about faculty retention rates

Compiling and analyzing data from program records about staff retention rates

Compiling and analyzing data from program records about structures hindering effectiveness of online programs

Compiling and analyzing data from program records about policies hindering effectiveness of online programs

Compiling and analyzing data from program records about policies hindering transferability of online coursework

Compiling and analyzing data from program records about faculty selection process to teach online

Section 4: Collecting data from alumni:

Collecting data from alumni about employment rates

Collecting data from alumni about satisfaction with online teaching/learning

Collecting data from alumni about satisfaction with the degree obtained

Collecting data from alumni about satisfaction with skills obtained

Section 5: Collecting data from employers:

Collecting data from employers about satisfaction with overall performance of graduates from online education programs

Collecting data from employers about recurrent problems of graduates from online education programs

No matter what data you get, the evaluation process and practices will require that your school of nursing be involved in a number of specific activities to enhance evaluation utilization in order to make improvements to or decisions about your program. We are interested in to what extent is your school of nursing engaging in the following program evaluation utilization practices in order to enhance evaluation use.

1 = Not at all

2 = Somewhat

3 = Greatly

4 = Totally

Section 6: Engaging in utilization practices:

Engaging in utilization practices by identifying stakeholders in the evaluation process

Engaging in utilization practices by assessing program readiness for evaluation

Engaging in utilization practices by identifying questions to be answered by the evaluation process

Engaging in utilization practices by using evaluation methods that are appropriate for the questions identified (e.g., qualitative, quantitative, mixed methods)

Engaging in utilization practices by keeping stakeholders informed of interim evaluation findings

Engaging in utilization practices by deciding how evaluation results are to be disseminated to the stakeholders

Engaging in utilization practices by reviewing evaluation results with faculty to identify best practices in online education

Engaging in utilization practices by reviewing evaluation results with staff to identify best practices in online education

Engaging in utilization practices by using evaluation data to improve online programs

Engaging in utilization practices by using evaluation data to establish new online programs

The final section of the survey is designed to collect information about you and your school of nursing (SON).

Section 7: Background information:

59. Which of the following best describes your SON? (*Check one*)

Public

Private

60. Which of the following setting or location best describes your SON? (*Check one*)

Urban

Rural

61. From which of the following does your master's degree level program hold accreditation? (*Check all that apply*)

NLN

CCNE

62. Approximately how many students are in your master's degree level nursing program? _____

63. Approximately what percentage of full time students are in your master's level degree program? _____

64. Approximately what percentage of full time faculty teaches in your master's level degree program? _____

65. Approximately how many years have your master's degree level program been in existence?

66. Approximately how many years have your master's degree level program used online course ware (e.g., Web-ct, blackboard, eLive, etc.)?

67. Which of the following best describes the method of online education delivery provided at the master's degree level? (*Check all that apply*)

Web enhanced

Web based

68. Which of the following best describes where online education is used in the master's degree level curriculum? (*Check all that apply*)

Core curriculum courses

Program specialty curriculum courses

69. During the past 2 years has tuition for your organization:

Increased

Decreased

Stayed the same

70. Indicate the **most recent purpose** for which your SON evaluated their master's degree level program. _____

71. Indicate the **most recent year** in which your SON evaluated their master's degree level program. _____
72. Approximately how many **total** years have you been a nurse educator? _____
73. What is your current administrative job title? _____
74. Approximately how long have you held this current job title? _____
75. Have you had formal training in program evaluation? _____
76. In what year were you born? _____
77. What is your gender? _____
78. What is your race/ethnicity? _____

APPENDIX H

PILOT STUDY SURVEY WITH SUPPORTING DOCUMENTS

Pilot Study Contact Letter

Name

Title

Institution

Dear [Nurse Administrator]:

My name is Eva Horne and I am a doctoral student in Adult Education at the University of Georgia in Athens, GA under the supervision of Dr. Lorilee Sandman. I am conducting a graduate study of systematic program evaluation of online nursing education and how its practice and use is occurring in schools of nursing across the nation. I am in the pilot phase of the research project. Your school of nursing, name, and email address were retrieved from the National League for Nursing Accrediting Commission and the Commission on Colleges in Nursing Education web-list as providing accredited nursing education at the master's degree level.

