Organization, Information and Learning Sciences ETDs

Electronic Theses and Dissertations

2-1-2016

Bibliometrics and Social Network Analysis of Doctoral Research: Research Trends In Distance Learning

Jason Skinner

Follow this and additional works at: https://digitalrepository.unm.edu/oils_etds

Jason Skinner

Candidate

Organization Information and Learning Sciences

Department

This dissertation is approved, and it is acceptable in quality and form for publication:

Approved by the Dissertation Committee:

Dr. Charlotte N. (Lani) Gunawardena , Chairperson

Dr. William Bramble

Dr. Victor Law

Dr. Vincent Flor

BIBLIOMETRIC AND SOCIAL NETWORK ANALYSIS OF DOCTORAL

RESEARCH: RESEARCH TRENDS IN DISTANCE LEARNING

BY

JASON KIRTLAND SKINNER

B.A., English, University of Oregon, Eugene, OR., 1999
M.A., Curriculum Teaching Learning, Sonoma State University,
Sonoma, CA., 2006

DISSERTATION

Submitted in Partial Fulfillment of the
Requirements for the Degree of

Doctor of Philosophy
Organizational Learning and Instructional Technology

The University of New Mexico
Albuquerque, New Mexico

December 2015

DEDICATION

The work found in this study, as well as all of the learning that coincided, is dedicated to my love Stephanie. I am filled with the utmost gratitude for the years of love and support. Come waste your time with me.

I also send profound love and thanks to my son Cedar who is joyfully becoming a voracious lover of learning and liver of life. And to my parents, thank you for instilling a hearty appetite for knowledge.

ACKNOWLEDGMENTS

Dr. Lani Gunawardena, my advisor and dissertation chair encouraged and inspired me through both my course and dissertation work. I am very grateful for her support. The worldview that she has helped me to co-construct will remain a guiding force in my professional work to come.

I also thank my committee members, Dr. Bill Bramble, Dr. Victor Law, and Dr. Nick Flor, for their time and valuable advisement during my degree studies and dissertation research.

Gratitude is also extended to Jane Lorand and Bruce McKenzie. Many thanks for the support and opportunity to join the Systemic and Critical Thinking work at Future Insight Maps, Inc.

BIBLIOMETRIC AND SOCIAL NETWORK ANALYSIS OF DOCTORAL

RESEARCH: RESEARCH TRENDS IN DISTANCE LEARNING

by

Jason Kirtland Skinner

B.A., English, University of Oregon, Eugene, OR., 1999
M.A., Curriculum Teaching Learning, Sonoma State University,
Sonoma, CA., 2006
Ph.D, Organizational Learning and Instructional Technology, Albuquerque, NM,
2015

ABSTRACT

The study investigated research topics of doctoral dissertations that examined

issues in distance learning from 2000-2014. Twelve reviews of research on distance

learning, spanning from 1997-2015, were identified. It was found that only one of these

reviews of research (Davies, Howell, & Petri, 2010) looked at doctoral dissertations. The

authors noted that investigating dissertations was complicated and daunting because 1)

only a fraction made full text available and 2) there were a large number of dissertations

in the area. To counter for these complications the current study utilized bibliometric and

social network analysis to investigate dissertation database listings, including abstracts,

keywords, classifications, and other bibliographic data.

Bibliographic data for dissertation listings (n=3,954) was exported from the

ProQuest Dissertations & Theses A&I (PQDT) database. Software developed for the

study formatted the data and imported it into a series of databases. Natural language

processing techniques were utilized to pull emergent keywords from dissertation

abstracts. Department and University types were analyzed. Dissertation reference

sections were investigated utilizing co-citation analysis. Author generated keywords and

emergent keywords from abstracts were investigated utilizing keyword co-occurrence network analysis.

Findings indicated that dissertations came from 17 department types including education-oriented department types, such as Educational Leadership, Educational Technology, and Educational Psychology, as well as non-education-oriented departments, such as Business, Psychology, and Nursing. Seven research topics were found to be pervasive in dissertations from 2000-2014: Student, Instructor, Interaction, Administration and Management, Design, Educational Context, and Technological Medium. No change was found over time; rather these seven topics remained the most central nodes in each of the keyword co-occurrence networks.   Finally this method of investigation relied heavily on algorithms developed for the study to aid in data formatting and analysis. The merits of this highly automated SNA approach were discussed. Use of abstracts and natural language processing enabled a much higher n size (n=3954) to be investigated than in comparison with the only other study to analyze distance education dissertations Davies et al. (2010) where n=100. This method enabled the heavy lifting to be dedicated to the interpretation of the results, rather than data preparation.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER I

INTRODUCTION

Background and Context for the Study

Distance Education

Distance education is not a 21st century phenomenon. "Distance education has

been a part of the educational experience for well over 100 years, with the introduction of

print-based correspondence study in the late 1800s" (Fudell & Hardy, 1998, p. 1).

Garrison (1985) identified three generations of distance education: the *Correspondence*

Generation, the *Telecommunication Generation*, and the *Computer Generation*.

The Correspondence Generation relied on the postal system to deliver the printed

learning materials and written communications. The Telecommunication Generation

added voice, video, and data to their modes of communication. Although Garrison made

these distinctions 30 years ago, one may argue that we still reside in the Computer

Generation.

Ritzhaupt, Stewart, Smith, and Barron (2010) further delimited the timeline of

distance education into three distinct time periods within the computer generation that are

defined by the World Wide Web: the *Pre-Web*, *Emerging Web*, and *Maturing Web*. The

authors used the first publication dates for the *American Journal of Distance Education*

(ADJE) 1987 and the *Journal of Distance Education* (JDE) 1986 as the beginning of the

current era of distance education research. The Pre-Web period, 1987 – 1993, marked

the interval in which the Web was not significant to distance education and largely not

available to researchers. The emergence of the Web initially impacted distance education

from 1994 – 1999 when learning management systems such as WebCt first became

available. The third time period, since 2000, has seen the maturation of web-based distance education and subsequently its effect on distance education research. Ritzhaupt et al. (2010) investigated distance education research through 2005 and so end their definition of the maturing Web period at that point. Yet since 2005 new technologies that have influenced distance education, such as desktop video conferencing technologies, mobile technologies, and cloud-based computing, have emerged. Therefore the Maturing Web time period of distance education, which is the time period of focus for this study, is here characterized as 2000 to the present.

Reviews of Distance Education Research

Researchers recognize that analysis and understanding of trends and issues found in the literature base are critical for the advancement of distance education research (Lee, Driscoll, & Nelson, 2004). Reviews of research provide a broad systems-based view of the state of a field by reporting key contributors, methodological inclinations, and areas of focus over time (Anglin, & Morrison, 2000). Therefore an important starting point for considering the status of distance education research is to review what has been put forth in scholarly journals and to review dissertations (Moore & Kearsley, 2005). During the past two decades several articles have reviewed findings in distance education literature (Koble & Bunker, 1997; Anglin, & Morrison, 2000; Berge & Mrozowski, 2001; Rourke & Szabo, 2002; Lee et al., 2004; Zawacki-Richter, Bäcker, & Vogt, 2009; Ritzhaupt, et al. 2010; Davies, Howell, & Petri, 2010, Zawacki-Richter & Anderson, 2011; Cho, Park, Jo, & Suh, 2013; Zawacki-Richter & Anderson, 2014; and Bozkurt et al., 2015).

Koble and Bunker (1997) looked at trends in the *ADJE* during the Pre-Web time period of 1987 – 1995. Berge and Mrozowski's (2001) large study reviewed articles

from four distance education research journals and dissertations from 1990 – 1999 based on the categorization system that originally came from Sherry's (1995) study.

Zawacki-Richter, Bäcker, and Vogt (2009b), continuing from Berge and Mrozowski's (2001) study end point, reviewed articles published between 2000 – 2008 categorizing research topics with their Classification of Research Areas in Distance Education (CRADE) framework. Bozkurt et al. (2015) reviewed research topics in articles ranging from 2009 – 2013 using both the CRADE framework and Social Network Analysis (SNA).

Studies that review distance education research have largely focused on articles from one or a combination of five journals: *Open Learning (OL), Distance Education (DE), ADJE, JDE, and the International Review of Research In Open and Distance Learning (IRRODL)*. The research trends observed in studies from these journals often become the basis for more in depth reviews included in handbooks of research such as the Handbook of Distance Education (Moore ed., 2013). This is the case for the Zawacki-Richter (2009a) study that forwarded the CRADE framework. Zawacki-Richter and Anderson (2013) followed with an edited book using the research areas covered by the CRADE as the chapter outline for more in-depth reviews of research in each area.

Since the objective of reviews of research studies is often to investigate research topic trends reviews usually include some categorization framework of research methods. For example, Berge and Mrozowski (2001) classified articles into Phipps and Merisotis's (1999) four research methods: descriptive research, case study, correlational research, and experimental research.

Reviewers also utilize a variety of frameworks to classify research topic areas. For example, Zawacki-Richter (2009a) utilized the CRADE framework that emerged from a Delphi study asking 25 distance education researchers from 11 countries to classify common areas that are or should be covered in distance education research. From these responses three broad meta-levels of distance education research were derived:

Macro Level: Distance education systems and theories

Meso Level: Management, organization and technology

Micro Level: Teaching and learning in distance education

(Zawacki-Richter, 2009a, p. 22).

Within these three levels fifteen sub-categories of research areas were identified. Zawacki-Richter et al. (2009b) used these 15 sub-categories as the framework for a content analysis of distance education topics and as the basis for a handbook of research (Zawacki-Richter & Anderson, 2014).

Content analysis is the most common approach researchers use to classify both the method and topic area within distance education research reviews (Koble and Bunker, 1997; Anglin, and Morrison, 2000; Berge and Mrozowski, 2001; Rourke and Szabo, 2002; Lee et al., 2004; Zawacki-Richter et al., 2009b). This approach enables multiple raters to counterbalance the subjectivity of coding and verify that reported data is accurate.

Ritzhaupt, et al. (2010) found that using a SNA approach enabled research topics and themes to objectively emerge from the data and not be bound to the categories of an a priori framework. Cho et al. (2013) also used a SNA approach without a preexisting

research topic framework and report five research themes that emerged from their citation network analysis. Bozkurt,(2015) utilized SNA to let research topics emerge from a keyword co-occurrence network. They grounded these findings by comparing the results to Zawacki-Richter (2009a) CRADE framework. The current study will use a similar approach investigating research topics that emerge from keyword co-occurrence networks. The results will be compared to the CRADE and similarities or differences discussed.

Reviews of Graduate Research on Distance Education

Abstracts to graduate dissertations and theses are included in Berge and Mrozowski's (2001) review of distance education research that covers Pre-Web through the Emerging Web periods, 1990 – 1999. But there has been little investigation into graduate research in the field during the Maturing Web period since 2000.

Davies et al. (2010) reviewed a sample of 100 dissertations and theses from three years 1998, 2002, and 2007. Davies' team analyzed the most commonly addressed topics and methods utilized by graduate researchers to determine changes over the investigated time period. A content analysis approach was used and topics were coded to Lindsay, Wright, and Howell's (2004) research topic framework. The authors found that across time dissertations focusing on face-to-face versus distance learning and quality of technology had decreased.

Moore and Kearsley (2005) point out that it is difficult to access graduate student research on distance education. Davies et al. (2010) noted that dissertation databases make the task possible, but that full-text access to manuscripts is often not available. At the time of this study this remains largely true. Yet as shown in Berge and Mrozowski

(2001), Zawacki-Richter & Anderson (2011), Cho, Park, Jo, & Suh (2013), and Bozkurt (2015), the bibliometric data that are available in dissertation databases, including abstracts, keywords, subject classifications, and cited references, may provide for an interesting gaze into the landscape of graduate research in the field of distance education.

The current study will investigate English language doctoral dissertations that focus on distance education. The dissertation sample will be gathered from the ProQuest Dissertations & Theses A&I (PQDT) database. The PQDT indexes English language dissertations from around the world. It receives 97.2% of all dissertations and theses from research universities in the United States and 87.2% of those from Canadian research universities (Davies, 2010). This number decreases dramatically for dissertations in countries outside of North America. Of the initial 4027 dissertations listed in the PQDT considered for this study only 73 were from countries other than the United States and Canada. This limited listing of dissertations from countries like the UK and Australia was judged to not be representative of the full body of distance education focused dissertations from these countries for the past fifteen years and so will not be included in the study. Rather the current study will analyze distance education focused dissertations from North American institutions that were published in English. This delimitation means that the sample will include dissertations from the United States and Canada, as none from Mexico are listed in the PQDT.

Bibliometrics and Social Network Analysis

Bibliometrics is a field founded on the concept that citations can be used as indicators of past and present practices in scientific work (Lee & Su, 2010). It can be characterized as a quantitative method that uses statistics to analyze bibliographic

information found in publications (Holden & Barker, 2005). Traditional bibliometric measurements enable the analysis of trends through frequency ranking of document variables, such as author or publication title. For example, Zawacki-Richter, et al. (2009b) report bibliometric frequencies for: number of authors by year, male versus female researchers, and number of references. To investigate the characteristics of the graduate researcher sample addressed by this study bibliometric data including department and university will be used. To address citation trends dissertation reference citation sections will be utilized. Finally to address research topics of dissertations bibliometric data including titles, abstracts, classifications, and keyword data will be used. These bibliometric data points are similar to the previous studies that use bibliometric and SNA approaches (Ritzhaupt, et al. 2010; Zawacki-Richter & Anderson, 2011; Cho, Park, Jo, & Suh, 2013; Bozkurt et al., 2015).

Citation analysis inspects how many times a publication is cited in a sample of article reference sections. Some researchers work from the premise that heavily cited articles tend to have greater influence on the field than less frequently cited publications (Culnan, 1986; Sharplin & Marby, 2007). Keyword analysis examines the frequency of keywords assigned to a sample of texts. Keyword frequencies point towards topics and areas of interest that have greater influence on a field (Lee & Su, 2010). While frequency rankings reveal trends of high utilization in a field, Leong (1989) points out that frequency cannot describe the structure of *influence* within a field. To understand influence in a field an approach beyond frequency is needed to reveal the relationships and structural patterns between elements within a larger system.

Social network analysis enables one to investigate the relational and structural attributes of complex data groupings. Two SNA approaches are analysis of keyword co-occurrence networks and co-citation networks. They enable complex bibliometric data relationships to be investigated. Relationships can be described and analyzed mathematically by calculating an adjacency matrix for the data set. This will be further discussed in chapter II. There are also a number of toolsets that enable networks to be represented and investigated visually

Distance Education Research Utilizing Bibliometric and Social Network Analysis

Ritzhaupt et al. (2010) utilized a word co-occurrence network analysis technique to investigate topic trends and themes found in the *AJDE* and the *JDE*, between 1987 and 2005. The authors noted that previous studies that review distance education research utilize a priori sets of themes and topic frameworks, as well as subjective interpretation of the data through content analysis. They found that a SNA approach enables objective investigation where resultant themes and trends objectively emerge from the data. This is an important distinction from previous reviews of distance education literature. These studies often used content analysis methods and either developed or utilized pre-existing frameworks (Koble & Bunker, 1990, Berge & Mrozowski, 2002; Zawacki-Richter et al., 2009b) to categorize research article topics. Ritzhaupt et al. (2010) note that by letting topics emerge from the data through network analysis techniques the topics are not constrained to limitations or scope of pre-existing frameworks. The key difference of this approach is that topic categories not conceived of by the researcher before analysis have the opportunity to surface independently.

Cho et al. (2013) utilized SNA to investigate citations from distance education studies. The researchers coded studies into topic areas and were able to demonstrate the relationships between authors that have studied similar research topics. Bozkurt et al. (2015) also utilized SNA to investigate research topics and after analyzing the emergent results compared the findings back to the areas of research identified in the Zawacki-Richter (2009a) CRADE framework.

The present study takes its lead from Ritzhaupt el al. (2010) and will utilize SNA to investigate themes and topics that emerge from doctoral research in distance education published in English. This emergent approach will be of particular value to apply to the doctoral researcher sample. Since dissertations that focus on distance education are largely uninvestigated, topic areas that differ from those reported as occurring frequently in journal articles may emerge. A comparison of the emergent research areas results will be made with those found in the CRADE (Zawacki-Richter et al., 2009b; Zawacki-Richter & Anderson, 2014). As found in Bozkurt et al. (2015). This approach allows for research topics to emerge independently from the sample and then be compared back to the research areas found in an established and commonly utilized framework.

Purpose of the Study

The primary purpose of this study is to investigate the research topics found in doctoral dissertations addressing distance education from North American institutions published in English. A secondary purpose of this study is to assess how these topics have changed over time: 2000-2004, 2005-2009, and 2010-2014. These two general purposes will be addressed through two modes of investigation, bibliometrics and social network analysis. Bibliometric analysis will enable a picture of the doctoral researcher

sample to emerge. Frequency ranking will detail the universities and departments

research has come from, as well the most common topics investigated and sources cited.

SNA will enable investigation of the relationships of research topics through a keyword

co-occurrence network and relationships between citations through a co-citation network.

<div align="center">Research Questions</div>

General Research Questions

1. What research topics can be identified in doctoral dissertation research on

    distance education published in North America in English between 2000-

    2014?

2. How have the research topics changed over time, specifically from 2000-

    2004, 2005-2009, and 2010-2014?

3. How do the research topics compare to the research areas forwarded in

    Zawacki-Richter (2009a) CRADE framework?

Research Sub-Questions: Bibliometric and Citation Analysis

A. What departments did graduate researchers come from?

B. What universities did graduate researchers come from?

C. Which authors received the highest frequency of citation in dissertation

    reference sections?

D. What journal articles were cited with the highest frequency in dissertation

    reference sections?

E. What journal publications were cited with the highest frequency in

    dissertation reference sections?

F. What book titles were cited with the highest frequency in dissertation reference sections?

G. What are the relationships among dissertation reference citations using co-citation SNA?

Research Sub-Questions: Dissertation Research Topics

H. What dissertation database classifications received the highest frequency?

I. What dissertation keywords received the highest frequency?

J. What are the relationships among the topics identified in dissertation abstracts using co-occurrence SNA?

Theoretical Framework

Bibliometric studies operate with the theoretical perspective that examination of citations enable an understanding of growth of contributions within a field, the ability to identify when major publications were written, and how the popularity of the works fared over the period investigated (De Bellis, 2009). If publications continue to be cited over time a historical value is assumed and the work is considered to be a primary reference. Changes in citation rates over time enable scholars to identify when major changes in the field occur (Pilkington & Meredith, 2009). Likewise, from a bibliometric position the popularity of keywords indicates the influence of the concept or topic contained in those keywords on the field. A continuing use of a keyword across numerous works and time indicates important areas, concepts, or topics in a field. Changes in keyword usage over time indicates changes in research topics due to competing topics or paradigm shifts in the field (Ritzhaupt et al, 2010). From this foundation, this study will utilize a bibliometric perspective to investigate dissertation reference sections in order to identify

the primary authors, articles, and publications that doctoral students utilize to support their research in distance education. Also the keywords and classifications doctoral students chose to characterize their dissertations will be analyzed as indicators of popular research topics for doctoral dissertations. While a bibliometric approach enables key citations and research topics to be identified, it does not provide information on the relationships between the citations or research concepts.

A Social Network Analysis approach picks up this lack. SNA enables the relationships between elements in a network to emerge and thus provides a more holistic picture (Marin & Wellman, 2011). Network analysis (social network theory) is the study of how the social structure of relationships around a person, group, organization, or concept effects beliefs or behaviors (Kadushin, 2011). A social network analysis approach should be primarily conceived and investigated from the view of the properties of relations between and within units instead of the properties of these units themselves (Hanneman & Riddle, 2011). It is a relational approach. This study will utilize a co-citation network to investigate the relationships between studies and authors. This approach assumes that citations that occur in reference sections frequently are related (Moed, 2005). Through SNA calculations a more global picture of the relationships between studies can emerge. Likewise this study will utilize a co-occurrence network approach to calculate and visualize the relationships between keywords in order to analyze the topics that dissertations on distance education have addressed over the past fifteen years.

Method

A sample, n=3954, of dissertation records that address distance education and range between the years 2000 – 2014 will be gathered from ProQuest Dissertations & Theses A&I (PQDT) database. This large sample size makes traditional content analysis of dissertations daunting (Davies, 2010). However, utilizing a bibliometric and SNA approach to analyze bibliographic data makes it possible to investigate this large sample of dissertations. The data for the study include dissertation listings from the PQDT and bibliographic information held in each listing including the attributes university, department, title, author, classification, keywords, abstract, and reference citations when available

Distance education related dissertations are here characterized as having the following keywords in the title, abstract, or subject fields of the database record: "distance education", "distance learning", "online learning", "electronic learning", "e-learning", "network learning", "distributed learning" or "web-based learning".

The collected sample will be exported from the PQDT database. The raw data for the collected sample will then be imported into a database created specifically for this study. A series of web-based PHP algorithms will be created to query the study database to separate, format, and organize the data into a series of subsequent databases each designed to hold the information needed to get results for each of the study's research sub-questions. Natural language processing will be performed using the Python based Natural Language Tool Kit (NLTK) to standardize keywords and emergent keywords gathered from dissertation abstracts. A subsequent set of algorithms will be created to calculate the results for each of the bibliometric sub-questions.

An algorithm will be created to export citation and keyword data into the ISI format. This data will then be imported into the SITKIS (Schildt, 2002) bibliometric analysis application. This application will be used an intermediary program to create adjacency matrixes for network analysis. Once the adjacency matrixes are calculated they will be imported into the UCINET (Borgatti, Everett, & Freeman, 2002) SNA application. UCINET will be used to create the two networks and calculate network centrality statistics. Finally the NetDraw (Borgatti, Everett, & Freeman, 2002) application will be utilized to visualize the networks.

One network will investigate research topics through a co-occurrence network of emergent keywords gathered from dissertation abstracts. The other network, a co-citation network, will be based on dissertation reference sections. Only 10% of the retrieved dissertation listings include reference sections. Therefore the co-citation network will only include a small portion of the total sample.

Results will then be reported, analyzed, and discussed. A comparison of the research topics found in dissertations will then be compared with the research areas from Zawacki-Richter (2009a) CRADE framework. Findings and discussion will be presented in chapters IV and V.