We are turning to you for assistance in piloting the survey instrument for this study. As we refine the final instrument, the data from your responses may be included in the final study but you will not be contacted again to participate in the main study. All individually identifiable responses will be treated with confidentiality and only summarized data will be published. An executive summary of the research findings will be provided to participants upon completion of this study. The executive summary will be emailed to the email account used for this correspondence.

We invite you to complete the pilot survey by following the link below for online completion. A PDF version can be obtained by emailing emh56@uga.edu. On the first page of the survey (via survey monkey) is the consent form. By clicking on the "next" option at the end of the informed consent you are agreeing to participate in the research project. The survey works best in Internet Explorer browser. Link:

This link is uniquely tied to this survey and your email address; please do not forward the message for other individuals to complete. We ask that if you are not the right contact, would you kindly provide the designated faculty's contact information by return email.

Name:

Title:

Email:

Phone:

Thanks in advance for your valuable time and assistance. We appreciate your participation in this study.

Sincerely,

Eva M. Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602
Phone: 770-761-8310 Email: emh56@uga.edu

Dr. Lorilee R. Sandmann
Professor and Program Chair, Adult Education
Department of Lifelong Education, Administration, & Policy
The University of Georgia

Information Letter

Dear [Nurse Administrator]:

I am a doctoral student under the direction of Dr. Lorilee Sandmann in the Department of Lifelong Education, Administration, & Policy at The University of Georgia. I invite you to participate in a research study entitled *Systematic Program Evaluation of Online Nursing Education at the Master's Degree Level: Current Practices*. The purpose of this study is to investigate to what extent schools of nursing (SONs) are currently performing systematic program evaluation of online education and how they are utilizing the results. Using a quantitative approach the following questions will guide this study; (a) to what extent are SONs systematically evaluating their online education activities at the master's degree program level; (b) what aspects of the program are being evaluated; (c) to what extent do organizational personal predictor factors affect evaluation practice and use of SONs online graduate programs; and (d) to what extent are evaluation results utilized.

You have been invited to participate because you were designated by your chief nurse administrator as the contact faculty member responsible for or best able to answer questions regarding program evaluation practices at your school of nursing.

Your participation will involve you responding to 81 items on a survey and should only take approximately 20 minutes of your time. Your involvement in the study is voluntary, and you may choose not to participate or to stop at any time without penalty or loss of benefits to which you are otherwise entitled. Your name and other records will be kept confidential to the fullest extent allowed. Only the research team will have access to the information you provide. Information may also be shared with those who make sure the study is done correctly (UGA IRB). Data will be labeled with a code rather than your name. The key to the code will be in a password protected file. The coded data file will be maintained on a separate computer in a locked cabinet in the researcher's office. The results of the research study may be published, but your name will not be used. In fact, the published results will be presented in summary form only. Your identity will not be associated with your responses in any published format.

The findings from this project may provide information on state of the art online program evaluation. Schools of nursing who wish to establish or refine their evaluation system will be able to see what is being done at a majority of schools and what is being done at particularly excellent schools. There are no known risks or discomforts associated with this research.

If you have any questions about this research project, please feel free to call me at 770-761-8310 or Dr. Lorilee Sandmann at 706-542-4014 or send an e-mail to emh56@uga.edu. Questions or concerns about your rights as a research participant should be directed to The Chairperson, University of Georgia Institutional Review Board, 629 Boyd GSRC, Athens, Georgia 30602; telephone (706) 542-3199; email address irb@uga.edu.

By clicking on the "next" option at the end of the informed consent you are agreeing to participate in the research project and will gain access to the questionnaire.

Thank you for your consideration! Please keep this letter for your records.

Sincerely,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
850 College Station RD
Athens, GA 30602

Lorilee R. Sandmann Ph.D.
Professor and Program Chair, Adult Education
The University of Georgia
Athens, GA 30602

Survey 1

Introduction: In this questionnaire we have identified five main constructs that provide guidelines for good systematic program evaluation of online education. Online education for this study is described as web-enhanced (partially online) **and/or** web-based (totally online) teaching/learning. Web-enhanced education is defined as teaching/learning events that combine aspects of online and face-to-face education or is partially online. Web-based education is defined as no face-to-face meeting with the instructor and the teaching/learning experience is totally online.

Does your school of nursing provide online (web-enhanced and/or web-based) educational activities at the master's degree level?

Yes

No

Instructions: One of your responsibilities as chief nurse administrator or master's degree level program director/coordinator is to systematically evaluate your school of nursing (SON) online education program at regularly designated intervals.

Below you will find a list of activities and practices used to evaluate online education programs. *We are interested in **how many times** within the **past 5 years** did your school of nursing engage or participate in each of the following practices of systematic program evaluation of online (web-enhanced and/or web-based) master's degree level nursing programs. As you complete the survey please base your responses on program evaluation in your current school of nursing. For sections 1-6 choose only one answer for each response.*

Never

1 or 2 times

3 or 4 times

5 or more times

Section 1: Collecting data from students:

Collecting data from students about online learning needs (e.g., needs assessment)

Collecting data from students about technological knowledge in preparation for learning online

Collecting data from students about satisfaction with technology support for online learning

Collecting data from students about satisfaction with the admission process to online programs

Collecting data from students about adequacy of the technology learning system
(Blackboard, Web-ct, eLive, etc) used for online learning

Collecting data from students about perceptions of the online learning social environment

Collecting data from students about satisfaction with online learning

Collecting data from students about perceptions of teaching effectiveness

Collecting data from students about satisfaction with online library support and access

Collecting data from students about perceived quality of academic advisement

Collecting data from students about perceived quality of financial aid support

Collecting data from students about cost-benefit of online learning

Collecting data from students about perceived value of online course content

Section 2: Collecting data from faculty:

Collecting data from faculty about online teaching materials being up to date

Collecting data from faculty about technological knowledge in preparation to teach online

Collecting data from faculty about student technological preparation for online learning

Collecting data from faculty about responsiveness to students' needs in the teaching/learning environment

Collecting data from faculty about control over online course content

Collecting data from faculty about satisfaction with the courseware used for online education (e.g., blackboard, Web-ct, etc.)

Collecting data from faculty about satisfaction with university technology support for online education

Collecting data from faculty about reasonableness of workload when teaching online

Collecting data from faculty about instructor satisfaction with student learning

Section 3: Compiling and analyzing data from program records:

Compiling and analyzing data from program records about applicants' standardized test score results (e.g., GRE, MAT)

Compiling and analyzing data from program records about applicants' prior academic achievement results (e.g., GPA)

Compiling and analyzing data from program records about demographic data on students applying to online programs

Compiling and analyzing data from program records about compliance of online programs with requirements from certifying bodies

Compiling and analyzing data from program records about comparative cost effectiveness of online versus face-to-face instruction

Compiling and analyzing data from program records about comparisons between online versus traditional end-of-course evaluations

Compiling and analyzing data from program records about trends in online course enrollment

Compiling and analyzing data from program records about online course failure rates

Compiling and analyzing data from program records about online course completion/withdrawal rates

Compiling and analyzing data from program records about program enrollment rates

Compiling and analyzing data from program records about program retention rates

Compiling and analyzing data from program records about program graduation rates

Compiling and analyzing data from program records about national certification pass rates

Compiling and analyzing data from program records about faculty qualifications to teach online

Compiling and analyzing data from program records about faculty retention rates

Compiling and analyzing data from program records about staff retention rates

Compiling and analyzing data from program records about structures hindering effectiveness of online programs

Compiling and analyzing data from program records about policies hindering effectiveness of online programs

Compiling and analyzing data from program records about policies hindering transferability of online coursework

Compiling and analyzing data from program records about faculty selection process to teach online

Section 4: Collecting data from alumni:

Collecting data from alumni about employment rates

Collecting data from alumni about satisfaction with online teaching/learning

Collecting data from alumni about satisfaction with the degree obtained

Collecting data from alumni about satisfaction with skills obtained

Section 5: Collecting data from employers:

Collecting data from employers about satisfaction with overall performance of graduates from online education programs

Collecting data from employers about recurrent problems of graduates from online education programs

In order to get the data mentioned in the above five sections your school of nursing most likely needed to be involved in a number of specific activities in planning and implementing the evaluation. We are interested in to what extent does your school of nursing engage in the following program evaluation practices.