Significance of the Study

This study looks to build on the work of Davies et al. (2010) by investigating graduate research in the field of distance education during the Maturing Web period, from 2000 to 2014. Reviews of the work contributed by doctoral student researchers are largely missing from the literature. This study contributes demographic information on

the distance education graduate researcher population, as well as provides an analysis of the areas and topics these researchers investigated during the past 15 years.

Davies et al. (2010) note that the time and effort involved in reading and categorizing the full text of thousands of dissertations with a content analysis approach is a formidable task. However, utilizing a bibliometric and SNA approach to analyze abstracts and bibliographic information makes investigation of this large sample a possibility.

By comparing the findings of this current study to the research areas forwarded by Zawacki-Richter (2009a) and expanded on in Zawacki-Richter & Anderson (2014) this study will compare its findings to a framework currently well regarded by the distance education research community.

Definitions

Distance Education: Schlosser and Simonson (2009) define distance education as "institution-based, formal education where the learning group is separated, and where interactive telecommunications systems are used to connect learners, resources, and instructors" (p1). Doctoral dissertations that address distance education are defined in this study as containing one or more of the following keywords in the title, keywords, classification, or abstract: "distance education" distance learning", "e-learning", "electronic leaning", "network learning", "distributed learning", or "web-based learning".

Graduate Student Research: "Graduate student research refers to research reported in master's theses and doctoral dissertations" (Davies et al., 2010, p. 42). This study includes only doctoral research and reference to graduate student research herein reflects this delimitation. Berge and Mrozowski point out that "acceptance criteria for

journals versus dissertations may be quite different," (2001, p.8). Unlike the peer review process, completion of a dissertation enables all completed research to be indexed and reported publically to dissertation databases.

Bibliographic data: "The author, title, place of publication, and other such information about a document" (Diodato, 1994, p. 13). This study utilizes bibliographic data from PQDT dissertation listings including university, degree date, and department.

Bibliometrics: "The application of measurements and statistics to study the bibliographic data of documents and publications," (Shimp, 2007, p.14). Citation analysis and keyword analysis are examples of bibliometric methods.

Social Network Analysis: "Social network analysis studies a set of socially relevant nodes linked by one or more relations," (Marin & Wellman, 2011). This approach requires a set of assumptions about how best we can describe and explore social phenomena of interest. Contrary to conventional quantitative analysis, SNA does not assume that environments, attributes, or contexts affect social actors independently. SNA is concerned about the interrelatedness of social elements or entities (Hanneman & Riddle, 2011).

<div align="center">Delimitations of the Study</div>

The scope of the dissertation research examined in this study will be confined to English language doctoral dissertations completed in North America between the years 2000 – 2014. Masters theses, journal articles, or conference proceedings will not be included in the study sample. All dissertations will be collected from the ProQuest Dissertations & Theses A&I database. The PQDT receives 97.2% of all dissertations and theses from research universities in the United States 87.2% of those from Canadian

research universities (Davies, 2010). The number of English language dissertations from other countries listed in the PQDT are quite small, 73 out 4027. This small number is considered to not be representative of distance education scholarship in these other countries and so only dissertations from United States and Canada will be included.

To determine if a dissertation is based in the field of distance education seven search terms will be used as indicators for database searchers: "distance education", "distance learning", "online learning", "electronic learning", "e-learning", "distributed learning", and "network learning". For inclusion in this study, one or a combination of these search indicators must be present in one of the following database fields: title, abstract, classification, keywords, or subject.

Research methods are often not made explicit in the abstract or keywords of a dissertation. Therefore an analysis of the methods used in the dissertations will not be made. Dissertation full texts are available for roughly 6% of the study sample of 3954 dissertations. Therefore, data gathered will be limited to bibliographic data listed for each dissertation database record and reference citations when available. Analysis will focus on bibliographic information found in PQDT dissertation listings as opposed to the full text.

<div align="center">Limitations of the Study</div>

The sample of distance education focused dissertations in North America may be incomplete, rather only those studies listed and retrieved by the ProQuest Dissertations & Theses A&I database will be included. Only 10% of the retrieved dissertation records include references cited in the study. Therefore, the citation analysis and co-citation analysis will only include 10% of the total sample. Any findings or trends illustrated

through these analyses will only be indicative of this sub-sample and cannot be generalized beyond.

Errors in the reference lists impact the reliability and accuracy of the data. Moed (2005) notes that misspellings, incorrect volume numbers, and different versions of an author's name affect the accuracy of citation data. The *n* size of reference citations for the dissertations in this study, in the tens of thousands, disables manual review and correction of this citation data. Yet some algorithmic identification of error is possible and if errors are detected in the citations they will be corrected.

Collected dissertation data will be transferred out of the ProQuest Dissertations & Theses A&I database and into a database created to house the data sample for this study. During this transfer and ingest, a number of algorithms will be utilized to section and chunk bibliometric data into categories (such as title, author, abstract, year, etc.) before the data is imported into the study database. Algorithms will be tested to ensure the accuracy of data processing before the complete study sample is ingested into the study database. This is discussed in chapter III.

Organization of the Study

This study is organized in a five-chapter format. Chapter I presents a topical background, purpose of the study, significance of the study, research questions, delimitations, and limitations. Chapter I also addresses the framework for analysis of topics and identifies the methods of the study. Chapter II discusses previous literature related to distance education research reviews, as well as a synthesis of relevant findings. This chapter also describes bibliometrics and network analysis and how these techniques have been utilized in previous similar studies. Chapter III addresses the use of these

methods in the current study. This chapter also details the design of the study and the

data sample. Chapter IV discusses the analysis of the collected data and presents the

results of the study. Chapter V presents a summary of the findings, interpretations,

implications, recommendations, and conclusions.

CHAPTER II

LITERATURE REVIEW

This chapter reviews literature related to analyzing trends in distance education research. It is organized into three main sections. The chapter begins with an overview and synthesis of previous studies that review distance education research. The section also addresses reviews of graduate student research on distance education as well as reviews of research employing social network analysis. The second section reviews bibliometric analysis as a research method. The final section addresses SNA as a research method. Finally a summary of relevant findings from the literature is presented.

Reviews of Distance Education Research

One method to gain access to trends in distance learning research is to look at previous studies that have reviewed the literature base. Numerous studies have reviewed distance education literature during the past twenty years (Koble and Bunker, 1997; Anglin, and Morrison, 2000; Berge and Mrozowski, 2001; Rourke and Szabo, 2002; Lee et al., 2004; Zawacki-Richter, Backer, and Vogt, 2009b; Ritzhaupt et al., 2010; Simonson et al., 2011; Bozkurt, et al., 2015). These studies analyzed attributes such as topics, methods, and bibliographic data of journal articles. This section inspects existing studies that review distance education literature and a synthesis of their findings follows.

Koble and Bunker (1997) assessed 129 articles in the *AJDE* between 1987-1995. The review investigated authors, audience, article topics and methodology. The authors utilized subject headings from the International Center for Distance Learning at the British Open University as a framework to classify article topics and evolved these to work as classification of article topics. Article abstracts were coded to determine topic

classification and methodology. Full text was analyzed for articles where abstracts were not adequate to determine topic or method.

The resulting categorized article topics included: Theory, Policy, and Development (25.65%), Media and Delivery Systems (20.9%), Institution, Staff and Management (15.5%), Student, Psychology, Motivation, and Characteristics (14.7%), Faculty Participation and Instructional Process (10.9%), Course Design and Curriculum Development (10.1%), and Student Administration and Support (2.3%). Main research contributors were from higher education institutions in the US (70%) and Canada (30%). Trends in topics areas showed that authors addressed effectiveness of distance education, the move from correspondence to telecommunications technology, and interactions in distance education.

Anglin and Morrison (2000) evaluated 383 studies, 222 in the *ADJE* and 161 in *DE*, between 1987–1999. For each article author, name, publication date, article type, topic, audience described, and type of data were collected. They reported that many studies were anecdotal and bound to a particular distance education context. Studies did not base their approach in theory (distance education, learning, or instructional). Also researchers did not distinguish between delivery technologies and instructional technologies. They call for a systems approach to investigate components including learners, content, organization context, process staff, and communications technology in order to go beyond developing *pockets* of knowledge.

Anglin and Morrison found seven study types and report the percentage of these found in the study sample: Primary Research (38.12%), Conceptual/Theory (30.03%), Review (3.66%), Evaluation (3.66%), Lessons Learned (13.58%), How To (2.35%), and

Other (8.62%). While they detail the types of articles published in the journals of interest, they fail to detail article topics or areas of focus.

Berge and Mrozowski (2001) reviewed 1,419 journal articles and distance education dissertations from 1990 – 1999. These were sampled from *AJDE* (142), *DE* (185), *JDE* (111), and *OL* (289). The authors included dissertations (692) and "acknowledge that acceptance criteria for journals versus dissertations may be quite different "(Berge and Mrozowski, 2001, p.8). The authors found that 62.7% of journal articles were based in research (descriptive, case study, correlational, experimental) and a high percentage, 93.4%, of dissertations were research oriented. Of these 74% of dissertations and 75% of journal articles used descriptive research.

Berge and Mrozowski (2001) used Sherry's (1996) topic categorization system as a framework for content themes found in the sample. The results were: Design issues (21.57%), Learner characteristics (16.96%), Strategies to increase interactivity and active learning (15.61%), Technology selection and adoption (10.89%), Policy and management issues (10.67%), Redefining roles of key participants (8.65%), Operational issues (7.64%), Learner support (5.50%), Equity and accessibility (1.57%), Cost/benefit trade-offs (0.89%). The authors found the majority of the studies focused on design issues, strategies to increase interactivity and active learning, as well as learner characteristics. Like Anglin and Morrison (2000), the authors found that research addressed individual courses rather than on full academic programs. They also note that studies addressed the impact of individual technologies as opposed to the interaction of multiple technologies.

Rourke and Szabo (2002) investigated the *JDE* from 1986 – 2001. While analysis of trends and methods were included in their research questions, one of their main

objectives was to do a comprehensive review of the journal's content and compare the information with the mission of *JDE*. All content (235 items) including book reviews, dialogue, editorials, letters to the editor, and research articles were included in the analysis.

The authors utilized an inductive process to extract topic categories and used these in conjunction with several distance education textbooks table of contents to create the typology topics. Results showed a nearly linear decrease across the ten categories Foundations (13.2%), Technology/Media (13.2%), Administration (12.3%), Instructional Design (10.6%), Learner Characteristics (9.8%), Editorial (8.1%), Instructional Perspectives (6%), Faculty Characteristics (4.7%), Evaluation (4.3%), and Student Support Services (1.3%).

Lee, Driscoll, and Nelson (2004) examined 383 articles from the *ADJE*, the *JDE*, *DE*, and *OL* between 1997–2002. Research topics, research methods, statistical methods, and citation indexes were gathered. A categorization system was developed based on Sherry (1995), Phipps and Merisotis (1999), and Khan (1997). The six categories and frequency rankings found were: Theory and research (31%), Design (27%), Development (9%), Management (11%), Evaluation (12%), Institution and operation (10%) .

A thematic analysis was also performed. Each article was assigned three keywords and frequency was used to show trends over the six-year period (see Table 1).

1997	1998	1999	2000	2001	2002
Interaction (12)	Program quality (8)	Learners' perception (5)	Collaboration (10)	Cross-cultural issues (11)	Problem-based learning (8)
Program Evaluation (8)	Learners' perception (7)	Video-conferencing (4)	Learners' perception (8)	Faculty support (5)	Interaction (7)
Collaboration (8)	Program's effectiveness (5)	Learners' participation (4)	Video-conferencing (5)	Video-conferencing	Learners' attitude (6)
Videoconferencing (7)	Faculty support (5)	Learners	Program's effectiveness	Collaboration (4)	Learners' perception
Learning Outcome (6)				Barrier to	
Online learning					

model (5)	Interaction (4)	attrition (4)	(4)	online learning (3)	(4)
Program quality (4)	Instructor's (leadership (4)	Collaboration (3)	Learners' achievement (4)	Assessment of outcomes (3)	Flexible learning (4)
Faculty support (4)	Theory development (4)	Program evaluation (3)	Need analysis (4)	Interaction (3)	Learners' satisfaction (4)
Learners' perception (3)	Institution issues (4)	Metacognition (3)	Program evaluation (4)	Learners' performance (3)	Tutor (4)
Learners satisfaction (3)	Learning strategy (3)	Learners' achievement (3)	Faculty support (3)	Program quality (3)	Program evaluation (3)
	Assessment of outcome (3)	Faculty support (3)	Self-directed learning (3)	Learners' persistence (3)	Faculty support (3)
		Tutor (3)	Cost issues (3)		Scaffolding (3)

Note: Frequency of the key words in parentheses

Table 1. Distance Education Research Specific Topics by Key Words (Lee et al., 2004, p. 232)

Unlike the previous reviews, Lee et al. analyzed the citations (8,409) found in the reference lists of the examined articles (383). In an effort to identify the landscape of primary authors and publications that had contributed to the distance education research the authors used frequency counts to analyze the most frequently cited authors (by journal), books titles, and book chapter/article titles. The authors noted that frequency bias in article authors' high ranking were based on studies being extensively cited in one or two journals as opposed to an author who had multiple studies cited in all four journals. Also, a limitation to the study was that secondary authors were not considered in the citation analysis.

Zawacki-Richter, Bäcker, and Vogt (2009b) reviewed 695 journal articles published between 2000–2008. The authors chose 2000 as the cut-off year noting that Berge and Mrozowski (2001) performed a large-scale review with articles published between 1990 – and 1999. Five journals were reviewed for the study: *OL, DE, AJDE, JDE, and IRRODL*. Excepting *IRRODL*, which was first published in 2000, these

journals have been used as data sources in previous reviews, including Berge and

Mrozowski's (2001) review of the previous decade.

In an earlier study (Zawacki-Richter, 2009a) performed an international Delphi

study to develop the *Classification of Research Areas in Distance Education (CRADE)*.

Based on analysis of 25 experts from 11 countries, three broad meta-levels of distance

education research were derived.

Macro Level: Distance education systems and theories

Meso Level: Management, organization and technology

Micro Level: Teaching and learning in distance education (Zawacki-Richter,

2009a, p. 22).

Rank	Research Area	Level	F	%	Cum %
1	Interaction and communication in learning communities	3	122	17.6	17.6
2	Instructional design	3	121	17.4	35.0
3	Learner characteristics	3	113	16.3	51.2
4	Distance teaching systems and institutions	1	62	8.9	60.1
5	Educational technology	2	48	6.9	67.1
6	Quality assurance	2	41	5.9	72.9
6	Professional development and faculty support	2	41	5.9	78.8
7	Access, equity, and ethics	1	31	4.5	83.3
8	Theories and models	1	24	3.5	86.8
9	Learner support services	2	23	3.3	90.1
10	Management and organization	2	18	2.6	92.7
11	Research methods in DE and knowledge transfer	1	13	1.9	94.5
11	Globalization of education and cross-cultural aspects	1	13	1.9	96.4
11	Innovation and change	2	13	1.9	98.3
12	Costs and benefits	2	12	1.7	100.0
	Total		695	100	
Level: 1=macro, 2=meso, 3=micro, F=frequency, Cum. %=cumulative %					

Table 2. Ranking of Research Areas by Number of Articles by Research Area (N = 695)
(Zawacki-Richter et al., 2009b, p. 26) * See Appendix 1.

Zawacki-Richter et al. (2009b) coded 695 studies to topic areas (see Table 2).

Coders also coded research methods classified as: quantitative, qualitative, triangulation,

or other. The authors found that the Micro Level (Teaching and learning in distance

education) accounted for over 50% of articles. These included the top three areas of

study: interaction and communication in learning communities (17.6%), instructional

design (17.4%), and learner characteristics (16.3%) (see Table 2). The only discernible

trend in research methods was a modest upward trend in qualitative methods. Frequency

counts for authorship patterns were also reported, including frequency rankings for

leading contributors, author gender, country-wise distribution, and number of references

per article by journal.

Synthesis of Distance Education Research Reviews

Table 3 lists each of the distance education reviews covered in this chapter.

Attributes of each study are presented for comparison. Discussion of these studies'

findings in regards to research topic trends follows.

Authors	Koble & Bunker	Anglin & Morrison	Berge & Mrozowski	Rourke & Szabo
Date of Review	1997	2000	2001	2002
Journals & Years Reviewed	*AJDE* 1987 - 1995	*AJDE* 1987 – 1999 *DE* 1991 – 1999	*AJDE* *DE* *JDE* *OL* Dissertations 1990 - 1999	*JDE* 1986 - 2000
Number of Articles Reviewed	129	383	890	235
Analyzed or Identified	Authors Audience Article topic Research	Authors Article topic Type of article Type of data	Authors Research methods Research	Authors Type of article Article topic

	methods		problem Article	
Review Method	Content Analysis	Content Analysis	Content Analysis	Content Analysis
Framework	International Center for Distance Learning at the British Open University	Developed by Author	Sherry (1996)	Developed by Author Textbook table of contents
Analysis of Citations	No	No	No	No
Analysis of Keywords	No	No	No	No

Authors	Lee, Driscoll, & Nelson	Zawacki-Richter, Backer, and Vogt	Ritzhaupt, Stewart, Smith, & Barron	Davies, Howell, Petrie
Date of Review	2004	2009	2010	2010
Journals & Years Reviewed	AJDE DE JDE Open Learning 1997 - 2002	AJDE DE JDE Open Learning IRRODL 2000 - 2008	AJDE JDE 1987 - 2005	ProQuest Dissertation Theses Database (PQDT) 1998, 2002, 2007s
Number of Articles Reviewed	383	695	517	308
Analyzed or Identified	Authors Research methods Article topic	Authors Research Area Research Method Article Topic	Abstracts	Authors Research methods Article topic
Review Method	Content Analysis	Content Analysis	Bibliometric Social Network Analysis	Content Analysis
Framework	Sherry (1996), Phipps & Merisotis (1999) Khan (1997)	Zawacki-Richter (2009a)	Emergent through Social Network Analysis	Developed by Author
Analysis of	Yes	No	Yes	No

Citations				
Analysis of Keywords	Yes	No	Yes	No

Authors	**Bozkurt, Akgun-Ozbek, Yilmazel, Erdogi, Ucar, Guler, Sezgin, Karadeniz, Sen-Ersoy, Goksel-Canbe, Dincer, Ari, Suleyman, & Aydin**
Date of Review	2015
Journals & Years Reviewed	AJDE DE EURODL JDE JOLT Open Learning IRRODL 2009-2013
Number of Articles Reviewed	861
Analyzed or Identified	Keywords Research Areas Conceptual Background Research Design Collection Instrument Focused Variable Targeted Population References Cited Authors
Review Method	Content Analysis Bibliometric Analysis Social Network Analysis
Framework	Zawacki-Richter (2009a)
Analysis of Citations	Yes
Analysis of Keywords	Yes

Table 3. Attributes of Reviews of Distance Education Research

Table 4 shows the topics and themes, as well as the frequency percentage of each of these topics, found to be present in the aforementioned Reviews of Distance Education Research.

Koble & Bunker (1997)		Berge & Mrozowski (2001)	
- Theory, Policy, and Development	25.60%	- Design issues	21.57%
		- Learner characteristics	16.96%
- Media and Delivery Systems	20.90%	- Strategies to increase interactivity and active learning	15.61%
- Institution, Staff and Management	15.50%		
- Student, Psychology, Motivation, and Characteristics	14.70%	-Technology selection and adoption	10.89%
		- Policy and management issues	10.67%
- Faculty Participation and Instructional Process	10.90%	- Redefining roles of key participants	8.65%
- Course Design and Curriculum Development	10.10%	- Operational issues	7.64%
		- Learner support	5.50%
- Student Administration and Support	2.30%	- Equity and accessibility	1.57%
		- Cost/benefit trade-offs	0.89%

Rourke and Szabo (2002)		Lee, Driscoll, & Nelson (2004)	
Foundations	13.20%	Theory and research	31%
Technology/Media	13.20%	Design	27%
Administration	12.30%	Development	9%
Instructional Design	10.60%	Management	11%
Learner Characteristics	9.80%	Evaluation	12%
Editorial	8.10%	Institution and operation	10%
Instructional Perspectives	6.00%		
Faculty Characteristics	4.70%		
Evaluation	4.30%		
Student Support Services	1.30%		

Zawacki-Richter, Backer, and Vogt (2009)		Bozkurt, Akgun-Ozbek, Yilmazel, Erdogi, Ucar, Guler, Sezgin, Karadeniz, Sen-Ersoy, Goksel-Canbe, Dincer, Ari, Suleyman, & Aydin (2015) *	
Interaction and communication in learning communities	17.6%	Interaction and communication in learning communities	13%
Instructional design	17.4%	Learner characteristics	12%

Learner characteristics	16.3%	Instructional design	11%
Distance teaching systems and institutions	8.9%	Educational technology	15%
Educational technology	6.9%	* Only partial percentages reported	
Quality assurance			
Professional development and faculty support	5.9%		
Access, equity, and ethics	5.9%		
Theories and models	4.5%		
Learner support services	3.5%		
Management and organization	3.3%		
Research methods in DE and knowledge transfer	2.6%		
	1.9%		
Globalization of education and cross-cultural aspects	1.9%		
	1.9%		
Innovation and change			
Costs and benefits	1.7%		

Table 4. Research Topics Found in Reviews of Distance Education Research

Koble and Bunker (1990) utilized subject headings from the international Center for Distance Learning at the British Open University as a framework to classify articles. Lee et al. (2004) developed a framework based on previous studies.   There are notable similarities in these frameworks even though they are separated by nearly a decade and a half (see Table 5).

Koble and Bunker (1990)	Lee et al. (2004)
Theory, Policy, and Development	Theory and research
Course Design and Curriculum Development	Development
Institution, Staff and Management	Design
Student Administration and Support	Management
Media and Delivery Systems (Effectiveness/Evaluation/Methods)	Institution and operation
	Evaluation
Student Psychology, Motivation, and Characteristics	
Faculty Participation and Instructional Process	

Table 5. Comparison of Koble and Bunker (1990) and Lee et al. (2004) Research Topic Frameworks

These two studies represent investigation of the literature during the Pre-Web period (Koble and Bunker, 1990) and the Emerging and Maturing Web Periods (Lee et al., 2004).   From the pre-web period, Koble and Bunker's (1990) top three reported topics were Theory, Policy, Development (25.6%), Media and delivery Systems (20.9%), and Institution, Staff, and Management (15.5%).