Strongly disagree Disagree Agree Strongly Agree

Section 6: My school of nursing:

My school of nursing identifies stakeholders in the evaluation process

My school of nursing assesses program readiness for evaluation

My school of nursing identifies questions to be answered by the evaluation process

My school of nursing uses evaluation methods that are appropriate for the questions identified (e.g., qualitative, quantitative, mixed methods)

My school of nursing keeps stakeholders informed of interim evaluation findings

My school of nursing decides how evaluation results are to be disseminated to the stakeholders

My school of nursing reviews evaluation results with faculty to identify best practices in online education

My school of nursing reviews evaluation results with staff to identify best practices in online education

My school of nursing uses evaluation data to monitor the program's performance at regular intervals

My school of nursing uses evaluation data to improve (make better) the online program

My school of nursing uses evaluation data to render judgment of the program (i.e., determine the overall merit, worth, significance, or value)

My school of nursing uses evaluation data for accountability to someone or something (e.g., an accrediting or certifying body, funding organizations)

My school of nursing *uses evaluation data to generate knowledge about the program (i.e., to identify patterns of effectiveness, trends in participation, general problems with online learning)*

My school of nursing *uses evaluation data to make developmental changes (make changes) to the program (i.e., alternative course sequencing, innovative delivery)*

The final section of the survey is designed to collect information about you and your school of nursing (SON).

Section 7: Background information:

63. Which of the following best describes your SON? (*Check one*)

Public

Private

64. From which of the following does your master's degree level program hold accreditation? (*Check all that apply*)

NLN

CCNE

Other_____ (please specify)

65. Approximately how many students are in your master's degree level nursing program? _____

66. Of all the students in your master's degree level program approximately what percentage are full time students? _____

67. Of all the courses taught in your master's degree level program approximately what percentage are taught by part time faculty? _____

68. Approximately how many years have your master's degree level program been in existence?

69. Approximately how many years have your master's degree level program used online course ware (e.g., Web-ct, blackboard, eLive, etc.)?

70. Which of the following best describes the method of online education delivery provided at the master's degree level? (*Check all that apply*)

Web enhanced or partially online

Web based or fully online

Other_____ (please specify)

71. Which of the following best describes where online education is used in the master's degree level curriculum? (*Check all that apply*)

Core curriculum courses

Specialty tract curriculum courses

72. During the past 2 years has tuition for your organization:

Increased

Decreased

Stayed the same

73. Indicate the **most recent year** in which your school of nursing evaluated their online master's degree level program. _____

74. Approximately how many **total** years have you been a nurse educator? _____

75. What is your current administrative job title? _____

76. Approximately how long have you held this current job title? _____

77. In what year were you born? _____

78. What is your gender? _____

79. What is your race/ethnicity? _____

Section 8: Overall reaction to the survey:

80. Did you have any trouble completing this survey? If so, please describe.

81. Were there any problems items that needed to be improved?

Week three of Pilot – Second reminder

Dear [Nurse Administrator]:

An online survey was sent to you on May 22, 2012 seeking your assistance in piloting the survey instrument. We would like to know perspectives regarding program evaluation practice and use of online nursing education at the master's degree level. Your willingness to share your expertise provides important information into better understanding these practices.

If you have not yet responded, we look forward to your participation in this 20 minute survey. We invite you to complete the pilot survey by following the link below for online completion. On the first page of the survey (via survey monkey) is the consent form. By clicking on the "next" option at the end of the informed consent you are agreeing to participate in the research project. The survey works best in Internet Explorer browser. Link:

This link is uniquely tied to this survey and your email address; please do not forward the message for other individuals to complete. We ask that if you are not the right contact, would you kindly provide the designated faculty's contact information by return email.