Lee et al. (2004) presented findings for each year between 1997 and 2002 (see Table 4). The authors reported, similar to Bunker and Koble (1990), that Theory and Research (31%) is the top topic category for each of these years. The second top topic category is Design (27%). These two categories accounted for more than 50% of topics Lee et al. addressed.

A comparison of the two studies shows that the Theory based category has remained the primary concern across the three web periods. In the Pre-Web period, Koble and Bunker (1990) reported Media and delivery systems (20.9%) and Institution, Staff, and Management (15.5%) were the next highest topics addressed in the literature. In contrast by 1997 and through 2002, Lee et al. (2004) reported that studies addressing Management (11%) issues were among the lowest represented category of study. Institution and Operations (10%) (i.e. technology delivery systems) were also low in representation. During the emerging and maturing web periods, Lee et al. (2004) report that distance learning Design (27%) was one of the top issues researched each year. This may indicate that during the early years of distance education work was done to lay the technical and business foundations for distance education programs. By 2004 and the

Lee et al. study, this work had been largely completed and concern had moved away from business and infrastructure and could center more on course design.

Koble and Bunker (1990) listed the research topics Learner Characteristics (14.7%) and Course Design (10.1%) as secondary to the previously mentioned business oriented categories. Lee et al. (2004) reported Design as the second highest topic of interest between 1997-2002, but did not differentiate a category addressing learner characteristics. Two other studies Berge and Mrozowski (2001) and Zawacki-Richter et al. (2009b) did report Learner Characteristics as one of the most researched topics of the decade. This further indicates the move away from business-oriented topics and towards research addressing learners and courses.

Zawacki-Richter (2009a), who looked at studies between 2000-2008, set out to continue from where Berge and Mrozowski's (2001) study left off, 1990-1999. Combined the two studies provide a nearly two-decade span of the topics addressed in the major journals that publish distance education related research. The most salient point in comparing these two studies is that the top three categories are the same for both decades. Berge and Mrozowski's (2001) top topic categories were Design Issues (21%), Learner Characteristics (17%), and Strategies to increase interactivity and active learning (16%) (see Figure 1). Zawacki-Richter et al. (2009b) reported similar findings nearly ten years later with the top categories being Interaction and communication in learning communities (17.6%), Instructional Design (17.4%), and Learner Characteristics (16.3%) (see Table 2). While order may slightly vary, these data strongly indicate that the focus of journal article research for the past two decades has trended towards studies on instructional design, online learners, and their interaction.

Technology related categories hold the next positions for most addressed topics in both studies. Zawacki-Richter et al. (2009b) differentiated between delivery-based technology Distance Teaching systems and institutions (8.9%) and Educational technology (6.9%) meant to aid in the learning process. Berge and Mrozowski's (2001) did not differentiate between delivery and educational technology by reporting the category Technology selection and adoption (11%). This confirms what Rourke and Szabo (2002) noted in their review. Earlier studies did not differentiate between educational and instructional technology as topics of research. This is not the difference between disciplines. Rather the distinction between technologies that are meant aid in the learning process (educational technology) and the technologies used to deliver distance education (instructional technology). Later studies made the distinction. Concentration on *selection and adoption* of technology in the earlier study may have been because institutions did not have technology solutions in place and were still in a technological evaluation phase. Later studies do not have this evaluation or adoption of technology focus, and therefore may indicate that institutions have systems in place and are past the initial fitting of distance systems.

There was a decrease in studies focusing on management-oriented categories in the later Zawacki-Richter et al. study. Berge and Mrozowski's (2001) reported Policy and Management Issues (11%), while Zawacki-Richter et al. (2009) reported Management and Organization (2.6%).

Learner support services are mentioned in both studies, Berge and Mrozowski's (2001) category was Learner support (6%) and Zawacki-Richter et al. (2009) was Learner support services (3.3%). The later study also included a category for Professional

Development and Faculty support (5.9%). This may indicate that as distance education

became more pervasive within institutions, there was more need to address training

instructors than in the earlier years while technology adoption was still the focus.

Cross-cultural issues are not mentioned in Berge and Mrozowski's (2001) earlier

study. Zawacki-Richter et al. (2009) include Globalization of education and cross-cultural

aspects (1.9%) as a category in their framework. While the ranking of this category was

low, it is of note to include it in this discussion. It may be addressed to a higher degree in

the doctoral dissertation literature during the past decade than in the research journal

context. It may also become a category of interest during this next subsequent decade of

research.

Graduate Distance Education Research and Social Network Analysis

As discussed above, there have been numerous studies reviewing research journal

literature in the field of distance education spanning the past twenty years and during all

three of the web related periods. Studies worked off of previous studies and confirmed

continuance of trends in the field. There are two distinct areas where little analysis has

been performed: first in the area of graduate research and second utilizing a non-content

analysis approach, specifically a SNA approach.

Graduate Distance Education Research

Abstracts to graduate dissertations and theses were included Berge and Mrozowski's

(2001) review of distance education research that covered Pre-Web through the Emerging

Web periods of 1990 – 1999. While the researchers reported the methods utilized in

dissertations, they limited their report of topics to journal articles. Therefore there is little

that can be shared regarding topics of dissertations in that study.

Davies et al. (2010) reviewed a sample of 100 dissertations and theses from three years, 1998, 2002, and 2007. The researchers analyzed the most commonly addressed topics and methods to determine changes over the investigated time period. The authors found a notable trend away from instructional media studies that compare distance education with traditional instructional practices. Also they reported a decrease in the number of studies that focused on technology issues, particularly about the quality of technology and the ability of distance educators to provide acceptable learning experiences. The authors also found that studies focusing on student issues and faculty issues increased over the three years studied.

Social Network Analysis

Both Lee et al. (2004) and Zawacki-Richter et al. (2009b) reported bibliometric frequency rankings for primary authors, book chapters, and journal articles. Lee et al. (2004) coded each article with three keywords and reported keyword frequency as indicating research topics. In 1997 the top keywords were Interaction, Program Evaluation, and Collaboration. In 2002 the top key words were Problem-Based Learning, Interaction, and Learners' Attitudes. Further discussion of these results will be made in chapter IV.

Ritzhaupt et al. (2010) utilized a co-word occurrence SNA method to investigate topic trends and themes found in the *AJDE* and the *JDE*, between 1987 and 2005. They found that network analysis, using a word co-occurrence network approach, enabled topics to objectively emerge from the data. The researchers reported that during the Pre-web era, 1987 – 1993, strong lexical relationships included *Distance-Education-Course* and *Need-Development-Quality* and indicated that the need to develop quality courses

and policies was a central trend. The 1994 – 1999 time period, the Emerging-Web, placed emphasis on the study of distance education, *Study-Distance-Learning*, and the development of theory, *Development-Theory-Information.* During the maturing web period between 2000 – 2005 the authors reported *Study-Distance-Student* as emphasizing the study of distance education and *Interaction-Communication-Tool* as emphasis on strategies for communication and interaction. These emergent lexical groupings indicated key areas of research during the specified periods and triangulated findings mentioned in many of the previous reviews of distance education research.

Zawacki-Richter, Anderson, & Tuncay (2010) utilized a bibliometric approach to demonstrate the growing impact of open access journals, such as IRRODL, on the distance learning literature base. Zawacki-Richter & Anderson (2011) investigated the journal network of 12 journals from the field of distance education. Bibliometric frequencies are reported placing AJDE, the Asian Journal of Distance Education, and Distance Education as the most frequent producers of research in the field between 2003-2008. Network analysis illustrated a directed that showed ties citing out of a journal and ties citing into a journal.

Bozkurt (2015) utilized a content analysis, bibliometric, and social network analysis approach to review distance education literature in seven journals from 2009-2013. A co-citation and keyword co-occurrence analysis was performed. The authors found a high similarity with the CRADE research area framework (Zawacki-Richter, 2009a). Particularly they found interaction and communication in learning communities, learner characteristics, instructional design, and educational technology to be the research areas most predominant in their network analysis.

Bibliometrics

Bibliometrics is a quantitative method that uses statistics to analyze bibliographic information found in written publications (Borgman & Furner, 2002; Moed, 2005, Diodato & Gellatly, 2013). "There does seem to be a clearly delineated body of research involving physical units of publications, bibliographic citations, and surrogates for them. The measurement of these items is called, logically, bibliometrics" (Broadus, 1987, p. 377).

Application/ Purpose of Analysis

Traditional bibliometric investigation is based on ranking elements by frequency, such as author characteristics, research methodology, and citation attributes (Diodato & Gellatly, 2013). Two common approaches are *citation analysis*, which works to demonstrate top contributors in a field, and *keyword analysis*, which attempts to show the dominant conceptual areas in a given research field (Diodato & Gellatly, 2013). Bibliometric techniques that use network analysis methods include word and citation co-occurrence (Moed, 2005). These are discussed further in chapter III.

Measurement

Borgman and Furner (2002) state that bibliometrics are concerned "with the measurement specifically of properties of documents" (p. 7). Measurements are frequency counts of document variables. Moed (2005) says, "Citation analysis comprises a variety of ways to analyze references cited in scholarly publication" (p. 20). Others have defined citation analysis as a method to rank citations according to the frequency they are cited in the reference and bibliography lists of publications (Waugh & Ruppel,

2004; Adams, 2014). In this study citation analysis and keyword analysis will refer to the frequency ranking method.

Citations and Metadata: Bibliographic Data Quality

Research methodology may be an important factor to be considered in citation analysis (Palmer, Sese, & Montano, 2005; Swyhart-Hobaugh, 2004). Palmer et al. (2005) excluded descriptive studies and limited analysis to quantitative studies and determined the frequencies by type of quantitative method. Swyhart-Hobaugh (2004) concluded in her study of sociology literature that quantitative publications primarily cite quantitative literature, while qualitative publications cite quantitative and qualitative publications. Citation context effects companion citations in a bibliometric data set.

Many studies lacked rigor or a theory-based perspective (Moore, 2003). Poor methodology and citing a large percentage of secondary sources could affect the validity and reliability of the research. The continued citing of poor quality research by researchers only perpetuates the problem. High citation counts should also not necessarily translate to quality research or researcher status (De Bellis, 2009). Determining the use of primary and secondary sources, author's source for citation, and size of citation are other citation characteristics that could be included in the analysis (Diodato & Gellatly, 2013).

Waugh and Ruppel (2004) conducted citation analysis of graduate student publications within their academic department. The purpose of their research was to provide information to assist their library's efforts to acquire and maintain journals within their academic discipline. A reference list of graduate student papers was compiled and a list of publications was ranked by frequency of citation. Waugh and Ruppel (2004)

concluded that graduate research citations are variable in quality. Beile, Boote, and Killingsworth (2004) in a study of education dissertations from three universities concluded there were differences in the currency, scholarliness, and appropriateness of the citations used by the doctoral students.

Age of the citation may also be a factor in bibliographic data quality. The citations are more valuable within the first few years of publication, thereafter moving towards obsolescence (De Bellis, 2009).

Social Network Analysis

Networks are "sets of socially-relevant nodes connected by one or more relations" (Marin & Wellman, 2011, p.12). Members of networks are called nodes or actors. Ties connect nodes and represent relationships (Borgatti, Everett,& Johnson, 2013). Patterns of relationship between groups of nodes are the focus of analysis. Traditionally social networks enable investigation of complex relationships between people or organizations (Hanneman & Riddle, 2011). Marin & Wellman (2011) note that in principle any unit that can be connected to other units can be studied as nodes, including data such as web pages (Catanese, De Meo, Ferrara, Fiumara, & Provetti, 2011, May), journal articles (Ding, 2011), journals (Zawacki-Richter, O., Anderson, T. (2011) and keywords (Ritzhaupt et al., 2009, Bozkurt, et al., 2015).

Relationships

Relationships in networks are measured between two nodes. The sum of the relationships of all pair nodes within a data sample makes the network. Understanding the broader patterns of ties within a network is needed to understand the effect and meaning of individual ties between two nodes (Barnes , 1972; Kadushin, 2011). The

context of a node pair's relation to other pairs within the network enables analysis of patterns. Assuming that each pair acts independently hides network processes that are created by larger patterns in the network (Reffay & Martínez-Monés, 2013).

Ties can signify four categories of relationships: similarities, social relations, interaction, and flows (Borgatti, Mehra, Brass, Labianca, 2009). Similarities occur when two nodes share attributes such as demographic characteristics, attitudes, locations, or group memberships. Marin & Wellman (2011) notes that these are the attributes frequently studied in variable-based statistical approaches. Social relations include kinship, role relations (e.g., friend, student), affective ties (e.g., feelings, liking, disliking), or cognitive awareness (e.g., knowing). Interactions are based on behaviors between two actors in a network, such as speaking with, helping, or inviting. Flows are relations based on exchanges or transfers between nodes. These may include relations in which resources, information, or influence flow through networks (Borgatti, Everett,& Johnson, 2013).

SNA studies have been used to map authorship, topic structures, and knowledge structures within a variety of fields including; Business (Backhaus, Lügger, & Koch, 2011), engineering (Ho, 2012), medicine (Pagel & Hudetz, 2011), and psychology (Burt, Kilduff, & Tasselli, 2013). Bozkurt et al. (2015) used co-citation analysis to investigate authorship and topic structures in distance learning literature. Ritzhaupt et al (2009) utilized co-occurrence social network analysis to investigate research topics in distance education abstracts. This study will follow Rizthaupt et al. (2009) and Bozkurt et al.(2015) and investigate the attributes of two types of networks, co-citation and keyword co-occurrence.

Focus of Analysis

A traditional statistical approach to investigating attributes (such as race, gender, education) treats causation as something that comes from within individuals. Common attributes act independently on individuals to produce similar outcomes (Marin & Wellman, 2011). Researchers sort individuals by common attributes and measure "which outcomes are disproportionately common to individuals with particular attributes" (Marin & Wellman, 2011, p. 14).

SNA positions causation as not from within individuals, rather in the social structure they belong to (Reffay & Martínez-Monés, 2013). "While people with similar attributes may behave similarly, explaining these similarities by pointing to common attributes misses the reality that individuals with common attributes often occupy similar positions in the social structure" (Marin & Wellman, 2011, p.14). People with similar attributes often have similar social network positions. Their similar outcomes are caused by the constraints and opportunities created by these similar network positions. Therefore by positioning the focus of study on the network one can see that people are not acting similarly just because they are similar, but also because of their location to one another within the larger social structure (Kadushin, 2011).

Whole Networks vs. Ego Networks

Whole networks provide the full view of relationships between all nodes within the data sample (Hanneman & Riddle, 2005). These networks begin from a list of included nodes and include data on the presence or absence of relations between every pair of nodes. Highly related groups of nodes, clusters or neighborhoods, can be investigated. But there is not an individual central node in a whole network.

On the other hand, egocentric networks focus on the network surrounding a central node, the ego. Nodes within this kind of network share relations(s) with the ego and between other nodes in the network (Hanneman & Riddle, 2011). Unlike whole network analysis, which commonly focuses on one or a small number of networks, ego network analysis typically samples large numbers of egos and their networks (Marin & Wellman, 2011). Commonly whole networks are analyzed and emergent central nodes are then investigated as ego nodes for subsequent egocentric analysis.

Adjacency Matrix

Network data uses matrixes, most commonly a square adjacency matrix with as many rows and columns as there are nodes in the data set (Hanneman & Riddle, 2005). Table 6 shows a simple binary matrix where a one in the cell indicates a tie and a zero indicates that there is no relationship between the nodes.

	A	B	C	D
A	---	1	0	0
B	1	---	1	0
C	1	1	---	1
D	0	0	1	---

Table 6. Four-by-Four Matrix (Hanneman & Riddle, 2005)

Nodes are usually not related to themselves, node A is usually not related to node A and so on. This is indicated in Table 6 by the *main diagonal* symbol ---. Matrixes can be progressively more complex indicating, for example, value data in the cell instead of binary ones and zeros. Also, direction of relationship can be represented. If node A is related to node B and node B is also related to node A (and the same is true for all nodes in the network) the matrix is symmetrical. If relationships are directional the matrix is

asymmetrical. In this instance the matrix rows represent the source of the directed tie and the columns hold the target (Hanneman & Riddle, 2005).

Adjacency matrixes for network analysis can be very large. They are typically generated using software to convert the data set into a matrix. UCINET can convert imported data into matrixes (Borgatti, Everett,& Johnson, 2013). SITKIS is specialized software that converts citation-based data into adjacency matrix files. Both will be utilized for this study.

Analyzing Network Data

Once adjacency matrix files are imported network position data such as node location, node pair position, and the position of the network as a whole can be calculated. Positions are determined from elements such as the number of ties a node has and the extent to which the node is a bridge between other nodes (Freeman, 1979; Kadushin, 2011). Node pairs, or dyads, are measured by strength or reciprocity of their tie. Their placements in relations to each other are dependent on factors such as similarity of the two nodes (homophily), their content, or the number of relation types shared (multiplexity) (Marin & Wellman, 2011).

Whole networks are analyzed by investigating areas of density, where large proportions of node pairs are related. Average path length necessary to connect pairs of nodes, the average tie strength, and the centralization of specific nodes in relations to other are of particular interest to begin to see patterns within a network (Freeman, 1979; Kadushin, 201). In addition whole networks can be divided into subgraphs. For example components are sets of nodes that are highly related to each other, but not related other sets of nodes within the network (Marin & Wellman, 2011). Analysis of one component

may reveal patterns not observed in other components of the same network. Network analysis measurement and analysis techniques for this study will look to understand the patterns present in citation and research topic oriented networks. These will be further discussed in chapter III.

Summary

This chapter discussed previous reviews of distance education research. There are numerous studies that have reviewed journal articles and reported a trend that instructional design, learner characteristics, and interaction have been highly addressed topics over the past two decades. But there are very few studies that review topics or trends within graduate student research on distance education. The field of bibliometrics was discussed. The traditional form of bibliometric analysis is frequency ranking of citation attributes. A number of issues regarding citation quality were also addressed. SNA in this study will focus on co-citation and co-occurrence of keywords. Through the calculation of pair relationships across a data set patterns of all the relationships within the network become available for a researcher to investigate. Position of data within the network enables investigation of the relationship structure of highly related concepts. This study will utilize bibliometric and SNA procedures to review graduate student research in the area of distance education.

CHAPTER III

METHODS

This chapter describes the methodological process used to investigate the research questions identified in chapter I. The chapter details sample collection, data preparation, and data analysis. The objective of this research is to explore the topics and trends present in doctoral research in the area of distance education from 2000 to 2014. The general research questions for the current study were:

Research Questions

General Research Questions

1. What research topics can be identified in doctoral dissertation research on distance education published in North America in English between 2000-2014?

2. How have the research topics changed over time, specifically from 2000-2004, 2005-2009, and 2010-2014?

3. How do the research topics compare to the research areas forwarded in Zawacki-Richter (2009a) CRADE framework?

Research Sub-Questions: Bibliometric and Citation Analysis

A. What departments did graduate researchers come from?

B. What universities did graduate researchers come from?

C. Which authors received the highest frequency of citation in dissertation reference sections?

D. What journal articles were cited with the highest frequency in dissertation reference sections?

E. What journal publications were cited with the highest frequency in dissertation reference sections?

F. What book titles were cited with the highest frequency in dissertation reference sections?

G. What are the relationships among dissertation reference citations using co-citation SNA?

Research Sub-Questions: Dissertation Research Topics

H. What dissertation database classifications received the highest frequency?

I. What dissertation keywords received the highest frequency?

J. What are the relationships among the topics identified in dissertation abstracts using co-occurrence SNA?

Research Design

While there have been numerous reviews of distance education focusing on research journals (Berge & Mrozowski, 2001; Zawacki-Richter et al., 2009; Ritzhaupt, et al. 2010; Zawacki-Richter & Anderson, 2011; Cho, Park, Jo, & Suh, 2013; Bozkurt et al., 2015) there have been very few studies addressing doctoral dissertations that research distance education (Davies et al., 2010). The design of this study was exploratory and looked to provide an initial view into this under investigated area. Using a methodology that is novel to the field furthers the exploratory design. Instead of categorizing dissertation research with an a priori framework through content analysis, as the bulk of previous reviews of research, this study used a bibliometric and network analysis approach that enabled keywords and topics to emerge directly from the data.

The unit of analysis for the current study was a single *dissertation publication listing* indexed in the ProQuest Dissertations & Theses A&I database. Full text was available for only 222 of the 3954 (6%) study sample. Therefore full text is not investigated. Rather this study focused on bibliometric information included in the dissertation database listing including attributes such a title, abstract, author, and keywords. The full list of data gathered from a single dissertation database listing is detailed in Table 7. Dissertation reference sections were a secondary data source collected when available from a dissertation database listing. Table 7 illustrates what data points, analysis method, and comparison points were used to address the research sub-questions.

Data Sources and Research Sub-Questions	Comparison Points	Analysis Methods

Dissertation Bibliographic Data
  A. Departments
  B. Universities
  C. Database Classifications
  D. Keywords
	A. By Year
	B. By Year
	C. By Year &
	    Dept. Type
	D. By Year &
	    Dept. Type
		Frequency Ranking

Dissertation Reference Citations
  E. Authors
  F. Journal Articles
  G. Journals
  H. Books
	E – H. By Year	Frequency Ranking

Dissertation Bibliographic Data I. Emergent Keywords	I. By Year & Dept Type	Network Analysis • Emergent Keyword Co-Occurrence
Dissertation Reference Citations J. Journal Articles & Book Chapters	J. By Year	Network Analysis • Co-Citation

Table 7. Data Points, Analysis Methods, and Comparison Points

Sample and Data Collection

Sampling Procedure: ProQuest Dissertations & Theses A&I Query

The sample of dissertation publication listings was gathered from the ProQuest Dissertations & Theses A&I (PQDT) database. PQDT (formally known as UMI) is a commercial database housing a searchable archive of published dissertations and theses (www.proquest.com). PQDT receives 97.2% of all dissertations and theses from research universities in the United States (276 of 284) and 87.2% (41 of 47) of those from Canadian research universities (Davies, 2010).