Name:

Title:

Email:

Phone:

As an alternative or supplement to the online survey, a PDF file has been created. This file will allow you to complete or simply view the PDF version of the survey. A PDF version can be obtained by emailing emh56@uga.edu. If you decide to complete the survey in the paper form please send your response to the address below and include your contact information as a separate document in order to allow me to update the database.

Thank you for contributing your time and insight to our PILOT study regarding practice and use of program evaluation of online nursing education at the master's degree level. We look forward to your response!

Respectfully,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602

Lorilee R. Sandmann Ph.D.
Professor and Program Chair, Adult Education
Department of Lifelong Education, Administration, & Policy
The University of Georgia

Please note: If you do not wish to receive further emails from us, please click the following link and you will be automatically removed from our mailing list: [remove link]

Week Five of Pilot – Third reminder

Dear [Nurse Administrator]:

An online survey was sent to you asking for your valuable input into investigating the extent of systematic program evaluation of online nursing education at the master's degree level. WE NEED YOUR ASSISTENCE SPECIFICALLY IN PILOTING THE SURVEY INSTRUMENT FOR THE STUDY. To the best of our knowledge, the survey has not yet been completed. If you feel our records are incorrect please respond directly to emh56@uga.edu.

We are continuing to collect data for our study on this very important subject of program evaluation. Therefore, we are contacting you again for your response in order to collect meaningful data and provide you an opportunity to contribute to the success of this study. The survey is to be completed based on your knowledge of program evaluation practices within your schools of nursing. This 20 minute survey can be accessed through the link below.

As an alternative or supplement to the online survey, a PDF file has been created. This file will allow you to complete or simply view the PDF version of the survey. If you decide to complete the survey in the paper form please send your response to the address below and include your contact information as a separate document in order to allow me to update the database. A PDF version can be obtained by email: emh56@uga.edu.

Thank you for contributing your time and insight to our PILOT study regarding practices and utilization of program evaluation of online nursing education at the master's degree level. For those who have not yet responded, we look forward to your response.

Respectfully,

Eva Horne MN, BSN, FNP-BC
The University of Georgia
Graduate Student, Adult Education
416 Rivers Crossing 850 College Station RD, Athens, GA 30602
emh56@uga.edu

Lorilee R. Sandmann Ph.D.
Professor and Program Chair, Adult Education
Department of Lifelong Education, Administration, & Policy
The University of Georgia

APPENDIX I

TESTING OF PILOT INSTRUMENT DISTRIBUTION AND RELIABILITY OF KEY MEASURES

Distribution and Reliability of Key Measures

Scale	Number of Items	M	SD	Mean Item Means	Alpha
Source of Evaluation					
Students	13	38.05	9.67	2.927	.91
Faculty	9	22.37	9.66	2.49	.96
Records	20	51.50	13.50	2.57	.98
Alumni	4	11.93	18.35	2.98	.96
Employers	2	4.93	2.01	2.46	.84
Area of Evaluation					
Input	20	50.20	14.98	2.51	.92
Process	11	31.93	7.66	2.90	.87
Output	13	36.78	7.50	2.83	.78
Impact	4	10.73	3.49	2.68	.79
Utilization of Evaluation					
Use	14	46.73	6.55	3.33	.91

APPENDIX J

RESPONDENT ADMINISTRATIVE JOB TITLES

Current Administrative Job Titles of Respondents

Title
Institutional Specialist
Director MSN
Chair, School of Nursing
Assistant Professor
Dean
Department Chair
Dean
Associate Dean Academic Affairs
Chair
Interim Chair and Graduate Program Director
Director, School of Nursing
Chair
Interim Associate Dean
Graduate Nursing Program Director and Assistant Professor
Graduate Nursing Program Coordinator
Dean
Dean
Chair of Department
Associate Dean Academic Affairs
Dean
Director MSN Program
Director, MSN-Nurse Educator Program
Chair, Graduate Nursing
Dean
Associate Dean
Director, school of nursing
Associate Dean
Dean and Professor
Prof and Dean
Graduate Chair
Associate Dean/Director, SON
Director of Nursing
Dean
Dean
Director
Program Coordinator
Chair
Dean
Dean
Associate Dean
Associate Dean for Graduate Nursing Programs
Professor

Dean and Professor
 Director, School of Nursing
 Graduate Chair
 Associate Dean Grad Programs
 Director of Graduate Nursing Programs
 Assistant Dean for Academic Affairs
 Interim Dean
 Associate Dean
 Associate Dean
 Director MSN Program
 Assistant Dean
 Director Masters and DNP Programs
 Associate Dean Graduate Programs
 Associate Dean
 Master's Program Coordinator
 Dean, College of Health Professions
 MSN Program Director
 Associate Dean
 Associate Professor Nursing/MSN Advisor
 Associate Dean
 Professor, Chair, Department Nursing Health
 Associate Dean for Graduate Nursing Programs
 Director, Grad Nursing
 Associate Dean
 Professor of Nursing
 Director of Graduate Nursing Program
 Program Director (MSN)
 Associate Dean Graduate Programs
 Dean
 Dean
 Coordinator MS FNP Program
 Dean
 MSN Program Director
 Professor of Nursing and Department Chair
 Associate Vice Chancellor for Nursing and Founding Dean
 Assistant Dean
 Assessment Coordinator
 Chair, Department of Nursing
 Dean
 Dean
 Director School of Nursing
 Program Director
 Acting Dean and Professor
 Associate Dean for Academics
 Director, School of Nursing
 Director of Distance Education

Director of Graduate Nursing
Assistant Professor - Graduate Coordinator
Associate Dean for Research
Director, Graduate Nursing Studies
Director of the Online RN to MSN Program
Associate Dean
Director Educational Strategies, Clinical Instructor
Chairperson, Department of Nursing
Graduate program Director
Chair
Director of the School of Nursing
Chair Graduate Nursing
Master's Coordinator

APPENDIX K

FREQUENCY TABLE – EVALUATING PRACTICE ITEMS

Frequency Table - Evaluating Practice Items

Items	n	Mean	SD	Rank of Means	Frequencies			
					Never	1 or 2	3 or 4	5 or more
8.Collecting data from students about perceptions of teaching effectiveness	103	3.64	.72	1	3/ 2.8%	6/ 5.6%	17/ 15.9%	79/ 73.8%
7.Collecting data from students about satisfaction with online learning	103	3.35	.87	2	3/ 2.8%	19 17.8%	21/ 19.6%	62/ 57.9%
3.Collecting data from students about satisfaction with technology support for online learning	104	3.33	.94	3	6/ 5.6%	17/ 15.9%	19/ 17.8%	64/ 59.8%
5.Collecting data from students about adequacy of the technology learning system (e.g., Blackboard, eLive, etc.) used for online learning	104	3.27	.95	4	8/ 7.5%	13/ 12.1%	27/ 25.2%	58/ 54.2%
32.Compiling and analyzing data from program records about program enrollment rates	100	3.10	1.06	5	11/ 10.3%	20/ 18.7%	19/ 17.8%	52/ 48.6%
34.Compiling and analyzing data from program records about program graduation rates	98	3.09	1.04	6	9/ 8.4%	22/ 20.6%	20/ 18.7%	49/ 45.8%
45.Collecting data from alumni about satisfaction with the degree obtained	102	2.99	1.06	7.5	9/ 8.4%	31/ 29.0%	16/ 15.0%	48/ 44.9%
9.Collecting data from students about satisfaction with online library support and access	102	2.99	1.03	7.5	10/ 9.55	25/ 23.4%	25/ 23.4	44/ 41.1%
46.Collecting data from alumni about satisfaction with skills obtained	102	2.97	1.04	9.5	9/ 8.4%	30/ 28.0%	20/ 18.7%	45/ 42.1%
43.Collecting data from alumni about employment rates	102	2.97	.98	9.5	6/ 5.6%	33/ 30.8%	23/ 21.5%	42/ 39.3%
33.Compiling and analyzing data from program records about program retention rates	99	2.95	1.09	11	13/ 12.1%	23/ 21.5%	21/ 19.6%	44/ 41.1%