The sample included all doctoral dissertations returned from a PQDT advanced search using the following search constraints:

- Doctoral dissertations, exclude masters theses or conference proceedings

- Publication date between 2000 - 2014

- Written in English

- From North American institutions

- Search descriptors "distance education", "distance learning", "online learning", e-learning", electronic learning", "network learning", "distributed learning", or "web-based learning".

- Search for descriptors in fields "Anywhere except full text – All"

When the query is performed n=4765 dissertation records are returned, yet 811 were determined as irrelevant to the sample. Details for exclusion from the sample are discussed below.

Removal of Irrelevant Dissertation Records

The researcher manually reviewed each of the returned dissertation records to determine if inclusion in the study sample was appropriate. From the total returned

records 811 were identified as not pertaining to distance education as characterized in this study. The discrepancy stems largely from overlapping uses of key language in different discourses. For example both "online learning" and "network learning" are commonly used computer science terms utilized to discuss attributes of machine learning algorithms. Dissertation records not pertaining to the focus of this study were identified and removed from the study sample.

Study Data Sample Overview

The study data sample was n=3954**.** A breakdown of the sample by year is displayed in Figure 1. There were 182 dissertations in 2000, the first year of the sample. This number increased yearly and peaked in 2008 with 344 dissertations. The number then decreased until 2014 during which 247 dissertations were published in the PQDT.
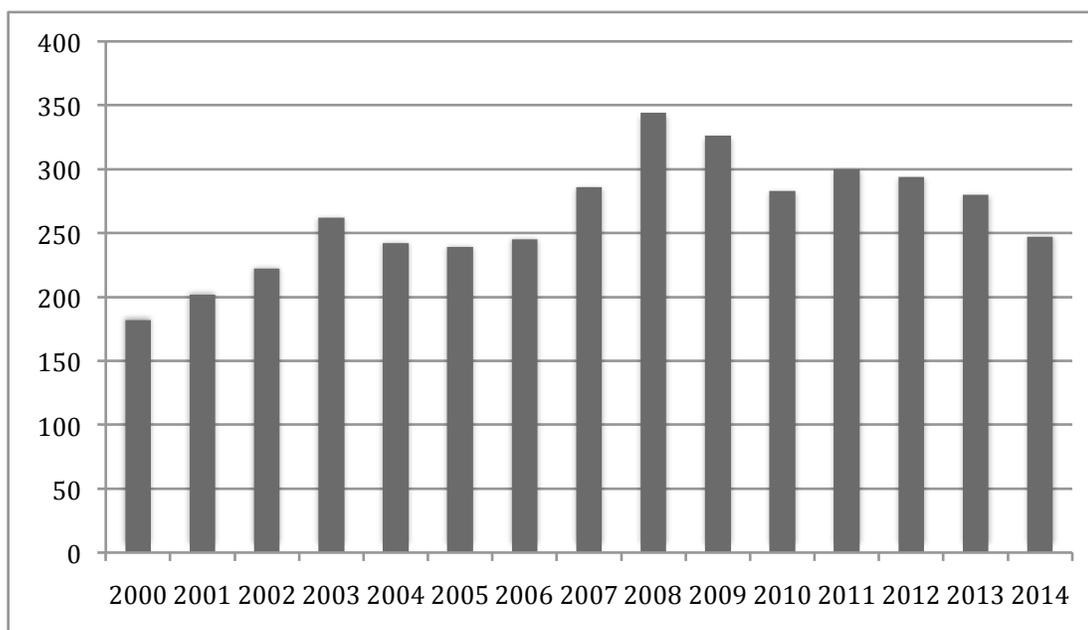
Figure 1. Number of Distance Education Dissertations by Year 2000 – 2014

Data Collection: Bibliographic Information Export

Each dissertation listing in the PQDT has 30 bibliographic fields associated with the dissertation. Table 8 displays all available fields and denotes fields used in the study. Most fields are requisite for inclusion in the PQDT. Fields that are not required for inclusion in the PQDT are Advisor, Committee member, and Department. Department was a data point utilized within the study. It is of note that it was not included as an indexing option in the PQDT until 2006.

Title *	Advisor
Author *	Committee member
Abstract *	University/Institution *
Publication Info	Department *
Subject *	University Location *
Classification *	Degree*
Identifier/ Keyword *	Source Type
Number of Pages	Language*
Publication Year	Document Type
Degree Date *	Dissertation/Thesis Number*
School Code	ProQuest Document ID*
Source	DOI
Place of Publication	Document URL
Country of Publication *	Copyright
ISBN	Database

Table 8. ProQuest Dissertation Database Fields. (* denotes fields used in this study)

Once a PQDT search was performed and the irrelevant records removed the sample was exported. The PQDT allows search records to be exported to both online and desktop-based citation management software formats including Endnote, Reference Manager, ProCite, or RefWorks. Returned search records may also be exported into RTF, TXT, PDF, RIS, and HTML file formats. The RIS file format is specifically standardized to deliver bibliographic information (Reference Manager, 2011). It was an ideal dissertation record export file format for this study because the tag format, similar to rss or xml, makes it easy to transcode data into other file formats or to ingest data into

a database. Unfortunately the RIS file format does not include all of the bibliographic

fields indexed by the PQDT. For example university department is not included in the

RIS export. Therefore RIS was not utilized as the export format for the study sample.

Instead HTML was used as the sample export format because it is natively

formatted for the web and can be transcoded, albeit with further processing steps

required, for ingestion into the study database. HTML exports of the PDQT study sample

search results were taken for each year of the study, as well as a full export that included

all years 2000-2014 in a single file.

Data Collection: Reference Citation Export

Inclusion of a dissertation's reference section is not a requirement for inclusion in

the PQDT and did not become available as an option until 2006. Therefore not all

dissertation records in the study sample included citation sections. Of this study's

n=3,954 sample, 338 (9%) of the dissertation listings included reference section citations.

Unfortunately the PQDT does not have a mechanism to bulk export citation

sections like it does for the dissertation listings. Instead each reference section was

manually saved into HTML and collected for ingest into the study database.

Data Transcoding and Database Ingest

A PHP-based HTML scraper was created to transcode the collected HTML files

into a format that could be ingested into the study database. The study database included

all fields listed in Table 11 and also included fields for UniversityType and

DepartmentType, fields unique to this study not the PQDT (see Appendix 2)**.** The HMTL

file exported from the PQDT was loaded into the scraper algorithm. The scraper then

worked on a series of automated loops by first identifying a single dissertation listing

within the HTML code. Then all of the attributes for that record (title, author, keywords,

abstract, etc.) were extracted into individual variables and imported into the study

database and marked as associated with the dissertation listing. Once data was imported

into the study database it could be queried and exported into formats required by the

bibliometric and SNA tools.

The data scraping and database import process was automated. Before the full

dataset was processed and imported into the database, the accuracy of the HTML scraper

was tested. A test sample of 25 dissertation records and 10 citation reference sections

were scraped and imported into the database. Both processing and import tests of the

sample dissertation records and citation records were 100% accurate. From these tests it

was determined the scraper was valid, reliable, and adequate for ingesting the data

records into the study database without error.

A similar process was undertaken for the secondary data sample, the dissertation

citations. A similar set of HTML scrapers was designed to isolate citation fields,

including reference type, author(s), publication, volume, and publication house. These

citation variables were imported into the study citation database. Full field details are

listed in Appendix 3**.**

<div align="center">Data Analysis</div>

Bibliometric Questions

Borgman and Furner (2002) state bibliometrics is concerned with the

measurement of document properties. To address the study's bibliometric sub-questions

A- F, comparison of frequency counts of document variables was the primary method of

analysis.

To generate these counts data was called from the study database using a series of SQL queries developed to calculate frequencies of interest for each of the study questions. For example, research sub-question B was "What universities did graduate researchers come from?" To answer this question the following PHP executes a SQL query:

```
function universityCounts(){

    $query = "SELECT university, COUNT(*) AS uniNum FROM studyDatabase
            GROUP BY university";

        $result = mysql_query($query);

        while ($row = mysql_fetch_array($result))

        {echo "row[university] =  row[uniNum]";}

    }
```

The query defined as *$query* says select the university and then count all of the dissertations records from that university. The results for each university are then stored in an array. The *while* statement starts a loop and for each loop through the array the *row[university]* variable returns the name of the university and the *row[uniNum]* variable returns the associated number of dissertations. In this way a list of universities and associated dissertation frequencies were populated. This example query does not limit the call by year, so the result will list all universities and the number of dissertations from each year spanning 2000-2014.

Results for each sub-question were broken up for analysis by comparison points, either by year, department type, or both, see Table 7. So to further investigate results for sub-question B "What universities did graduate researchers come from?" nineteen

variations to the above PHP functions and SQL query were needed. Fifteen of these

variations limit the returned data to a single year, limited to 2003 for example. Three of

the variations limited the returned data to the five-year spans 2000-2004, 2005-2009, and

2010-214.   Finally the last variation has no year limitation, as seen in the function above,

and returned data from the full span of years. To navigate through these variations a web

interface was developed so that the researcher could move between the different data

results.

Figure 2. Web Interface to Investigate Study Data

Figure 2 is a screenshot of the web interface that details the results of sub-question B from 2010-2014. The results show that 1404 dissertations were published during the time span. Of these 734 were from public institutions. Then individual institutions are listed with the number of dissertations shown by department type. The top navigation displays a link to each of the nineteen variations for the sub-question. The Home button returns the user to a menu where the data for each of the sub-questions may be accessed. The interface is available at http://jasonskinner.net/diss2. This interface also allows the results to be exported as excel spreadsheets.

Keyword Analysis: Keywords and Classifications

The research sub-question questions H and I addressed dissertation topics and were investigated using dissertation keywords and abstracts. Abstracts and keywords attached to a study are of particular value because they are the descriptors the researcher self-chooses to identify the topics addressed (Lee & Su, 2010). Berge & Mrozowski (2001), Lee et al. (2004), and Bozkurt et al., (2015) included abstract or keywords as topic indicators in their reviews of distance education research. Ritzhaupt et al. (2010) used a network analysis approach to extract emergent keywords from abstracts. These emergent keywords were utilized as indicators of research topics. This study utilized both keywords and emergent keywords, derived from dissertation abstracts, as indicators of research topics.

Keyword and citation data both required further preparation before frequency analysis was made. Each dissertation record in the PQDT had a comma-delimited list of keywords associated with the study. In preparation for keyword analysis this keyword list was broken into an array of keywords and keyword phrases for each dissertation

record. These keywords and keyword phrases were imported into a study keyword database so that they could be accessed for further standardization.

These collections of keywords were standardized through a process called stemming. Stemming is used to "remove morphological affixes from words, leaving only the word stem" (Natural Language Tool Kit, 2012a). Affixes may include separable prefixes and suffixes (pre-, -ness), tense (-ed), or plurals (-s). Stemming is used to standardize words into root stems so that counting variations of the same word separately does not skew frequency count. The Natural Language Tool Kit (NLTK) is an open source Python library for Natural Language Processing (Bird, Loper, & Klein, 2009). The NLTK was used to stem the keyword records for the study. Once standardized, frequencies of keywords were calculated. Reported frequencies for keywords were compared both over time and by department type.

Keyword Analysis: Emergent Keywords

Words and phrases found in the abstracts of dissertations may also provide valuable information regarding the topic of a research study (Ritzhaupt et al., 2010). This study investigated abstracts of dissertations for *emergent keywords*.

For each dissertation record the following procedure was performed on the abstract text to create a list of emergent keywords. First the text was processed by the NLTK using a process called chunking. Chunking or chunk parsing is used to identify non-overlapping linguistic groups (such as noun phrases) in text (Natural Language Tool Kit, 2012b). For example, the chunk structure for noun phrase chunks in the sentence "I saw the big dog on the hill" is:

(SENTENCE:

(NP: <I>)

<saw>

(NP: <the> <big> <dog>)

<on>

(NP: <the> <hill>)) (Natural Language Tool Kit, 2012b).

By chunking, elements such as noun phrases are left intact, as opposed to considering each word in the sentence as a separate keyword. This is important for maintaining an author's initial intent. The term *instructional design* carries a specific intent that neither *instructional* or *design* indicate as precisely when separated.

Chunk groups were further processed through the removal of stop words. Stop words are frequently occurring, insignificant words that appear in a database record, article, or web page (MIT Libraries, 2012). Common stop words include: a, an, the, in, of, on, are, be, if, into, and which. These words do not contribute to our understanding of dissertation topic and therefore were removed as emergent keywords. Further, stop words were removed to standardize keyword chunks. In this way "*the instructional design*" and "*instructional design*" are not counted as different. The frequently used SMART system (Salton, 1971) stop word list was utilized and can be accessed at http://jmlr.csail.mit.edu/papers/volume5/lewis04a/lewis04a.pdf (Lewis, Yang, Rose, and Li, 2004). SQL queries were run on the keywords and keyword chunk groups. If stop words were identified they were removed from the study keyword list.

Next stemming using the NLTK standardized the keywords and keyword phrases. The results were a group of emergent keywords and phases for each dissertation abstract. These emergent keywords were imported back into the study keyword database for

subsequent analysis. Emergent keyword frequencies were compared both over time and by department type.

Citation Analysis

Unlike keyword data, reference citations do not need grammar, tense, or plurality standardization. However citations with multiple authors need to be considered. Many researchers of citations decide to drop all but the first author. This is not necessary. In this study the author field for citations was algorithmically analyzed and individual authors broken out and listed individually in a separate database related back to a particular citation. In this way each author could be analyzed and considered separately yet still connected to the citation. Further standardization for authors' names included removing all but last name and first initial.

Citations entered into the PQDT needed to be checked to counter for error like duplicate counting. A SQL query created to loop through the citation records searched the citation dataset. If more than one record was found to be highly similar, but not exactly the same, the record was added to a log. The researcher then investigated the logged citations to see if they were the same record but with minor differences (such as a misspelling or miss-cited issue/volume number). If it was determined that the records were the same, errors were fixed in the database. One exception was book edition information. If multiple similar records were returned based on different editions of the same book, the researcher ensured the correct edition was present in the database. Citation frequencies, broken out in bibliometric research sub-questions C-F were compared and reported both over time.

Social Network Analysis

The first network investigated utilized a co-occurrence network analysis method. Co-occurrence measures the strengths of related terms in a sample from which a set of graphs that illustrate the strongest associations between various terms can be created (Coulter, Monarch, & Konda, 1998; Ritzhaupt et al., 2010). This network was created using co-occurring emergent keywords from dissertation abstracts. Keyword co-occurrence network analysis moves from the premise that if a pair of keywords tend to occur together frequently within the text, they are more likely to be related concepts than pairs or words that do not commonly occur together (Ritzhaupt et al., 2010).

The second social network investigated in this study, a co-citation network, was calculated using the reference sections of dissertations. Figure 3 illustrates that co-citation indicates a relationship between two citations that are cited in the same citing document (Garfield, 1988, p61). Studies C, D, E, and, F all cite both article A and article B. The strength of the relationship between A and B is based on the number of citing documents that cite both A and B in their reference sections (Waugh & Ruppel, 2004). Cited documents (A and B) are related because they are cited by the same citing document (C, D, E, and F) even if A and B don't cite each other.

Co-citation

Item A (cited)   Item B (cited)

Papers A and B are associated
because they are both cited by
papers C, D, E, and F.
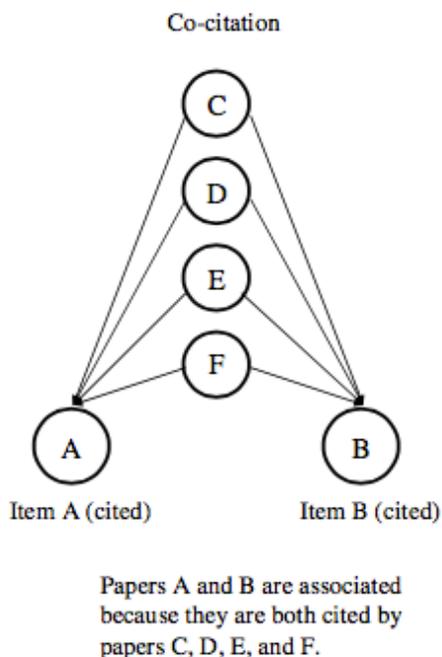
Figure 3. Co-citation Relationships

In co-citation analysis, there is a premise that if a pair of citations tend to be co-cited together frequently, that pair of publications are more likely to share related themes or concepts. In this sense, if collections of documents are grouped by co-citation counts, it enables scholars to find patterns of conceptual relationships among them (Benckendorff, 2009).

Social Network Analysis Data Preparation

The first step to prepare keyword and citation samples for network analysis was to export the data from the study database and transcode the returned data into the .ISI file format. The ISI file format is a proprietary output format for the Thompson Scientific publication databases, most notably the Web of Science (Alencar, 2012). ISI files contain bibliographic information including title, abstract, and other citation information. A PHP algorithm was created for this study to query the database and export keyword,

emergent keyword and co-citation data into ISI files. Specific fields exported for each of these ISI files are detailed in appendix 4.

SITKIS (http://users.tkk.fi/~hschildt/sitkis/index.html) is a Java based Windows application that imports ISI files into a Microsoft Access database. "The purpose of the program is to enable researchers to easily and quickly download and analyze bibliometric records," (Schildt, 2004). SITKIS includes a number of bibliometric analysis tools. But for this study it was used primarily as an intermediary program. First the ISI file data was imported into the SITKIS platform. Next network adjacency matrixes were calculated (see chapter II).

SITKIS has the functionality to export network adjacency matrix data into the .DL file format, the native file format for the UCINET software. Both emergent keyword and co-citation networks were imported into SITKIS. The adjacency matrixes were calculated and then exported into the .DL file format.

UCINET (https://sites.google.com/site/ucinetsoftware/home) is a social network analysis software suite. UCINET is one of the most utilized tool kits for social network measurement and visualization (Marin, 2011). This study utilized UCINET for network calculation as well as the program's NetDraw integration to create network graph visualization.

Co-Occurrence Frequency

Once the .DL files were imported into UCINET co-occurrence data for the networks were calculated. Frequency of co-occurring pairs are reported for each network. Just as frequency of individual key words or citations indicates that the term is an important concept or that the citation is an important source, frequently co-occurring keywords and

citation pairs give an initial indication of important couplings of concepts and publications within a network. Pairs that occur more frequently indicate that the relationships between the concepts or citations are more influential in the field, than the relationship between less frequently occurring pairs of concepts or citations (Waugh & Ruppel, 2004). Frequencies for co-occurring keywords and citation were reported by year.

Network Characteristics

Networks have a number of characteristics that determine how the relationships between data are measured. This study used three network characteristics; mode, directions, and value. *Mode* refers to how many types of data are connected in a network. A one-mode network looks at a network with only one type of data and conceivably every node could be connected to any other node in the network. Two-mode networks require the analysis of two kinds of nodes, for example organizations and organization members. In this example a node may belong to the other node type, members may belong to organizations, but not to the same node type, members cannot belong to other members (Marin, 2010).   In this study there was only one kind of data in each of the networks, keyword or citation. Since each could be related, this study used one-mode networks.

The next network characteristic is *Direction*. Directed ties are relationships where order of connection is specified and represented by an arrow pointing in the direction of the tie. Node A is connected to node B (arrow points towards B), but node B is not necessarily connected to node A. Reciprocity is possible between directed ties but not necessary. Undirected ties exist between two nodes in no particular direction. If a tie is

present between node A and Node B, then A is connected to B and B is connected to A (Marin, 2010). In this study order of connection was not specified. Keywords co-occurred in abstracts and did not have an ordered relationship where one pointed toward the other. Similarly citations co-occurred in reference sections. They did not have an ordered relationship. Therefore the study utilized undirected networks.

Value is the third network characteristic. Valued networks have ties with a set of variables, such that each carries some value. For instance, between nodes A and B the tie may have a value of 3. While between nodes A and C tie value is 1. The first pair is weighted heavier because of the higher value. On the contrary, unvalued network ties do not hold values; rather just indicate a base relationship between two nodes (Butts, 2008). As with direction, both keywords and citations had base co-occurrence relationships. They did not have value, rather simply co-occurred in abstracts or reference sections. Therefore the networks in this study were unvalued. This also meant that the adjacency matrixes were simple symmetrical binary matrixes that used ones to distinguish a relationship between nodes and a zero to signify no relationship (see Table 6).

Network Analysis Measurements

Centrality measures the influence of nodes in a network (Everett & Borgatti, 2005). A node with a higher centrality in a network tends to have a higher impact (Brass and Burkhardt, 1992). There are three types of centrality: degree centrality, betweenress centrality, and closeness centrality (Freeman, 1979).

Degree centrality is a measure between one node and the other nodes in a network (Hanneman & Riddle, 2005). In a social network nodes with more ties may be more advantaged because they may have many alternative ways to satisfy needs and hence are

less dependent on other nodes. These nodes may have access to, and be able to call on more of the resources of the network as a whole. Thus they have a higher degree of centrality and potential power within the network (Hanneman and Riddle, 2005). For this study, in the context of keywords for example, keywords with the highest score of degree centrality had the most number of direct connections to other keywords in the network and as such served as a hub or central concept within the network. Table 9 shows the formula for degree centrality. All centrality calculations were made using UCINET and the relationships with the highest degree centrality scores were reported.

Degree Centrality	**Betweeness Centrality**	**Closeness Centrality**
$d(i) = \sum_j m_{ji}$	$b(i) = \sum_{j,k=1} \dfrac{g_{jik}}{g_{jk}}$	$c(i) = \sum_{j=1}^{N} \dfrac{1}{d_{ji}}$
m = 1 if keyword i and keyword j are linked	gjk: the shortest path between keyword j and keyword k g: the shortest path between keyword j and keyword k that contains keyword I	dji = the shortest path between keyword j and keyword i

Table 9. Centrality Measurement Formulas (Hanneman & Riddle, 2005)

Betweeness centrality is a measure of how often a node is placed on the shortest path between any *other* two nodes in a network. A node with a higher betweeness centrality plays a critical role in the flow of information and resources within the network, and as such it is more likely to be placed at the center of the network. In the present context, it means that a keyword with a high betweeness centrality score linked important sets of keywords. Table 9 illustrates the formula UCINET uses to

calculate betweeness centrality. Relationships with the highest betweeness centrality scores were reported for each network.