13.Collecting data from students about perceived value of online course content	101	2.92	1.22	12	22/ 20.6%	14/ 13.1%	17/ 15.9%	50/ 46.7%
10.Collecting data from students about perceived quality of academic advisement	103	2.88	1.14	13	19/ 17.8%	18/ 16.8%	25/ 23.4%	43/ 40.2%
35.Compiling and analyzing data from program records about national certification pass rates	101	2.84	1.25	14	24/ 22.4%	17/ 15.9%	13/ 12.1%	49/ 45.8%
26.Compiling and analyzing data from program records about compliance of our online program with requirements from certifying bodies	99	2.77	1.14	15.5	19/ 17.8%	23/ 21.5%	21/ 19.6%	38/ 35.5%
24.Compiling and analyzing data from program records about applicants' prior academic achievement results (e.g., GPA)	102	2.77	1.18	15.5	21/ 19.6%	24/ 22.4%	17/ 15.9%	42/ 39.3%
25.Compiling and analyzing data from program records about demographic data on students applying to our online program	100	2.74	1.22	17	26/ 24.3%	14/ 13.1%	23/ 21.5%	39/ 36.4%
20.Collecting data from faculty about satisfaction with university technology support for online education	102	2.73	1.08	18.5	17/ 15.9%	27/ 25.2%	27/ 25.2%	33/ 30.8%
1.Collecting data from students about online learning needs (e.g., needs assessment)	102	2.73	1.16	18.5	22/ 20.6%	21/ 29.6%	24/ 22.4%	37/ 34.6%
44.Collecting data from alumni about satisfaction with online teaching/learning	102	2.71	1.15	20	19/ 17.8%	30/ 28.0%	17/ 15.9%	38/ 35.5%
6.Collecting data from students about perceptions of the online learning social environment	102	2.65	1.17	21	24/ 22.4%	23/ 21.5%	22/ 20.6%	35/ 32.7%
14.Collecting data from faculty about online teaching materials being up to date	102	2.62	1.23	22	28/ 26.2%	22/ 20.6%	16/ 15.0%	38/ 35.5%
49.Collecting data from employers about satisfaction with overall performance of graduates from our online education program	101	2.60	1.14	23	21/ 19.6%	32/ 29.9%	17/ 15.9%	33/ 30.8%

19. Collecting data from faculty about satisfaction with the courseware used for online education (e.g., Blackboard, eLive, etc.)	102	2.58	1.10	24	21/ 19.6%	31/ 29.0%	23/ 21.5%	29/ 27.1%
31.Compiling and analyzing data from program records about online course completion/withdrawal rates	99	2.57	1.24	25	28/ 26.2%	23/ 21.5%	14/ 13.1%	36/ 33.6%
2.Collecting data from students about technological knowledge in preparation for learning online	103	2.55	1.18	26	27/ 25.2%	25/ 23.4%	21/ 19.6%	32/ 29.9%
17.Collecting data from faculty about responsiveness to students' needs in the teaching/learning environment	102	2.53	1.19	27	29/ 27.1%	22/ 20.6%	22/ 20.6%	31/ 29.0%
36.Compiling and analyzing data from program records about faculty qualifications to teach online	101	2.48	1.17	28	26/ 24.3%	33/ 30.8%	13/ 12.1%	31/ 29.0%
15.Collecting data from faculty about technological knowledge in preparation to teach online	101	2.47	1.15	29	27/ 25.2%	28/ 26.2%	21/ 19.6%	27/ 25.2%
51.Collecting data from employers about the professionalism of graduates from our online education program	99	2.41	1.19	30.5	30/ 28.0%	29/ 27.1%	13/ 12.1%	29/ 27.1%
30.Compiling and analyzing data from program records about online course failure rates	99	2.41	1.26	30.5	35/ 32.7%	22/ 20.6%	12/ 11.2%	32/ 29.9%
4.Collecting data from students about satisfaction with the admission process to our online program	101	2.39	1.25	32	36/ 33.6%	22/ 20.6%	14/ 13.1%	31/ 29.0%
52.Collecting data from employers about essential educational requirements needed by graduates from our online education program	101	2.38	1.21	33	33/ 30.8%	28/ 26.2%	12/ 11.2%	30/ 28.0%
29.Compiling and analyzing data from program records about trends in online course enrollment	98	2.37	1.22	34	35/ 32.7%	20/ 18.7%	18/ 16.8%	27/ 25.2%

11.Collecting data from students about perceived quality of financial aid support	101	2.35	1.23	35	37/ 34.6%	22/ 20.6%	15/ 14.0%	29/ 27.1%
21.Collecting data from faculty about reasonableness of workload when teaching online	101	2.29	1.13	36	30/ 28.0%	38/ 35.5%	10/ 9.3%	25/ 23.4%
22.Collecting data from faculty about instructor satisfaction with student learning	102	2.27	1.12	37	33/ 30.8%	31/ 29.0%	19/ 17.8%	21/ 19.6%
18.Collecting data from faculty about control over online course content	101	2.21	1.27	38	41/ 38.3%	26/ 24.3%	9/ 8.4%	27/ 25.2%
23.Compiling and analyzing data from program records about applicants' standardized test score results (e.g., GRE, MAT)	102	2.14	1.22	39	46/ 43.0%	21/ 19.6%	13/ 12.1%	24/ 22.4%
39.Compiling and analyzing data from program records about structures hindering effectiveness of our online program	102	2.13	1.11	40	39/ 36.4%	30/ 28.0%	17/ 15.9%	18/ 16.8%
50.Collecting data from employers about common problems of graduates from our online education program	101	2.12	1.22	41.5	47/ 43.9%	21/ 19.6%	11/ 10.3%	24/ 22.4%
40.Compiling and analyzing data from program records about policies hindering effectiveness of our online program	102	2.12	1.09	41.5	40/ 37.4%	28/ 26.2%	20/ 18.7%	16/ 15.0%
16.Collecting data from faculty about student technological preparation for online learning	101	2.10	1.13	43	42/ 39.3%	27/ 25.2%	16/ 15.0%	18/ 16.8%
53.Collecting data from employers about job retention rates of graduates from our online education program	100	1.94	1.19	44	55/ 51.4%	18/ 16.8%	9/ 8.4%	20/ 18.7%
37.Compiling and analyzing data from program records about faculty retention rates	101	1.84	1.04	45	53/ 49.5%	24/ 22.4%	15/ 14.0%	11/ 10.3%
28.Compiling and analyzing data from program records about comparisons between online versus traditional end-of-course evaluations	101	1.71	.94	46.5	55/ 51.4%	33/ 30.8%	5/ 4.7%	10/ 9.3%

42.Compiling and analyzing data from program records about faculty selection process to teach online	100	1.71	1.03	46.5	61/ 57.0%	22/ 20.6%	7/ 6.5%	12/ 11.2%
38.Compiling and analyzing data from program records about staff retention rates	98	1.67	.99	48.5	60/ 56.1%	23/ 21.5%	7/ 6.5%	10/ 9.3%
12.Collecting data from students about cost-benefit of online learning	101	1.67	1.03	48.5	65/ 60.7%	19/ 17.8%	7/ 6.5%	12/ 11.2%
47.Collecting data from alumni about reduction of learning needs	93	1.65	1.07	50	62/ 57.9%	16/ 15.0%	3/ 2.8%	13/ 12.1%
27.Compiling and analyzing data from program records about comparative cost effectiveness of online versus face-to-face instruction	98	1.62	.92	51	59/ 55.1%	29/ 27.1%	3/ 2.8%	9/ 8.4%
48.Collecting data from alumni about elimination of learning needs	93	1.60	1.01	52	63/ 58.9%	17/ 15.9%	3/ 2.8%	11/ 10.3%
41.Compiling and analyzing data from program records about policies hindering transferability of online coursework	100	1.59	.97	53	68/ 63.6%	18/ 16.8%	6/ 5.6%	10/ 9.3%