Closeness centrality refers to the inverse of the average length of the shortest paths to/from all the other actors in the network (Lee & Su, 2010). A node with a higher closeness centrality can obtain information more easily. In the present context, a keyword with a high closeness centrality score (i.e., having shortest paths to other keywords) was closely linked to the other keywords in the network. Table 9 shows the formula for closeness centrality. Relationships with the highest closeness centrality values were reported.

Centrality measurements for degree, betweeness, and closeness were calculated and reported for the networks. These measurements in combination with the graph visualization of co-occurring pairs enable analysis and investigation into which keyword concepts and citation relationships were the strongest. Measurements for each of the networks were reported for the time periods 2000-2004, 2005-2009, and 2010-2014.

Network Visualization

One of the unique advantages of network analysis is the ability to graphically convey relationships of the elements being investigated (Freeman, 2011; Hanneman, & Riddle, 2011). Co-occurrence data can be visualized using UCINET's NetDraw functionality. There are three graph types available through UCINET: random, circle, and multidimensional scaling layout (MSL).   Random graphing places nodes randomly in the network. Circle graphing emphasizes nodes that are highly connected within the network and de-emphasizes nodes that are not highly connected. MSL graphing places nodes on a two dimensional non-metric scale according to the similarity of node's tie profiles

(Borgatti, Everett, Johnson, 2013). This enables nodes that are similarly connected to be placed in a similar region on the graph, thus enabling clusters or neighborhoods of relation to be visualized. Since visualizing similarity of nodes and their relationships is one objective of network analysis graphing for this study the networks were visualized with MSL graphs. Hanneman and Riddle (2011) note the benefit of network analysis is in drawing graphs that create clusters so that researchers can identify differences in the patterns of ties within and between groups.

Inclusion of all nodes and ties in a network graph can create an overwhelming network. Often a large portion of keyword or citation pairs occurred only once in the data. Researchers often set a co-occurrence minimum on pair inclusion. "If the co-occurrence minimum is set too high, few links may be formed; if it is set too low, an excessive number of links may be formed" (Coulter et al., 1998, p. 1211). Ritzhaupt et al. (2010) set the keyword co-occurrence minimum at 5 for inclusion in the network. Co-occurrence minimum is largely dependant on n size and frequency of pairs. It is common that a researcher visualizes the full network and then slowly increases co-occurrence minimum until the more import cluster structure of a network begins to emerge (Ritzhaupt et al. 2010). This approach is utilized for the present study. The co-occurrence minimum for emergent keyword pairs was set for 250. The co-occurrence minimum for the co-citation network was set to 3.

Comparative Analysis: Zawacki-Richter's Framework

Frequency and co-occurrence analysis of the classification, keyword, and emergent keyword data samples revealed topics of focus in distance education dissertations. Ritzhaupt et al. (2010) noted that results utilizing this form of analysis

objectively emerge from the data, unlike content analysis-based studies that use a priori frameworks. A first round of analysis and synthesis of the emergent data unconstrained by the boundaries of a framework were reported.

The results of this were then compared with Zawacki-Richter's (2009a) Classification of Research Areas in Distance Education (CRADE) framework research areas to see if the topics that emerged in this study fell within the 15 sub-category domains (see chapter II) identified by Zawacki-Richter. Also the trends and topics that are identified over time in this study are compared to Zawacki-Richter et al.'s (2009b) findings. In Zawacki-Richter et al.'s (2009b) study the researchers utilized Zawacki-Richter's (2009a) CRADE framework to investigate topics found in research journals between 2000-2008. While the sample populations differ, doctoral dissertations vs. journal publications, it is of interest to compare the similarities and dissimilarities of topics addressed. Comparisons were reported and discussed in the results section.

<div align="center">Validity and Reliability</div>

Citations are not objective measures of the information that flow from previously published literature. Rather it must be acknowledged that citations include differing contexts such as showing respect for pioneers, criticizing or correcting related work, identifying original sources for concepts or ideas, or following disciplinary trends for citing (Smith, 1981).

According to Garfield (1977), in citation analysis studies, inferences are made focusing on the first author rather than all authors. Thus researchers may miss significant contributions and collaborations of other authors. To address this issue, some researchers use the publication as the unit of analysis as opposed to the author (Pilkington &

Meredith, 2009). This study utilizes the publication as unit of analysis for co-citation analysis and thus the node variable for network analysis.

The source of keyword data is also a concern. Lee and Su (2010) note that author keywords may represent the concepts or paradigms of publications more clearly then the results of text mining algorithms. To account for this, the study keeps separate author generated keywords and emergent keywords mined through natural language processing techniques. Lee and Su (2010) also point out that results of text mining are not easily reproduced by other software due to the use of different coefficients or parameters in equations and algorithms. Effort was made to explain each of the steps and associated software so that this study's results can be re-examined and checked for repeatability.

Errors in citations cause issues of sample validity (Lee and Su, 2010). Citations were investigated algorithmically for errors. Citations that were found to be highly similar yet with small differences (such as spelling, volume, issue) were logged. Logged records were manually investigated and records corrected by the researcher.

Summary

This study explored the topics and trends in doctoral dissertation research in the area of distance education between the years 2000 - 2014 from North American institutions published in English. A data sample of graduate dissertation bibliographic records n=3954, was selected for analysis. This data sample was processed and imported into a database created specifically for the study. Bibliometric research sub-questions were addressed using frequency analysis of data from the study database. A citation analysis approach addressed frequency of publications and authors. A keyword and emergent keyword analysis investigated frequency of keywords assigned by authors as

well as emergent keywords that were algorithmically established from dissertation abstracts. Social network analysis was performed for two different networks: an emergent keyword co-occurrence network and a co-citation network. Network visualization analysis and centrality measurements were reported and discussed. All data were analyzed to address the general study questions regarding topics and trends in doctoral distance education research and their change over time. A comparison of these findings to the findings of Zawacki-Richter et al.'s (2009b) and Zawacki-Richter & Anderson (2014) was also discussed. Results, findings, and discussion are presented in Chapter IV.

CHAPTER IV

RESULTS AND DISCUSSION

Purpose of the Study

The primary purpose of this study was to investigate the research topics and themes found in doctoral dissertations addressing distance education from North American institutions published in English. A secondary purpose of this study was to assess how these trends changed over time: 2000-2004, 2005-2009, and 2010-2014. These two general purposes were addressed through two modes of investigation, bibliometrics and SNA. Bibliometric analysis enabled a picture of the doctoral researcher sample to emerge. Frequency ranking detailed the universities and departments graduate researchers came from, as well the most commonly investigated topics. SNA enabled investigation of the relationships between citations and research topics that emerged from dissertation abstracts. Results are presented in this chapter.

The main research questions were what research topics can be identified in doctoral dissertation research on distance education and how have these changed over time. Before these main questions were addressed the results from the research sub-question were reported. First results to sub-questions A-B, which give a picture of the graduate researchers, were reported. Secondly sub-questions C-G, which report on dissertation reference sections, are reported. Finally sub-questions H-J covering dissertation topics are reported. Using the results from these sub-questions the main research questions are discussed.

Research Sub-Questions: Bibliometrics

The bibliometric sub-questions gave a picture of the doctoral student population that wrote the distance education dissertations. Secondly the bibliometric sub-questions enabled a view into the research foundations found in dissertation reference.

"What departments do graduate researchers come from?" Sub-Question A.

Beginning in 2006 the PQDT database added "department" as an optional field graduate researchers could enter when filing their dissertation for inclusion in the PQDT. Before 2006 university was listed for each dissertation but not department. University department names differ greatly from institution to institution. For example there were 38 variations for education-based departments, including Department of Education, Curriculum & Instruction, and School of Educational Studies. This diversity of department names required standardizing of departments for analysis. A taxonomy of 17 department types were found. For a department type to be added to the taxonomy at least 5 dissertations from that kind of department needed to be found in the study sample. Departments that occurred only once or twice in the sample and did not belong to one of the established department types were added to the Miscellaneous department type. Departments included in this category ranged from Kinesiology to Aeronautics, to Geography. Not many inferences could be made about the Miscellaneous department type beyond that there are a diverse number of fields where one or two dissertations on distance education have been written over the past 15 years**.** Appendix 5 details specific departments assigned to each of the 17 department types.

The researcher manually sorted all dissertation listings that included department and assigned each a department type. Table 10 details the 17 department types and how many dissertations came from each type by year. The "none" column represents dissertations that did not have a department included in their PQDT listing. The majority of dissertations included in this study, 2579 or 65%, did not have a department associated with the dissertation listing. The remaining 45% or 1375 dissertation listings do have departments listed. Dissertations categorized as None were included for comparison purposes. Since 2011 there was a decline in dissertation listings with no department included. That year there were more dissertations from the Education department type listed than None. Each subsequent year the "None" listings decreased. If this trend continues, similar future studies will have a more complete department type sample.

Table 10 illustrates that for the first 6 years of the study, 2000 – 2005, graduate researcher department information was not available since it was not included in the PQDT.

	None	Education	Ed Lead	Ed Tech	Misc	Ed Psych	Business	Psych	Comp Sci	Nursing	High Ed	Org	Comm	Lang	Science	Spec Ed	Ag	Music
2000	182																	
2001	202																	
2002	222																	
2003	262																	
2004	242																	
2005	239																	
2006	242	1	1						1									
2007	213	48	8	5	1	1	5	2	1			2						
2008	212	81	11	11	5	11	4	2	4		2						1	
2009	195	68	20	16		13	6	1	3	1	1					2		
2010	125	83	25	17	6	6	9	1	2	1	2			5			1	
2011	106	89	33	22	6	9	9	4	3	3	3	2	3	1	4		1	
2012	65	100	37	26	12	11	5	11	4	4	2	4	2	3	2	3	3	2
2013	46	101	42	19	16	5	5	6	5	9	7	5	4	2	1	2	1	4
2014	26	108	33	11	16	4	9	9	7	7	5	4	3	1	2	1		1
Total	2579	679	210	127	62	60	52	36	30	25	22	17	12	12	9	8	7	7

Table 10. Dissertation Frequency by Department Type

Education was the highest represented department type across all three years, 679 or 49% of dissertations that reported department. Second and third respectively were Educational Leadership, 210 or 15%, and Educational Technology, 127 or 9% of dissertations that reported department. Miscellaneous was fourth, 63 or 5%, yet as noted previously should be considered as a diverse category that holds dissertations not represented by one of the other department types. Educational Psychology had the fifth highest frequency with 60 or 4% of dissertations that reported department.

Four of the top five department types were education oriented (excluding miscellaneous) and account for 1078 or 78% of dissertations that reported department. These results indicate that graduate researchers interested in distance learning predominately came out of education-based departments and therefore their research largely comes from the education research discourse.

The top non-educational department types were Business (52 or 4%), Psychology (36 or 3%), Computer Science (30 or 2%), and Nursing (25 or 1%). All had more than 25 dissertations across the years. Of these all but nursing had dissertations regularly listed each year after 2007. Yet even with this group of department types' 10% contribution, the sample remains largely represented by educational based researchers and the findings should be approached considering this.

"What universities did graduate researchers come from?" Sub-Question B.

The University field was mandatory for inclusion in the PQDT. Therefore, unlike department each dissertation in this study's sample included a university affiliation. Universities were segmented into three types: Public, Private, and For Profit.

	Public # of Diss	Public # of Univ.	Private # of Diss	Private # of Univ.	Profit # of Diss	Profit # of Univ.
2000-2004	841	154	220	72	49	6
2005-2009	929	179	268	86	243	6
2010-2014	734	172	330	101	338	8
All Years	2504	227	488	155	630	10

Table 11. Number of Dissertations and Number of Universities by Type and Time

Table 11 illustrates the number of dissertations published by time and university type. This table also shows how many universities were counted during each of the time periods. Across each time period more dissertations were published from public universities, with a total of 2504 or 63% of dissertations in the study sample, see Figure 4. The public dissertations came from 227 or 58% of the 392 total universities represented in the study sample, see Figure 5. The data showed that both number of public dissertations published and number of public universities represented increased from 2000-2004 to 2005-2009 and then decreased from 2005-2009 to 2010-2014.
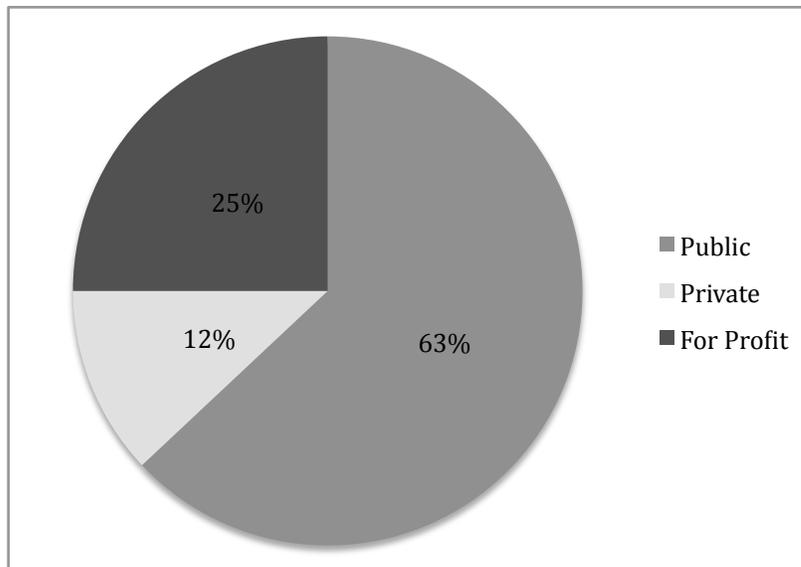
Figure 4. Percentage of Dissertations by University Type

Private universities accounted for 488 or 12% of the dissertations included in the study, Figure 4. These private dissertations came from 155 or 40% of the universities, Figure 5. Over each of the three time periods both number of private dissertations and number of private universities saw increased numbers.
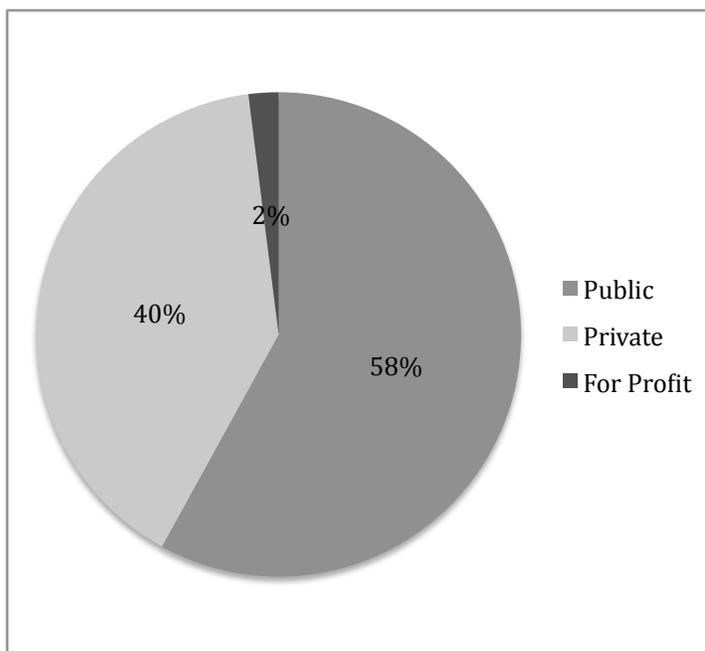
Figure 5. Percentage of Universities by University Type

For Profit Universities accounted for 630 dissertations, which was 25% of the study sample. Ten unique for-profit universities accounted for only 2% of the total universities in the study. While there was a large increase in dissertation numbers across the three time periods 49, 243, and 338 respectively, there was little increase in the number of universities over the years.

The data indicated that the majority, 63%, of graduate researcher in distance education came from public institutions. This remained the case across the three time periods.   For Profit universities came in at second with 25% and Private universities accounted for just 12% of graduate research in the field.

For Profit universities produced dissertations at a much higher rate per university. Note that 2% of the universities included in the study produced 25% of the dissertations. To further illustrate this Tables 12, 13, and 14 show the top 10 universities for each university type and number of dissertations over the three time periods.

Capella University increased number of dissertations from 28 over the 2000-2004 time period to 174 over the 2005-2009 time period. It increased to 194 over the 2010-2014 time period. Table 13 shows that during the 2005-2009 time period Capella University produced 6 times more dissertations than the next most frequent degree granting institution Nova Southeastern University, a private university with 29 dissertations. This trend remained true over the 2010-2014 time period when Cappella produced 3 times as many dissertations as the second most frequent, also a For Profit school, Walden University which produced 60 dissertations. In similar fashion, during that same time period, 2010-2014,Walden nearly doubled the third most frequent dissertation granting university, Nova Southeastern with 32.

Public		Private		Profit	
The Pennsylvania State University	25	University of Southern California	16	Capella University	28
Texas A & M University	22	The George Washington University	11	Walden University	11
Indiana University	21	Columbia University Teachers College	11	University of Phoenix	4
The University of Nebraska - Lincoln	20	Pepperdine University	11	University of Sarasota	3
Virginia Polytechnic Institute and State University	19	Nova Southeastern University	9	Argosy University/Sarasota	2
University of North Texas	17	University of San Francisco	9	Northcentral University	1
University of Central Florida	16	The Fielding Institute	7		
The Florida State University	15	The Union Institute	6		
Purdue University	14	New York University	6		
Arizona State University	13	Regent University	6		

Table 12. University Types 2000-2004 - University and Number of Dissertations

Public		Private		Profit	
The Pennsylvania State University	25	Nova Southeastern University	29	Capella University	174
Indiana University	23	Regent University	14	Walden University	22
University of Central Florida	23	Teachers College, Columbia University	13	Northcentral University	22
University of South Florida	21	University of Southern California	13	University of Phoenix	19
George Mason University	19	Fielding Graduate University	11	TUI University	5
Oklahoma State University	18	The George Washington University	11	Argosy University/Chicago	1
Arizona State University	17	Pepperdine University	10		
The University of Nebraska - Lincoln	17	Wilmington College (Delaware)	9		
Texas A & M University	17	Alliant International University, San Diego	7		
North Carolina State University	17	Union Institute and University	6		

Table 13. University Types 2005-2009 - University and Number of Dissertations

Public		Private		Profit	
Indiana University of Pennsylvania	18	* Nova Southeastern University	32	* Capella University	194
* Indiana State University	16	* Liberty University	24	* Walden University	60
* The University of Nebraska - Lincoln	15	* University of Southern California	24	* Northcentral University	51
Northern Illinois University	15	Pepperdine University	13	* University of Phoenix	18
The Florida State University	15	* Regent University	13	* TUI University	11
* Old Dominion University	15	Boston University	10	* Colorado Technical University	2
University of South Florida	14	Drexel University	10	* Argosy University/Phoenix	1
The Pennsylvania State University	14	* Fielding Graduate University	9	* Argosy University/Chicago	1
* Michigan State University	13	* Northeastern University	8		
* The University of Alabama	13	Teachers College, Columbia University	8		

Table 14. University Types 2000-2004 - University and Number of Dissertations
* Offered online PhD or EdD programs in one or more of department types found in this study

Public Universities that remained in the top ten across all three periods included

The Pennsylvania State University and The University of Nebraska – Lincoln. Research

at The Pennsylvania State University largely did not include department info (58

dissertations), but studies from Psychology (1) and Higher Education (2) were listed.

Dissertations from The University of Nebraska – Lincoln included the Education (14) and Ed Lead (11) department types.

Private Universities that are included in the top ten across all three time periods include Nova Southeastern, The Fielding Institute, and Regent University. Departments from Nova Southeastern included Educational Technology (41). The Fielding Graduate University included dissertations from EdLead (6) and Psychology (2).

For Profit Universities included across all three time periods included Capella University, Walden University, University of Phoenix and North Central University. Doctoral students at Walden graduated from Education (48), Business (13), and Psychology (16) departments. Departments represented by North Central University included Business (3) and not listed (60). The University of Phoenix included dissertations largely not classified with a department (38).

Capella University was the top ranking for four of the top six department types with the highest dissertation frequency detailed in Table 10, as well as having the most dissertations with no department recorded category. Capella's department types that granted the highest number of dissertations in this study included Education (268), Miscellaneous (11), Educational Psychology (13), and Business (25). There were 75 dissertations with no department recorded. Results indicated that Capella University granted more dissertations and was the top granter in more department types than any other institution over the past fifteen years. One out of every four dissertations investigating distance education came from Capella.

Capella was not the only university serving online doctoral programs where students wrote dissertations on distance education. Evans and Green (2013) discuss

professional doctorates, often mid-career professional people that are attracted to the

flexibility that online doctoral studies provide. Table 14 details that during the 2010-

2014 time period eight out of eight For Profit institutions granted online PhD or EdD

degrees. Online doctoral degrees were also granted by at least onn department type

included in this study by six out of ten private institutions and five out of ten public

institutions. The majority of degrees granted by private and public institutions were EdD

degrees. As demand for flexibility of study increases, more options to pursue doctoral

studies online will become available. Traditional brick and mortar universities will be

faced with the decision to make their programs available online or students may take their

study elsewhere.

Sub-Questions C-G: Reference Sections

Sub-questions C-G focused on dissertation reference sections.   Reference

sections were not available for dissertations listed in the PQDT before 2008. Table 15

details the number of dissertations with reference sections by year available for this

study.

2008	23
2009	30
2010	18
2011	35
2012	35
2013	34

2014	163
Total	338

Table 15. Total Number of Dissertations with Reference Sections Listed in PQDT

The table shows that only 16% of the 2064 dissertations from 2008-2014 included reference sections. Therefore the results addressing the reference section-based Sub-Questions C-G only represented a small sub-sample and not the full sample of 3954 dissertations. There was a marked increase in dissertations including reference sections in 2014, 163 out of 247 or 65% included reference sections. It is possible that in future more thorough analysis of dissertation reference sections can be made if this increase becomes a trend.

The small sub-sample of citations included in this study may act to point towards the types of sources graduate students used to support their research.

"Which authors received the highest frequency of citation in dissertation reference sections?" Sub-Question C

Frequently cited authors of both books and journal articles were compared between the 2005-2009 and 2010-2014 time periods, Table 16. The National Library of Medicine, the most frequently cited in both time periods, and the Sloan Consortium were the two institutional authors highly rated on both lists. Garrison, R., Allen, E., and Seaman, J. were the individual authors with the highest frequency of citation across both time periods. Of the initial twenty five authors from 2005-2009 eleven, indicated by *, were also listed in the top twenty five for 2010-2014.

2005-2009	6475	2010-2014	44626
* National Library of Medicine.	44	* National Library of Medicine.	394
* Garrison, R	31	* Seaman, J	306
* Allen, E	22	* Allen, E	273
* Seaman, J	22	* Sloan Consortium.	251
* Sloan Consortium.	22	* Garrison, R	121
* Anderson, T	22	* Rovai, A	114
* Rovai, A	21	Creswell, J	97
* Palloff, R	18	Bandera, A	95
* Pratt, K	18	* Anderson, T	88
* Knowles, M	18	Means, B	67
Bonk, C	17	Onwuegbuzie, A	66
Rourke, L	12	Palloff, R	65
Mezirow, J	12	Pratt, K	65
Merriam, S	12	Toyama, Y	65
Clark, R	11	Jones, K	63
* Archer, W	11	Murphy, R	62
Arbaugh, J	11	Bakia, M	62
Barb, S	10	Lincoln, Y	57
Gunawardena, C	10	Barbour, M	55
Simonson, M	10	* Knowles, M	54
Berge, Z	9	Patton, M	51
Jonassen, D	9	Cavanaugh, C	50
Anonymous.,	9	* Archer, W	50
Tu, C	8	Reeves, T	49
Kanuka, H	8	Mayer, R	48

Table 16. Top Twenty Five Frequently Occurring Authors 2000-2004 and 2005-2009
* Authors Listed during both time periods

Many of the authors listed during the first time period that were not shown in the top twenty five from 201-2014 still ranked within the top fifty cited authors (not shown); including Berge, Z., Bonk, C., Gunawardena, C, Simonson, M., and Arbaugh, J. Bandera, A. and Creswell, J. Of the authors new to the top twenty five listing for 2010-2014 only three were not listed in the top 50 (not shown) of the 2005-2009 time period; Toyama, Y., Murphy, R., and Barbour, M.

"What journal articles were cited with the highest frequency in dissertation reference sections?" Sub-Question D.

The top twenty most frequently cited journal articles listed in dissertation reference sections for the 2005-2009 time period were detailed in Table 17 and the 2010-2014 time period in Table 18. Only two articles from the first time period, Garrison's "Critical thinking, cognitive presence, and computer conferencing in distance education" and Rourke's "Assessing social presence in asynchronous text-based computer conferencing", were also included in the top twenty for the second time period. Four articles, denoted by #, listed between 2005-2009 did not make the top twenty yet made the top fifty. Newer articles rose to the top of the most frequently cited article list of a given time period and highly cited articles of a given time period were not guaranteed to carry over with as frequent a citation rate to subsequent time periods.

* Critical thinking, cognitive presence, and computer conferencing in distance education	6
Bridging the transactional distance gap in online learning environments	4
# Facilitating cognitive presence in online learning: interaction is not enough	4
# Mixed methods research: a research paradigm whose time has come	4
Self-regulation in a web-based course: a case study	4
* Assessing social presence in asynchronous text-based computer conferencing	4
Self-directed learning: toward a comprehensive model	4
Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing	4
Media will never influence learning	4
The learning styles, expectations, and needs of online students	4
# The future of online teaching and learning in higher education: the survey says...	4
Satisfaction, academic rigor and interaction: perceptions of online instruction	4
# Toward constructivism for adult learners in online learning environments	4
On defining distance education.	4
Situated learning	4
Cognitive style and self-efficacy: predicting student success in online distance education	4
Synchronous and asynchronous text-based cmc in educational contexts: a review of recent research	3
The role of students' cognitive engagement in online learning	3
Learning with invisible others: perceptions of online presence and their relationship to cognitive and affective learning	3
Cognitive presence in online learning	3

Table 17. Top Twenty Frequently Occurring Articles 2005-2009

Teaching courses online: a review of the research	20
Student barriers to online learning: a factor analytic study	19
How does distance education compare with classroom instruction? a meta-analysis of the empirical literature	19
* Critical thinking, cognitive presence, and computer conferencing in distance education	17
The reality of virtual schools: a review of the literature	17
A review of benefits and limitations of online learning in the context of the student, the instructor, and the tenured faculty	16
Self-efficacy: toward a unifying theory of behavioral change	16
Digital natives, digital immigrants part 1	15
Distance education trends: integrating new technologies to foster student interaction and collaboration	14
Factors that influence students' decision to dropout of online courses	14
As distance education comes of age, the challenge is keeping the students	14
What's the difference: a review of contemporary research on the effectiveness of distance learning in higher education	14
Going the distance with online education	13
Preparing instructors for quality online instruction	13
Instructor-learner interaction in online courses: the relative perceived importance of particular instructor actions on performance and satisfaction	13
Research and practice in k-12 online learning: a review of open access literature	13
* Assessing social presence in asynchronous text-based computer conferencing	13
Factors influencing adult learners' decision to drop out or persist in online learning	12
Blended learning and sense of community: a comparative analysis with traditional and fully online graduate courses	12
An overview of online education: attractiveness, benefits, challenges, concerns and recommendations	12

Table 18. Top Twenty Frequently Occurring Articles 2010-2014

The most frequently cited articles between 2005-2009 predominately focused on distance education research. Three articles reported research on learning not specific to the online environment; including situated and self-directed learning. One article from this time period focused on mixed methods research. All but one of the articles cited between 2010-2014 focused on an area of research other than distance learning. This one article forwarded a self-efficacy theory. This suggested that the articles graduate researchers chose were most frequently used to ground their dissertations in the distance education literature base. Studies that addressed theory, adult learning, or methodological issues out of the context of distance education were not found among the most frequently cited sources.

"What journal publications were cited with the highest frequency in dissertation reference sections?" Sub-Question E.

Tables 19 and 20 show the most frequently cited publications in distance education dissertations during the 2005-2009 and 2010-2014 time periods respectively. There was not much variation between the two time periods. There was some shifting of journal rank between two tables. But only three journals listed from 2005-2009 were not also found included between 2010-2014; International Journal on ELearning, Adult Education Quarterly, and Journal of Distance Education. The new additions to the 2010-2014 time period were Review of Educational Research, Adult Education Quarterly, and Journal of Distance Education.

* American Journal of Distance Education	53
* Quarterly Review of Distance Education	48
* Distance Education	47
* Educational Technology Research and Development	46
* British Journal of Educational Technology	38
* Internet and Higher Education	34
* Journal of Research on Technology in Education	29
* Online Journal of Distance Learning Administration	28
* International Review of Research in Open and Distance Learning	27
* TechTrends	26
* Computers & Education	26
International Journal on ELearning	24
* The Chronicle of Higher Education	22
Adult Education Quarterly	22
Journal of Distance Education	21

Table 19. Top Fifteen Frequently Occurring Publications 2005-2009

* Quarterly Review of Distance Education	311
* Distance Education	303
* Computers & Education	303
* Internet and Higher Education	248
* British Journal of Educational Technology	223
* Online Journal of Distance Learning Administration	217
* American Journal of Distance Education	216
* International Review of Research in Open and Distance Learning	193
Review of Educational Research	189

* Educational Technology Research and Development	160
* TechTrends	130
Journal of Asynchronous Learning Networks	130
* The Chronicle of Higher Education	124
* Journal of Research on Technology in Education	114
Journal of Nursing Education	109

Table 20. Top Fifteen Frequently Occurring Publications 2010-2014

During the 2005-2009 time period all but one journal did not include focus on distance education research. The Chronicle of Higher Education focuses on adult education research not specific to technology, yet does sometimes include research on distance education. Three journals not focused on distance education research were found among the between 2010-2014. Two focus on learning research, Review of Educational Research and The Chronicle of Higher Education. The third non-distance education focused publication was The Journal of Nursing Education. Again the predominance of distance education publications indicated that graduate researchers cited distance education oriented journal research more often than general education, learning, or methodological research.

The inclusion of the Journal of Nursing Education indicated that journals focused on disciplines are beginning to more frequently publish distance education research. These discipline-based journals may be of interest for future reviews of in distance education research. Results demonstrated that four department types regularly granted distance education focused dissertations from 2007 to 2014; Nursing, Psychology, Business, and Computer Science. These four disciplines may act a starting point to investigate discipline-based distance education research.

"What book titles were cited with the highest frequency in dissertation reference

sections?" Sub-Question F

The twenty books that graduate researchers cited most frequently during the two

time periods are shown in Tables 21 and 22. Nine of the books listed in the top twenty

from 2005-2009, indicated by *, were also listed in the top twenty from 2010-2014.

* Building learning communities in cyberspace: effective strategies for the online classroom. Josses-bass higher and adult education series	9
# Using multivariate statistics (5th ed.)	7
* Teaching and learning at a distance: foundations of distance education	7
# Psychology of learning for instruction	6
# Lessons from the cyberspace classroom: the realities of online teaching	5
* The foundations of distance education	5
* The modern practice of adult education: andragogy versus pedagogy	5
* Diffusion of innovations	5
* The adult learner. the definitive classic in adult education and human resource development. (5th ed.)	5
Thought and language	5
# Transformative dimensions of adult learning	4
* Experiential learning : experience as the source of learning and development	4
Learning as transformation: critical perspectives on a theory in progress. The josses-bass higher and adult education series	4
The virtual student. a profile and guide to working with online learners. The josses-bass higher and adult education series	4
# The world is flat: a brief history of the twenty-first century	3
Understanding by design (2nd ed.)	3
Designing and conducting mixed methods research	3

Applied multiple regression/correlation analysis for the behavioral sciences (3rd ed.)	3
* Qualitative evaluation and research methods	3
How people learn: brain, mind, experience, and school	3

Table 21. Top Twenty Frequently Occurring Books 2005-2009

Research design: qualitative, quantitative, and mixed methods approaches	51
* Building learning communities in cyberspace: effective strategies for the online classroom. josses-bass higher and adult education series	34
* Qualitative research and evaluation methods	32
* Experiential learning : experience as the source of learning and development	24
Mind in society: the development of higher psychological processes	24
* Teaching and learning at a distance: foundations of distance education	23
Naturalistic inquiry	22
* The modern practice of adult education: andragogy versus pedagogy	20
Democracy and education: an introduction to the philosophy of education	19
Educational research: competencies for analysis and application	19
* Designing and conducting mixed methods research	18
* The adult learner. the definitive classic in adult education and human resource development. (5th ed.)	15
Leaving college: rethinking the causes and cures of student attrition. (2nd ed.)	15
* The foundations of distance education.	15
E-learning in the 21st century: a framework for research and practice	15
How to design and evaluate research in education (2nd ed.)	14

Social learning theory	14
Theoretical principles of distance education	14
* Diffusion of innovations	14
Statistical power analysis for the behavioral sciences	14

Table 22. Top Twenty Frequently Occurring Books 2010-2014

Another five books from the earlier time period were also listed among the top fifty most cited books in the later time period, indicated by # in Table 21. Two books, indicated by # in Table 22, listed in the top fifty most cited books in the earlier time period were also listed among the top twenty in the later time period.

Unlike journal articles distance education was not the sole focus of the most cited books in either time period. Rather during the 2005-2009 time period books focused on learning (8), distance education (5), statistics (3), theory (2), and research methodology (2). During the 2010-2014 time period books focused on learning (6), methodology (6), distance education (5), theory (2), and statistics (1). These results indicated that the graduate researchers did utilize books to place their research in the greater distance education discourse. Yet more often, books were cited in dissertations to address other aspects of the research process; specifically learning research and theory, methodology, and statistics. Also books had a longer applicability as indicated by nearly half of the most cited books in the earlier time period also remaining highly cited during the second time period. Also books often addressed topics of longevity. The results indicated that learning communities, distance education foundations, diffusion of innovation, and adult learning are all topics of longevity in the distance education field.

"What are the relationships among dissertation reference citations using co-citation social network analysis?" Sub-Question G

Frequency analysis of dissertation citations enabled the identification of areas that graduate researchers used as the theoretical foundations for their dissertations. Graduate students tended to use journal articles to ground dissertations in distance education research, while books were often utilized to ground dissertations in methods, theory, and topics of longevity in the distance education field. Yet the structural relationships between citations cannot be grasped by frequency analysis alone. This study employed SNA to overcome this issue. First co-citation network features and centrality measurements were investigated. Then visualization was used to investigate co-citation network features.

Inclusion of a dissertation's reference section was not a requirement for inclusion in the PQDT and did not become available as an option until 2006. Therefore not all dissertation records included in the study were included in the secondary citation section sample. Of this study's n=3,954 sample, 338 (9%) of the dissertation listings included reference section citations. A further limitation with the co-citation network analysis was that the SITKIS software had the ability to only include one author. While bibliometric analysis of authors enabled multiple authors to be included, SNA analysis here included only the first author of a publication. To counter for this the study, not the author, was considered to be the unit of analysis for comparison.

Co-citation indicates a relationship between two citations that are cited in the same citing document. The strength of the relationship between A and B is based on the

number of citing documents that cite both A and B in their reference sections (Waugh &
Ruppel, 2004).

	Number of total elements		Frequently Occurred Pairs	
Period	Nodes	Ties	Nodes	Ties
2007-2009	71	1,860	20	46 (N>3)
2010-2014	932	2610	16	54(N>5)

Table 23. Description of the Co-Citation Network Elements

Table 23 describes the elements of the network. During the first time period,
2007-2009, a total of 71 nodes and 1,860 ties were identified. Among them, 20 nodes
with 46 ties occurred 3 times or more. During the second time period 2010-2014, a total
of 932 nodes and 2610 ties were identified. Among them, 16 nodes with 54 ties occurred
5 times or more. The frequency threshold for the pairs between 2005-2009 was initially
set to five. Only four nodes were included at this threshold. The threshold was lowed to
3 and the number of nodes and ties were found to be comparable to the 2010-2014 set.

Network Measurements: Centrality, Betweeness, and Closeness

Centrality is one of the most critical and popular conceptual tools in SNA in that
it can measure and analyze the influence of nodes in a network (Everett & Borgatti,
2013). According to Brass and Burkhardt (1992), a node with a higher centrality in an
organization tends to have a higher impact. Freeman (1979) suggested three different
types of centrality: degree centrality, betweeness centrality, and closeness centrality.
Degree centrality is defined as the number of direct linkages between a node and other
nodes in a network. In this study's context, degree centrality was related to the number of
citations associated with a given citation. A citation with the highest score on degree

centrality had the most number of direct connections to other citations in the network and as such served as a hub in the network.

The concept of betweeness is a measure of how often a node is placed on the shortest path between any other two nodes in a network. A node with a higher betweeness centrality plays a critical role in the flow of information and resources within the network, and as such is more likely to be placed at the center of the network. In the present context, it meant that a citation with a high betweeness centrality score linked important sets of citations.

Closeness centrality refers to the "inverse of the average length of the shortest paths to/from all the other actors in the network" (Lee & Su, 2010). A node with a higher closeness centrality can obtain information more easily. In this study's present context, a citation with a the high closeness centrality score (i.e., having shortest paths to other citation) was closely linked to the other citations in the network. Tables 24 and 25 detail citation centrality measurements across the two time periods.

	Degree	Closeness	Between
Palloff R-BUILDING LEARNING COMMUNITIES IN CYBERSPACE: EFFECTIVE STRATEGIES FOR THE ONLINE CLASSROOM. JOSSEY-BASS HIGHER AND ADULT EDUCATION SERIES.-1999-0	9	54	66
Garrison D-AMERICAN JOURNAL OF DISTANCE EDUCATION-2001-15	5	89	8
Clark R-EDUCATIONAL TECHNOLOGY RESEARCH AND DEVELOPMENT-1994-42	4	62	12.5
Tabachnick B-USING MULTIVARIATE STATISTICS (5TH ED.).-2007-0	3	70	12.5
Anderson T-JOURNAL OF DISTANCE EDUCATION-1999-14	2	92	0
Carr S-THE CHRONICLE OF HIGHER EDUCATION-2000-46	2	62	10
Freire P-PEDAGOGY OF THE OPPRESSED.-1987-0	2	92	0

Gunawardena C-JOURNAL OF EDUCATIONAL COMPUTING RESEARCH-1997-16	2	92	0
Holmberg B-THEORY AND PRACTICE OF DISTANCE EDUCATION. SECOND EDITION.-1995-0	2	64	0
Kanuka H-JOURNAL OF COMPUTING IN HIGHER EDUCATION-2004-15	2	92	0
Kozma R-EDUCATIONAL TECHNOLOGY RESEARCH AND DEVELOPMENT-1994-42	2	64	0
Mezirow J-LEARNING AS TRANSFORMATION: CRITICAL PERSPECTIVES ON A THEORY IN PROGRESS. THE JOSSEY-BASS HIGHER AND ADULT EDUCATION SERIES.-2000-0	2	64	12
Stein D-THE AMERICAN JOURNAL OF DISTANCE EDUCATION-2005-19	2	62	10
Bonk C-USDLA JOURNAL-2002-16	1	82	0
Clark R-REVIEW OF EDUCATIONAL RESEARCH-1983-53	1	74	0
Garrison D-ADULT EDUCATION QUARTERLY-1997-48	1	66	0
Garrison D-AMERICAN JOURNAL OF DISTANCE EDUCATION-2005-19	1	93	0
Hiltz S-COMMUNICATIONS OF THE ACM-1997-40	1	66	0
Keegan D-THE FOUNDATIONS OF DISTANCE EDUCATION.-1986-0	1	66	0
Rourke L-JOURNAL OF INTERACTIVE LEARNING RESEARCH-2002-13	1	76	0

Table 24. Co-Citation Centrality Measurements 2007-2009

	Degree	Closeness	Between
Creswell J-RESEARCH DESIGN: QUALITATIVE-2009-0	5	39	22
Patton M-QUALITATIVE RESEARCH AND EVALUATION METHODS-2002-0	4	40	18
Garrison D-AMERICAN JOURNAL OF DISTANCE EDUCATION-2001-15	2	54	1
Anderson T-JOURNAL OF DISTANCE EDUCATION-1999-14	1	55	0
Appana S-INTERNATIONAL JOURNAL ON ELEARNING-2008-7	1	46	0
Barbour M-COMPUTERS & EDUCATION-2009-52	1	57	0

Bean J-REVIEW OF EDUCATIONAL RESEARCH-1985-55	1	57	0
Cavanaugh C-INTERNATIONAL REVIEW OF RESEARCH IN OPEN AND DISTANCE LEARNING-2009-10	1	57	0
Gunawardena C-JOURNAL OF EDUCATIONAL COMPUTING RESEARCH-1997-16	1	55	0
Knowles M-THE MODERN PRACTICE OF ADULT EDUCATION: ANDRAGOGY VERSUS PEDAGOGY-1970-0	1	47	0
Lincoln Y-NATURALISTIC INQUIRY.-1985-0	1	47	0
Palloff R-BUILDING LEARNING COMMUNITIES IN CYBERSPACE: EFFECTIVE STRATEGIES FOR THE ONLINE CLASSROOM. JOSSEY-BASS HIGHER AND ADULT EDUCATION SERIES.-1999-0	1	46	0
Phipps R-JOURNAL OF DISTANCE EDUCATION-1999-14	1	46	0
Prensky M-ON THE HORIZON-2001-9	1	46	0
Rubin H-QUALITATIVE INTERVIEWING: THE ART OF HEARING DATA-2012-0	1	47	0
Tinto V-LEAVING COLLEGE: RETHINKING THE CAUSES AND CURES OF STUDENT ATTRITION. SECOND EDITION.-1993-0	1	57	0

Table 25. Co-Citation Centrality Measurements 2010-2014

Between 2007-2009 Palloff, 1999; Garrison, 2001; and Clark, 1994 all had high centrality measurements, suggesting that these citations acted as hubs within the network. Citations such as Anderson, 1999; Freire, 1987; Gunawardena, 1997; and other all showed a high closeness centrality of 92 suggesting that these citations may have belonged to a single or multiple cliques (or clusters) separate from the other citations. Finally Palloff, 1999 had a much higher betweeness centrality measurement than the other nodes. This indicated that a node that had a tie with Palloff, 1999 could get to any other node in the network faster. This in combination with Palloff's high centrality measurement suggested that this node was at the center of a larger cluster of citations within the network.

Between 2009-2014 Creswell, 2009 and Patton, 2002 were the most central nodes in the network. Nodes such as Barbour, 2009; Bean, 1987, and others with the high closeness score of 57 were likely to be in a separate cluster or clusters from the main cluster of citations. Creswell, 2009 and Patton, 2002 also had the highest betweeness measurements, suggesting these citations each would be at the center of a cluster of citations around them.

While centrality measurements indicated the key nodes within a network, they did not give the full picture of the relationship of nodes. Visualizations using MSL graphs with UciNet's NetDraw application were performed and detailed in Figures 6 and 7.
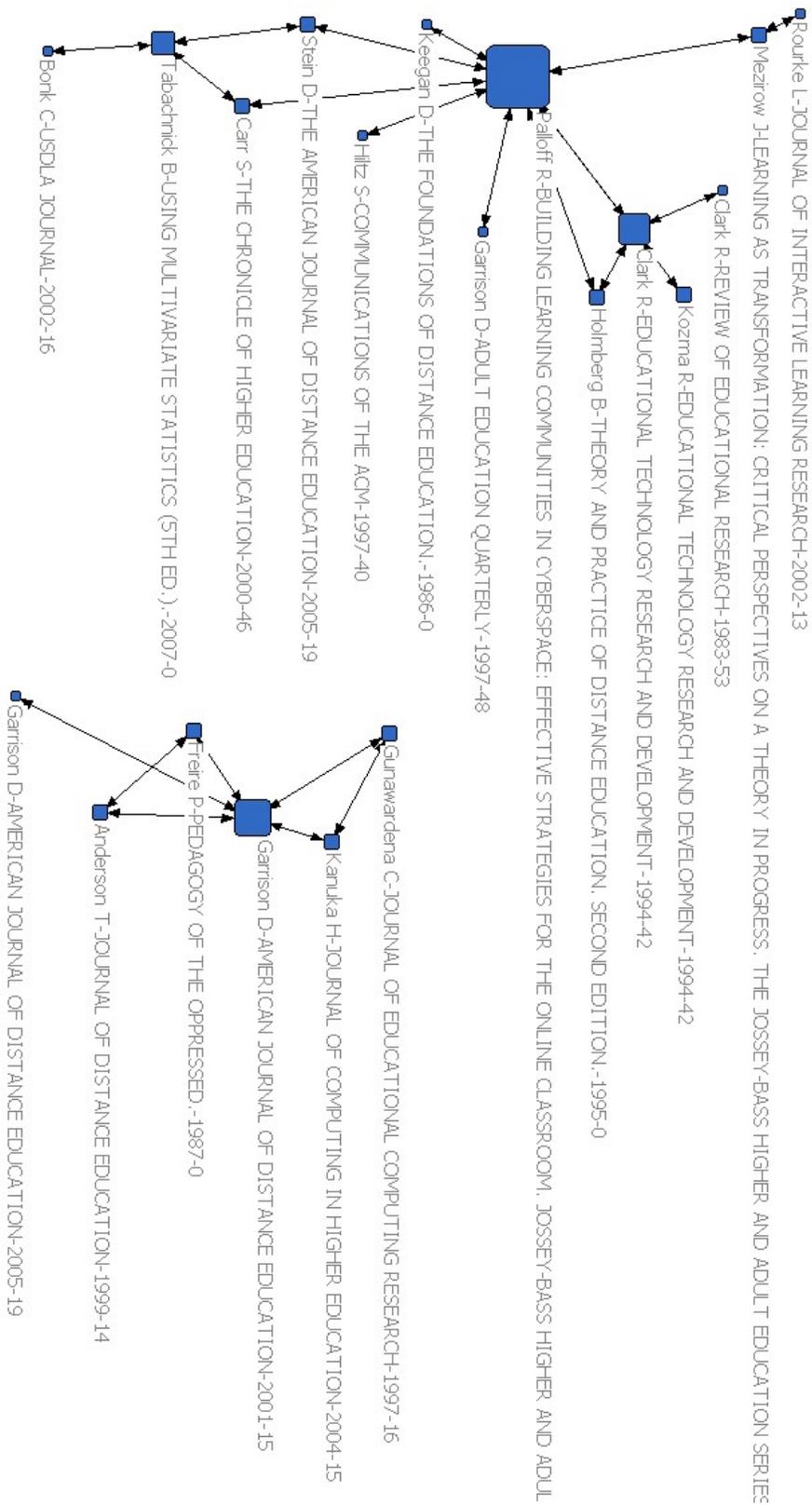
Figure 6. Co-Citation Network 2005-2009
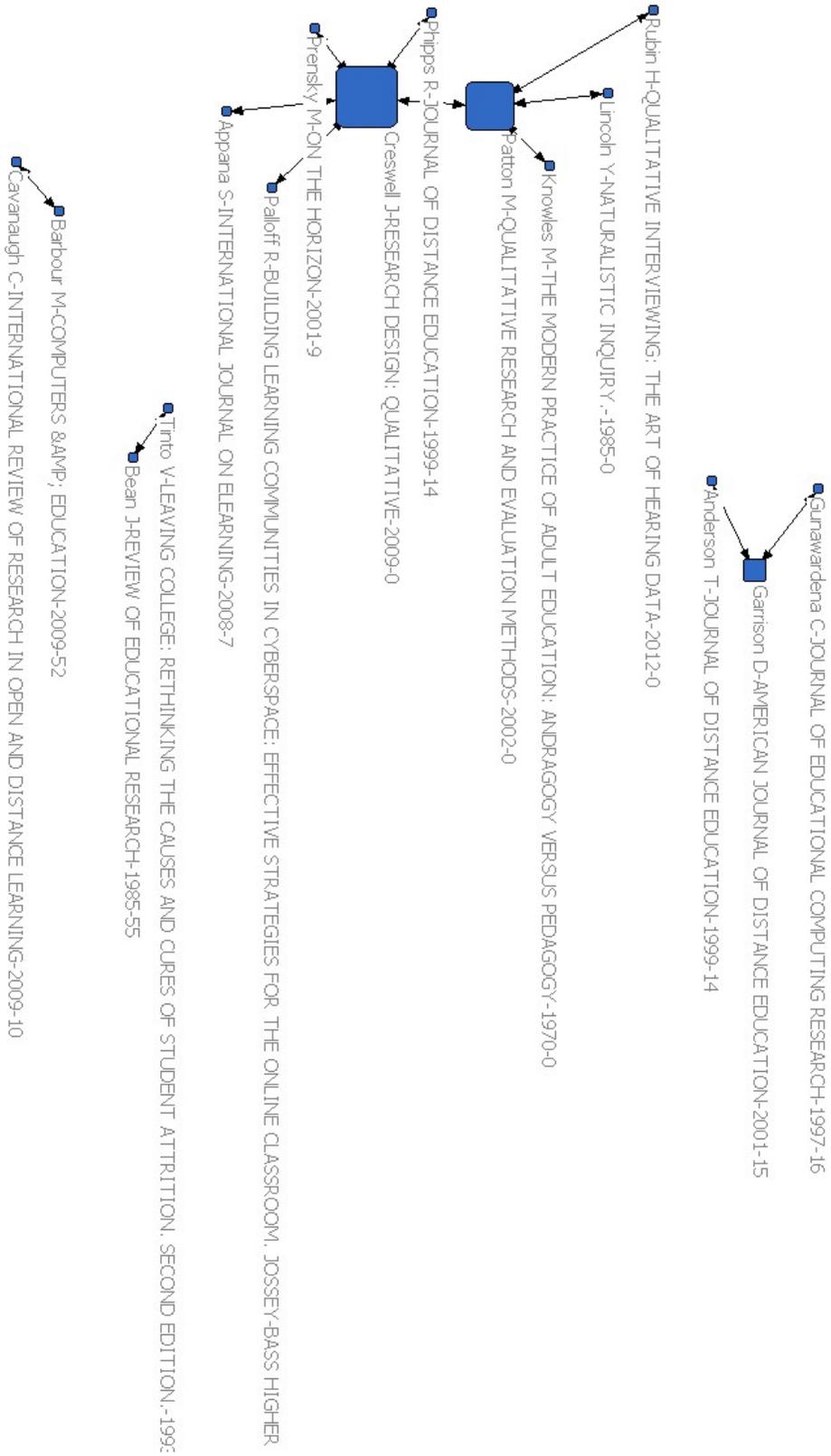
Figure 7. Co-Citation Network 2010-2014

The co-citation network from 2005-2009, Figure 6, was constructed of two clusters. The larger cluster to the left illustrated Palloff, 1999 as the central node. The second smaller cluster showed Garrison, 2001 at the center.

The Garrison, 2001 node was from a study focused on cognitive presence. The Kanuka, 2004 and Garrison, 2005 nodes were also cognitive presence oriented. The other nodes in this smaller cluster included social construction of knowledge (Gunawardena, 1997), co-creation of knowledge, Freire (1987) and constructivism (Anderson, 1999). These results indicated that the topics of cognitive presence and social-construction of knowledge were highly related. The two topics were commonly cited in studies together. And most frequently sources were cited alongside Garrison, 2001.

The larger cluster contained nodes that addressed a variety of topics. Pallof, 1999, at the center of the cluster, dealt with learning communities and instructional design. Stemming off this main node was a small cluster of Clark, 1992 & 1994, and Kozma, 1994. These nodes were from the debate between the two authors arguing if specific technology had cognitively relevant capabilities. Nodes that focused on theory of distance education included Holmberg, 1985 and Keegan, 1986. The only node addressing methods was Tabachnick's 2007 statistics book. Nodes addressing issues of design included Garrison, 1997 and Bonk, 2002. Issues of communication between learners were found in the Rourke, 2002 and Stein, 2005 nodes.

The network for citations between 2010-2014, Figure 7, included four clusters. The top cluster Gunawardena, 1997,; Garrison, 2001; and Anderson, 1999 illustrated that the strong relationship between social construction of knowledge and cognitive presence

remained important in the latter time period. This trend indicated that social construction of knowledge and cognitive presence remained two of the most popular research areas dissertations cited and that they were most frequently cited in tandem.

The larger cluster showed Creswell, 2009 and Patton, 2002 to be the most central nodes. Both of these were from methodological books. Other methods based nodes included Rubin, 2012 and Lincoln, 1985. All of the methods based nodes detailed qualitative methods specifically. Other nodes in this cluster focused on distance education context (Phipps, 1999), learners (Knowles, 1979; Prensky, 2001), and benefits and limitations (Appana, 2008). Of note Pallof, 1999, which focused on learning communities and instructional design, moved from being the central node in the earlier period to having only a single tie in the later period.

Two smaller clusters at the bottom included Tinto, 1993 and Bean, 1985. These authors both addressed attrition, a research area found to be common in dissertations that address distance learning management. Finally Cavanaugh, 2009 and Barbour, 2009 were related by both addressing distance education in the K-12 context.

Results from the two social networks show a trend that social construction of knowledge and cognitive presence remained highly related in dissertation reference sections across both time periods. Two nodes focusing on methods in the second network replaced a highly central node focused on learning communities in the first network. Only a single node addressing methods was found in the first network, while 44% of the nodes in the main cluster addressed qualitative methods in the second network. This indicated a trend during the latter time period that the most centrally related dissertation citations were from the area of qualitative methods. Issues of

technology were highly cited in the first network, but missing in the second. Citations that dealt with issues of online learning communities and instructional design were present across both time periods, although shown to be less central in the second time period.

Further discussion of the citation analysis and co-citation analysis findings are addressed in Chapter V.

Sub-Questions H-J – Research Topics: Classification, Keyword, and Emergent Keyword Analysis

Sub-questions H – J each addressed one indicator of dissertation research topic. Classifications were topics chosen by the doctoral student from a framework of categories provided by the PQDT. Keywords were chosen by the doctoral student with no constraints from the PQDT. Emergent keywords were algorithmically extracted from dissertation abstracts.

"What PQDT dissertation classifications received the highest frequency?" Sub-Question H

When graduate researchers submitted a dissertation to the PQDT database they were asked to characterize the dissertation subject using a framework developed by the PQDT. Authors were asked to select a primary classification that best described the research. They were allowed to choose two further classifications that were indexed as secondary classifications for the work. The classifications framework was arranged by subject categories and each main category had a number of sub-categories that was more

specifically focused within a discourse. The PQDT classification framework is listed in Appendix 6.

Classifications offer a unique opportunity for research to be categorized by the author, yet held to a common framework developed by the PQDT. The most frequently chosen classifications and counts for all dissertations spanning each of the three time periods are detailed in Table 26. The total number of classifications for all dissertations across years is 13,921.

Educational Software (2213 classifications across all years) and *Educational Technology* (1403 classification across all years) were the most frequently chosen classifications across all three-time periods. Educational technology was not available as an option for PQDT classification until the second time period 2005-2009. During the second two time periods, both Educational Software and Educational Technology were listed in the top three classifications.

2000-2004	5777	2005-2009	4232	2010-2014	3912	All Years	13921
Educational software	1524	Educational software	689	Educational technology	932	Educational software	2213
Higher education	667	Higher education	498	Higher education	388	Higher education	1553
Teaching	378	Educational technology	460	Instructional Design	228	Educational technology	1403
Curricula	378	Teacher education	236	Adult education	184	Teacher education	774
Continuing education	354	Adult education	228	Educational leadership	156	Adult education	578
Adult education	354	Curricula	200	Community college education	114	Curricula	578
School administration	199	Teaching	200	Teaching	114	Teaching	578
Community colleges	182	Continuing education	184	Continuing education	112	Continuing education	506
Teacher education	159	School administration	152	Teacher education	103	School administration	412
Educational psychology	153	Educational psychology	142	Educational psychology	86	Community colleges	408
Computer science	70	Curriculum development	137	Community colleges	76	Educational psychology	382
Information Systems	68	Secondary education	133	School administration	71	Community college education	263
Health education	64	Higher Education Administration	115	Curriculum development	65	Instructional Design	249
Nursing	48	Community colleges	81	Secondary education	64	Curriculum development	238
Business education	44	School administration	76	Community college education	50	Secondary education	181
Language arts	43	Secondary education	115	Community college education	50	Community college education	177
Secondary education	42	Community college education	63	Computer science	48	Educational leadership	163
Management	41	Educational psychology	86	Instructional Design	46	Computer science	154
Mathematics education	39	Bilingual education	81	Nursing	44	Nursing	136
Mathematics education	39	Mathematics education	41	Mathematics education	44	Mathematics education	123
Educational evaluation	31	Health education	53	Business education	42	Business education	119
Mass media	30	Pedagogy	46	Health education	39	Management	112
Academic guidance counseling	30	Education	40	Management	37	Health education	111
Science education	29	Science education	35	Science education	34	Science education	103
Multicultural education	26	Information systems	30	Information science	31	Language arts	99
Vocational education	26	Language arts	32	Computer science	30	Higher Education Administration	90
Bilingual education	26	Information science	26	Continuing education	30	Health education	83
Composition	25	Special education	27	Information Systems	31	Web Studies	75
Rhetoric	25	Multicultural education	24	Special education	24	Special education	74
Cognitive therapy	24	Vocational education	18	Bilingual education	23	Bilingual education	70
Agricultural education	22	Educational evaluation	18	Organizational behavior	21	Educational evaluation	66
Special education	22	Religious education	17	Educational evaluation	21	Vocational education	63
Communication	22	Literacy	13	English as a Second Language	21	Information science	60
Inservice training	22	Health education	13	Communication	21	Pedagogy	53
Womens studies	21	Rhetoric	13	Multicultural Education	19	Communication	52
Health education	20	Occupational psychology	13	Information science	19	Multicultural education	50
Library science	19	Academic guidance counseling	10	Vocational education	19	Information Technology	50
Educational sociology	18	Reading instruction	13	Pedagogy	19	Educational evaluation	49
Social psychology	17	Educational sociology	12	Library science	13	Rhetoric	45
Public administration	17	Elementary education	12	Educational sociology	14	Library science	45
Educational theory	15	Womens studies	12	Educational sociology	14	Educational sociology	43
Psychotherapy	15	Occupational psychology	10	Education	13	Education	43
Occupational psychology	14	Social psychology	10	Womens studies	13	Womens studies	42
Sociology	13	Educational tests & measurement	10	Elementary education	13	Religious education	41
Minority & ethnic groups	13	Social psychology	10	Religious education	11	Social psychology	40
Artificial intelligence	13	Cognitive therapy	13	Religious education	10	Social psychology	40

Table 26. Most Frequent Classifications and Dissertation Counts by time period

Table 27 and 28 detail the frequency of classification across department types. Here again Educational Software or Educational technology were the most frequently occurring classification for the 18 departments types except for Psychology, Nursing, Computer Science, Science, and Music. For each of these exceptions a classification specific to their department type (Educational Psychology, Nursing, Computer Science, Science Education, and Music Education) was most frequently chosen. Educational Technology is the second most frequently listed classification for each of the five above exceptions. This predominance of technology based classification across all years and department types indicates that graduate researchers characterized their research via technological *medium* more than classifications that focus on content, context, or specific discourses.

Higher Education (1553 classifications) was the second most frequent classification across the time periods. This remains true across the top three department types, Education, Ed Lead, and Ed Tech, as well as the None category. None indicated no department type listed yet is included for comparative purposes. These four types together account for 93% of the total classification across all years. The Higher Education classification droped to the third most frequent for the EdPsych, Business, and Psychology department types. The classification remained in the top twenty across all other department types except for Science, Special Education, and Agriculture. Graduate researchers predominately characterized their work as pertaining to the university context. It also indicated that the educational *context* was an important way to classify research.

To further support that the distance education *context* was of importance to graduate researchers, note that context-based classifications including Community College, Continuing Education, and Secondary Education are all represented in the top twenty classifications across the three time periods. These classifications were also among the top classifications represented in the Education-based department types such as Education, Ed Lead, Ed Tech, Ed Psych, and Higher Ed. These context based classifications were present in non-education department type, like Business, but not with as high a frequency.

Instructional Design (249 classifications), Curricula (578), and Curriculum Development (238) were among the top ten classifications across the three time periods. High frequencies of these were seen in the top department types None, Education, Ed Lead. They were present in all but the Nursing, Computer Science, Organizational, Special Ed, and Agriculture department types. Design-based classifications were more prevalent in education oriented department types, yet were present in some of the non-education departments. The predominance of these classifications in the sample indicated that Distinct Education *design* is another way graduate researchers characterize the focus of their research.

Teaching (578 classifications) ranked in the top ten classification in 2000-2004 and 2005-2009. Yet the classification was dropped by the PDQT during the 2010-2014. During this later time period Pedagogy (50 classifications) appeared as a classification choice. These indicated that issues stemming from the *instructor* and instructional situations were a focus for many Distance Education dissertations.

School administration (412 classifications) and Higher Education Administration (90) were among the top classifications across the three time periods. Educational Leadership was one the department types with the highest number of educational management oriented classifications, such as School Administration, Higher Education Administration, and Educational Policy. They were also present across the education-oriented department types and in department types like Business, Organizational, and Communications. With lower frequency the presence of these *management* oriented classification showed that many dissertations focused on administration of Distance Education.

none	10063	education	1928	edlead	601	edtech	333	misc	177	edpsych	169	business	151	psych	100
0710: Educational software	2094	0710: Educational technology	429	0710: Educational technology	150	0710: Educational technology	105	0710: Educational technology	29	0710: Educational technology	39	0710: Educational technology	35	0525: Educational psychology	21
0745: Higher education	1160	0745: Higher education	208	0745: Higher education	68	0745: Higher education	30	0525: Educational psychology	15	0525: Educational psychology	32	0454: Management technology	21	0710: Educational technology	9
0516: Adult education	586	0449: Instructional Design	128	0449: Instructional Design	50	0525: Educational psychology	27	0745: Higher education	9	0745: Higher education	15	0745: Higher education	20	0622: Clinical psychology	8
0727: Curricula	532	0516: Adult education	108	0514: School administration	46	0516: Adult education	14	0516: Adult education	9	0489: Information	10	0489: Information	10	0633: Cognitive	7
0727: Teaching development	532	0727: Curriculum	87	0489: Information	27	0449: Educational leadership	13	0449: Educational leadership	5	0688: Business education	8	0688: Business	8	0451: Social	5
0710: Educational technology	531	0710: Educational software	83	0723: Information	22	0646: Web Studies	11	0646: Web Studies software	4	0516: Adult education	7	0516: Adult education	6	0622: Clinical	5
0516: Continuing education	479	0530: Teacher Education	73	0646: college education	22	0443: Educational evaluation	9	0447: Instructional	4	0530: Teacher	4	0451: Social psychology	5	0727: Curricula	3
0514: School administration	320	0275: Community college education	55	0710: Educational software	21	0621: Psychology development	9	0525: Educational Technology	3	0710: Educational software	4	0530: Teacher education	4	0646: Web Studies	3
0530: Teacher education	294	0449: Educational leadership	50	0727: Curriculum development	15	0489: Information Technology	8	0533: Secondary education	3	0275: Community college education	3	0723: Information	3	0621: Psychology Developmental	2
0275: Community colleges	255	0533: Secondary development	48	0530: Teacher education	15	0727: Curriculum	7	0703: Organizational	3	0723: Information	3	0646: Web	3	0275: Community college education	2
0525: Educational psychology	253	0525: Educational psychology	38	0525: Educational education	13	0530: Teacher	7	0516: Continuing education	3	0646: Web Studies	3	0703: Organizational behavior	3	0516: Adult	2
0984: Computer science	121	0727: Teaching	37	0525: Educational psychology	11	0275: Community education	7	0456: Pedagogy behavior	3	0516: Continuing education	3	0624: Occupational	3	0529: Special education	2
0727: Curriculum development	118	0280: Mathematics education	36	0984: Computer science	10	0984: Computer behavior	6	0603: Counselling	3	0446: Higher Education	3	0275: Community college education	2	0733: Gender studies	2
0533: Secondary education	112	0443: Educational evaluation	36	0646: Web Studies	10	0680: Health education	4	0632: Quantitative psychology	2	0624: Occupational	3	0727: Teaching	2	0384: Behavioral	2
0350: Health education	103	0651: Continuing education	33	0443: Educational evaluation	9	0514: School administration	4	0620: Developmental psychology	2	0727: Teaching	2	0516: Educational	2	0628: Individual & family studies	1
0569: Nursing	96	0714: Science education	24	0514: School administration	9	0399: Library science	4			0514: School administration	2	0733: Gender studies	2	
0723: Information Systems	93	0458: Education Policy	21	0456: Pedagogy	6	0518: Early childhood education	4			0533: Secondary education	2	0384: Behavioral	2	
0688: Business education	85	0446: Higher Education	21	0446: Higher Education	6	0727: Curriculum	3			0651: Continuing education	2			
0454: Management	84					0569: Nursing psychology	3							

Table 27. Classifications by Department Type Part 1

Table 28. Classifications by Department Type Part 2

nursing — 66
- 13 — 0569: Nursing
- 10 — 0710: Educational technology
- 9 — 0680: Health education
- 4 — 0745: Higher education
- 3 — 0633: Cognitive psychology
- 3 — 0525: Educational psychology
- 2 — 0489: Information Technology
- 2 — 0646: Web Studies
- 2 — 0443: Educational evaluation
- 1 — 0451: Social psychology
- 1 — 0622: Clinical psychology
- 1 — 0343: Asian American Studies
- 1 — 0347: Mental health
- 1 — 0566: Health sciences
- 1 — 0382: Physical therapy
- 1 — 0450: Middle School education

highed — 66
- 13 — 0710: Educational technology
- 5 — 0275: Community college education
- 5 — 0446: Higher Education
- 5 — 0800: Artificial intelligence
- 3 — 0516: Adult education
- 4 — 0449: Educational leadership
- 3 — 0449: Educational leadership
- 3 — 0443: Educational evaluation
- 2 — 0489: Information Technology
- 2 — 0275: Community colleges
- 2 — 0342: Asian Studies
- 2 — 0455: Multicultural Education
- 1 — 0447: Instructional Design
- 1 — 0458: Education Policy
- 1 — 0447: Instructional Design
- 1 — 0514: School administration

compsci — 63
- 21 — 0984: Computer science
- 13 — 0710: Educational technology
- 8 — 0516: Adult education
- 4 — 0463: Statistics
- 5 — 0544: Electrical engineering
- 4 — 0800: Artificial intelligence
- 3 — 0723: Information science
- 3 — 0449: Educational leadership
- 2 — 0489: Community college education
- 2 — 0317: Neurosciences
- 2 — 0771: Robotics
- 2 — 0340: Educational sociology
- 1 — 0515: Education
- 1 — 0405: Mathematics
- 1 — 0548: Mechanical engineering
- 1 — 0537: Engineering
- 1 — 0521: Industrial arts education
- 1 — 0452: Social work
- 1 — 0533: Secondary education
- 1 — 0456: Pedagogy

org — 46
- 11 — 0710: Educational technology
- 3 — 0454: Management
- 4 — 0516: Adult education
- 8 — 0745: Higher education
- 3 — 0703: Organizational behavior
- 2 — 0449: Educational leadership
- 3 — 0530: Teacher education
- 2 — 0275: Community college education
- 2 — 0688: Business
- 1 — 0747: Vocational education
- 2 — 0714: Science education
- 1 — 0514: School administration
- 1 — 0525: Educational psychology
- 1 — 0533: Secondary education
- 2 — 0747: Vocational education
- 1 — 0460: Speech therapy
- 1 — 0515: Education
- 1 — 0290: Linguistics

communication — 46
- 8 — 0710: Educational technology
- 6 — 0745: Higher education
- 3 — 0459: Communication
- 3 — 0456: Web
- 3 — 0447: Instructional Design
- 2 — 0558: Multimedia Communications
- 2 — 0279: Language
- 2 — 0681: Rhetoric
- 1 — 0714: Science education
- 1 — 0525: Educational Design
- 1 — 0569: Nursing
- 1 — 0529: Special education
- 1 — 0460: Speech therapy
- 1 — 0515: Education

language — 42
- 9 — 0714: Science education
- 4 — 0710: Educational technology
- 3 — 0633: Cognitive psychology
- 3 — 0530: Teacher education
- 3 — 0441: English as a Second Language
- 3 — 0456: Pedagogy
- 3 — 0530: Teacher education
- 2 — 0471: Horticulture
- 1 — 0049: Information Technology
- 1 — 0447: Instructional Design
- 1 — 0378: Dance
- 1 — 0525: Educational care
- 1 — 0747: Vocational education
- 1 — 0727: Curriculum development
- 1 — 0688: Business education

science — 22
- 4 — 0714: Science education
- 2 — 0710: Educational technology
- 3 — 0633: Cognitive psychology
- 3 — 0530: Teacher education
- 3 — 0558: Multimedia Communications
- 2 — 0605: Physics
- 2 — 0471: Horticulture
- 1 — 0646: Web Studies
- 1 — 0490: Organic chemistry
- 1 — 0769: Health care
- 1 — 0573: Public health
- 1 — 0984: Computer science
- 1 — 0317: Neurosciences
- 1 — 0778: Veterinary services
- 1 — 0410: Microbiology

speced — 22
- 6 — 0522: Music education
- 3 — 0710: Educational technology
- 2 — 0529: Special education
- 2 — 0530: Teacher education
- 1 — 0529: Special education
- 1 — 0651: Continuing education
- 1 — 0340: Educational sociology
- 1 — 0447: Instructional Design
- 1 — 0455: Multicultural Education
- 1 — 0384: Behavioral psychology
- 1 — 0622: Clinical psychology
- 1 — 0679: Language
- 1 — 0535: Reading instruction
- 1 — 0745: Higher education
- 1 — 0489: Instructional Design

music — 21
- 6 — 0522: Music education
- 4 — 0710: Educational technology
- 3 — 0529: Special education
- 2 — 0456: Pedagogy
- 1 — 0515: Education
- 1 — 0530: Teacher education
- 1 — 0413: Music
- 1 — 0447: Instructional Design
- 1 — 0745: Higher education
- 1 — 0533: Secondary education

ag — 17
- 5 — 0516: Adult education
- 6 — 0710: Educational technology
- 5 — 0517: Agricultural education
- 2 — 0503: Agricultural economics
- 1 — 0518: Preschool education
- 1 — 0443: Educational evaluation
- 1 — 0489: Information Technology
- 1 — 0727: Teaching
- 1 — 0727: Curricula

Finally, many of the top classifications across the three time periods were discourse specific and focused on areas common to their department type. Top classifications included Computer Science (154 classifications), Nursing (136), Mathematics Education (123), Business Education (119), Management (112), Health Education (111), and Science Education (103). There was some cross over, Management showed up in Business and Organizational department types for example. But most of these *discipline specific* classifications were embedded solely in the department type that housed their field of study. As distance education moves from a predominantly education focused area other disciplines are more interested in delivering instruction in their field at a distance. Research into distance education in these disciplines is emerging in the literature of these fields. Again, the results from sub-question A indicated that every year since 2007 (when the PQDT began listing department) distance education focused dissertations came from the disciplines of Nursing, Business, Psychology, and Computer Science. These areas are recommended as a starting point for future investigation into discipline-based distance education research.

Tables 27 and 28 show that each department type had a series of classifications that are discipline specific. Yet across department types and the three time periods distance education graduate researchers characterized the topics of focus by technological medium, educational context, issues of design, issues around instructor, and issues of management.

"What dissertation keywords received the highest frequency?" Sub-Question I

Keywords, like classifications, were self-selected by the doctoral student. Unlike classification they were not bound by a framework, rather the graduate researcher was able to characterize their research using their own language choices. The PQDT allowed for up to six keywords or short phases for indexing in the database. All dissertations in the study provided at least one keyword.

The most frequently chosen keywords and counts for all dissertations spanning each of the three time periods were detailed in Table 29. Distance Education-based keywords, such as distance education (436/395/264), distance learning (214/235/140), and online-learning (59/215/408) were among the most frequently chosen keywords across all three time periods. Other distance education based keywords, like online education, e-learning, web-based instruction also made the list. This was of no surprise as these terms were used to isolate distance education based dissertations during sample collection from the PQDT. It is interesting to note the terms Distance Education and Distance Learning ranked most frequently during 2000-2004 and 2005-2009. Distance Education was used seven times more than Online Learning. Yet over the three time periods Online Learning has steadily increased in use (59/215/408) and become the most frequently used of these keywords to characterize distance education by graduate researchers. Bozkurt et al.(2015) found keywords such as ELearning and Mobile Learning highly represented in journal articles form 2009-2013, yet these were not found among the most frequently utilized by doctoral students.

The most frequent keyword education (1038/1384/1353) is utilized more than 50% than any other keyword across all three time periods. Keywords more specific to

educational context listed among the top twenty most frequently included higher

education (48/77/60), community college (36, 62, 42) and professional development (32,

47, 56). This further confirms what was seen in analysis of classification, that

instructional context is commonly chosen by graduate researchers to characterize their

studies. Higher education was the most frequently ranked across the three time periods,

as it was shown also to be the case in classification analysis.

2000-2004	6325	2005-2009	9888	2010-2014	9494
education	1038	education	1384	education	1353
distance education	436	online	503	online learning	408
distance learning	214	distance education	395	online	339
online	206	distance learning	235	distance education	264
web-based	120	online learning	215	social sciences	142
social sciences	97	social sciences	124	distance learning	140
communication and the arts	76	e-learning	120	communication and the arts	138
faculty	71	learning	94	e-learning	125
learning	67	communication and the arts	93	online education	118
health and environmental sciences	61	health and environmental sciences	80	applied sciences	97
online learning	59	faculty	79	health and environmental sciences	85
applied sciences	54	higher education	77	psychology	68
psychology	50	applied sciences	66	instructional design	61
higher education	48	instructional design	65	higher education	60
internet	39	psychology	63	professional development	56
community college	36	community college	62	blended learning	48
professional	32	online education	57	community college	42

development					
e-learning	32	web-based	49	technology	42
instructional design	31	professional development	47	motivation	36
teaching	29	technology	45	faculty	32
language	24	motivation	40	retention	31
technology	23	retention	34	learning styles	30
literature and linguistics	23	internet	32	student satisfaction	28
achievement	22	social presence	30	community of inquiry	24
adult learners	21	blended learning	29	persistence	24
internet-based	20	adult learners	27	self-efficacy	23
university	18	language	27	social presence	23
motivation	17	community colleges	26	student success	22
computer-mediated communication	16	face-to-face	26	leadership	21
community colleges	16	leadership	26	language	20
retention	15	online teaching	26	literature and linguistics	19
persistence	15	student satisfaction	25	computer-mediated communication	19
online education	15	teaching	25	interaction	19
self-efficacy	15	achievement	24	adult learning	18
learning environment	14	instructional technology	22	community colleges	18
teachers	14	collaborative learning	22	collaboration	17
online instruction	13	high school	22	web-based learning	17
collaboration	12	learning styles	22	online instruction	17
face-to-face	12	literature and linguistics	22	instructional technology	17
student satisfaction	12	graduate education	21	learning	16
television	11	critical thinking	20	training	16
graduate students	11	persistence	20	educational technology	16
texas	11	interaction	20	online courses	16

interactive television	11	web-based learning	19	virtual learning	15
web-based instruction	11	online instruction	18	mathematics	15
undergraduate	11	faculty development	18	high school	15
college students	11	learning environment	17	community	15
attrition	10	computer-mediated communication	16	student engagement	15
administrators	10	educational technology	16	credit recovery	15
learning styles	10	undergraduate	16	sense of community	15

Table 29. Frequency of Dissertation Keywords over Time

Community college focused studies were also shown as a common context in both keyword and classification analysis. There was no PQDT classification for Professional Development, but here we see that dissertations addressed Professional Development along with Training (NA/NA/16) and was the only non-academic specific contexts to emerge. With less frequency graduate researchers also characterized their dissertations as focusing on the undergraduate (11/16/NA), graduate (11/21/NA), and high school educational contexts (NA/22/15).

Instructional Design (31, 65, 66) is a keyword that was among the top twenty most frequent across all three time periods and was more frequently chosen in the later time periods. This supported what was shown in classification analysis, graduate researchers commonly choose keywords to characterize their research as addressing instructional design issues. Yet even though authors had the option to use more specific language in keyword choice, there are few keywords that address instructional design issues with more specificity. Some that do include blended learning (NA/21/48), collaboration (12/NA/17), and learning style (10,22,30).

Classification analysis showed that graduate researchers focused on issues of Administration and Instructor issues. Administration based keywords included administrators (10, NA, NA) and leadership (NA, 26, 21). Instructor based keywords included faculty (71, 79, 32) and teaching (29,5,NA). But as with instructional design based keywords, specificity of focus within areas of administration or instructor is not found in keywords. This indicated that while keywords may point towards a topic area, granularity within those areas is not often represented by graduate researcher keyword choices.

Discourse based keywords were frequently chosen across all three time periods. Social Sciences and Communications (97, 124, 142) and Communications and the Arts (76, 93, 138) were the first and second among these across all three time periods. Applied Sciences (54, NA, 97), Psychology (50, 63,68), and Health and Environmental Science (NA, 80, 85) also all increase over the three time periods. Similar to classification analysis, discourse specific keywords identified the area of study. Similar to other keyword areas found here, specificity beyond identifying the area of study was not present.

Finally, an area of focus that was not present in classification analysis but emerged in keywords was characteristics of learning. Among these achievement (22, 24, NA), motivation (17, 40, 36), retention (15, 34, 31), social presence (NA, 30, 23), student success (NA, NA, 22), student satisfaction (NA, 25, 28), self-efficacy (15, NA, 23), and persistence (15, 20, 24) all emerged as areas of the research focus.

Keyword analysis provided a view into graduate research on distance education that pointed at areas of focus yet did not enable much specificity of aspects within the

area of focus rather pointed at broader topics. Unlike the PQDT Classifications, which

are a pre-determined framework that researchers must fit within to categorize their

studies, keywords allows researchers to use any phrase to characterize their work. Yet

similar to what is shown with classifications the most frequently chosen keywords

characterized focus on educational context, design, instructor, management, and specific

disciplines (like nursing, business, psychology, or computer science). Classification

analysis and keyword analysis indicated that these were the predominant topic domains

for graduate dissertations in distance education.

"What are the relationships among the topics identified in dissertation abstracts using

co-occurrence social network analysis" Sub-Question J

Ritzhaupt et al., 2010 demonstrated a process where keywords and keywords

phrases were pulled out of research article abstracts. The authors noted that valuable

information regarding the topic of a research study may be housed in these keywords.

The current study utilized abstracts of dissertations to gather *emergent keywords* for

analysis. The process to gather the emergent keyword sample for the study included

using the NLTK natural language processing python toolset to chunk text into nouns and

noun phrases. Then words were stemmed to their root. This was useful for particular

instances, for example so that singular and plural form of a word were not counted as two

different keywords. Stop words such as "the" and "and" were removed. Finally the

distance education search terms used to identify dissertations for the sample and

keywords that are specific to methodology were removed. As methods were often not

distinguishable in abstracts, the removal of these words lead to a higher degree of

emergent keywords words pertaining to research topics. See appendix 7 for the full list of removed words.

Frequency analysis of dissertation Classification and Keywords enabled the identification of areas of focus for graduate researchers. Yet as seen in the previous section specificity in the topic areas was not represented. Also the structural relationships between topics could not be grasped by frequency analysis alone. This study employed network analysis to overcome this issue.   First frequently occurring pairs within the Emergent Keyword Co-occurrence Network (EKCN) were analyzed. Then visualization and network centrality measures were used to investigate EKCN features.

During the first time period, 2000-2004, a total of 2,081 nodes and 545,474 ties were identified in the EKCN. Among them, 200 ties with 69 nodes appeared 250 times or more. During the second time period 2005-2009, a total of 2,458 nodes and 688,662 ties were identified. Among them, 244 ties with 92 nodes appeared 250 times or more. During the third time period 2010-2014, a total of 2,487 nodes and 681,908 ties were identified. Among them, 254 ties with 100 nodes appeared 250 times or more. Table 30 presented a brief description of the EKCN elements in each of the time periods.

	Number of total elements		Frequently Occurred Pairs	
Period	Nodes	Ties	Nodes	Ties
2000-2004	2,081	545,474	69	200 (N>250)
2005-2009	2,458	688,662	92	244 (N>250)
2010-2014	2,487	681,908	100	254 (N>250)

Table 30. Description of the EKCN Elements

Frequency Analysis of Co-Occurring Pairs

What was immediately apparent when investigating the co-occurring keyword

pairs was that "Student" became a key factor. Table 31 lists the top fifty most frequently

co-occurring pairs across the three time periods.

2000-2004	2005-2009	2010-2014
COURS : STUDENT - 2167	COURS : STUDENT - 2411	COURS : STUDENT - 2474
INSTRUCTOR : STUDENT - 2029	INSTRUCTOR : STUDENT - 2036	PARTICIP : STUDENT - 1739
INTERACT : STUDENT - 1441	PARTICIP : STUDENT - 1880	INSTRUCTOR : STUDENT - 1712
PERCEPT : STUDENT - 1299	PERCEPT : STUDENT - 1848	PERCEPT : STUDENT - 1663
PARTICIP : STUDENT - 1296	LEARN : STUDENT - 1340	STUDENT : TEACHER - 1516
FACULTI : STUDENT - 1133	COMMUN : STUDENT - 1123	STUDENT : TECHNOLOG - 1232
EXPERI : STUDENT - 1106	EXPERI : STUDENT - 1111	COMMUN : STUDENT - 1185
STUDENT : TECHNOLOG - 1072	RELATIONSHIP : STUDENT - 1049	EXPERI : STUDENT - 1178
LEARN : STUDENT - 923	PROGRAM : STUDENT - 975	INTERACT : STUDENT - 1158
STUDENT : TEACHER - 855	INTERACT : STUDENT - 968	LEARN : STUDENT - 1133
GROUP : STUDENT - 832	STUDENT : TEACHER - 925	FACULTI : STUDENT - 1036
COMMUN : STUDENT - 830	FACULTI : STUDENT - 911	RELATIONSHIP : STUDENT - 971
PROGRAM : STUDENT - 751	STUDENT : TECHNOLOG - 811	PROGRAM : STUDENT - 871
INSTRUCT : STUDENT - 736	GROUP : STUDENT - 792	GROUP : STUDENT - 859
SATISFACT : STUDENT - 728	NUMBER : STUDENT - 764	NEED : STUDENT - 690
RELATIONSHIP : STUDENT - 693	SATISFACT : STUDENT - 700	NUMBER : STUDENT - 626
STUDENT : TIME – 637	STUDENT : UNIVERS - 693	MOTIV : STUDENT - 619
NEED : STUDENT - 601	STUDENT : TIME - 675	EXPERI : PARTICIP - 603
QUESTION : STUDENT - 537	NEED : STUDENT - 668	SATISFACT : STUDENT - 597
COURS : INSTRUCTOR - 519	QUALITI : STUDENT - 615	STUDENT : TIME - 588
STUDENT : UNIVERS - 507	GENDER : STUDENT - 589	SCHOOL : STUDENT - 586
COURS : PARTICIP - 503	QUESTION : STUDENT - 586	QUALITI : STUDENT - 570
QUALITI : STUDENT - 496	AGE : STUDENT - 555	PARTICIP : PERCEPT - 559
STUDENT : TYPE - 486	KNOWLEDG : STUDENT - 545	KNOWLEDG : STUDENT - 527
CLASS : STUDENT - 477	EXPERI : PARTICIP - 532	STUDENT : WAY - 522
INSTRUCTOR : INTERACT - 465	CLASS : STUDENT - 530	STUDENT : UNDERSTAND - 520
ATTITUD : STUDENT - 458	LEARNER : STUDENT - 528	ONLIN CLASS : STUDENT - 516
EXPERI : FACULTI - 458	STRATEGI : STUDENT - 515	PERFORM : STUDENT - 509
ACCESS : STUDENT - 454	ACTIV : STUDENT - 511	STUDENT : SUCCESS - 508
ADMINISTR : FACULTI - 454	STUDENT : SUCCESS - 510	LEARNER : STUDENT - 507
LEARNER : STUDENT - 450	PARTICIP : PERCEPT - 496	ENGAG : STUDENT - 500
NUMBER : STUDENT - 450	COURS : INSTRUCTOR - 493	OPPORTUN : STUDENT - 491
INTERNET : STUDENT - 446	COURS : PARTICIP - 484	TEACHER : TECHNOLOG - 490
STUDENT : TEACH – 445	GROUP : PARTICIP - 479	ONLIN ENVIRON : STUDENT - 483
AGE : STUDENT - 429	ENVIRON : STUDENT - 476	LEARNER : PARTICIP - 479
RESPONS : STUDENT - 422	RESPONS : STUDENT - 472	COURS : PARTICIP - 473
STUDENT : SUCCESS - 420	ONLIN ENVIRON : STUDENT - 465	STUDENT : YEAR - 473
COURS : FACULTI - 407	STUDENT : TYPE - 465	
EXPERI : PARTICIP - 392	PROBLEM : STUDENT - 460	
PROBLEM : STUDENT - 389	LEARNER : PARTICIP - 457	
FACULTI : PARTICIP - 383		

SEMEST : STUDENT - 380 FACULTI : PERCEPT - 376 LEARN STYLE : STUDENT - 373 GENDER : STUDENT - 366 COLLEG : STUDENT - 365 STUDENT : UNDERSTAND - 364 FACULTI : TECHNOLOG - 358 INSTRUCTOR : PARTICIP - 356 PARTICIP : TEACHER - 355	STUDENT : UNDERSTAND - 456 STUDENT : STYLE - 447 COURS : PERCEPT - 443 LITERATUR : STUDENT - 440 PERFORM : STUDENT - 439 ADMINISTR : STUDENT - 424 INSTRUCT : STUDENT - 423 ONLIN CLASS : STUDENT - 423 ABIL : STUDENT - 421 SEMEST : STUDENT – 418	INSTRUCT : STUDENT - 461 ENVIRON : STUDENT - 460 CLASS : STUDENT - 459 STUDENT : TYPE - 458 STUDENT : UNIVERS - 447 AREA : STUDENT - 441 LITERATUR : STUDENT - 430 PERCEPT : TEACHER - 421 DESIGN : STUDENT - 418 HIGH SCHOOL : STUDENT - 413 SKILL : STUDENT - 411 STUDENT : STUDENT SATISFACT - 401 GRADE : STUDENT - 400

Table 31. Frequently Occurring Emergent Keyword Co-Occurrence Pairs By Time Period

Student was one of the two co-occurring emergent keywords in 35 of the 50 pairs

or 70% during 2000-2004, 42 of 50 pairs or 84% during 2005-2009, and 44 of 50 pairs or

88% during 2010 to 2014. Student did not appear in keyword analysis or classification

analysis. Learner or student as focus of learning did begin to emerge as learning

characteristics in keyword analysis. But on analysis of this set of co-occurring pairs, it

became quite clear that locus of key concepts and topics held within dissertation abstracts

was highly related to the keyword Student. Further investigation of topics related to

Student will follow.

Other topics of focus that co-occur during all of time periods included the

keyword instructor and its relationship with student, course, administration, technology,

and interaction.

Visualization of the Emergent Keyword Co-Occurrence Network

In spite of the fact that we could recognize the central concepts through frequency

analysis, overwhelmingly student focused in this instance, it was hard to comprehend the

relationships among those pairs. Are there some pairs grouped together? Were there

many isolated pairs in the network? To answer these kinds of questions visualization of

the network was essential.   As addressed in chapter 3 the EKCNs were visualized using

multidimensional scaling layouts through UCINET's visualization app NETDRAW

(Borgatti, & et al., 2002). In these visual presentations some keywords were placed

closely, which implied that those keywords co-occurred more often than the words placed

further away. Furthermore, keywords were placed at the center of the diagram if they

were highly connected to most of the other keywords while some keywords were placed

at the periphery of the diagram as they were connected to a limited number of keywords

in the network. Figures 8, 9, 10 show the Emergent Keyword Co-Occurrence Network of

pairs that appeared at least 250 times across dissertation abstracts between 2000-2004,

2005-2009, and 2010-2014 respectively.

Figure 8. Emergent Keyword Co-Occurrence Network of pairs that appeared at least 250 times across dissertation abstracts between 2000-2004

Figures 9. Emergent Keyword Co-Occurrence Network of pairs that appeared at least 250 times across dissertation abstracts between 2005-2009
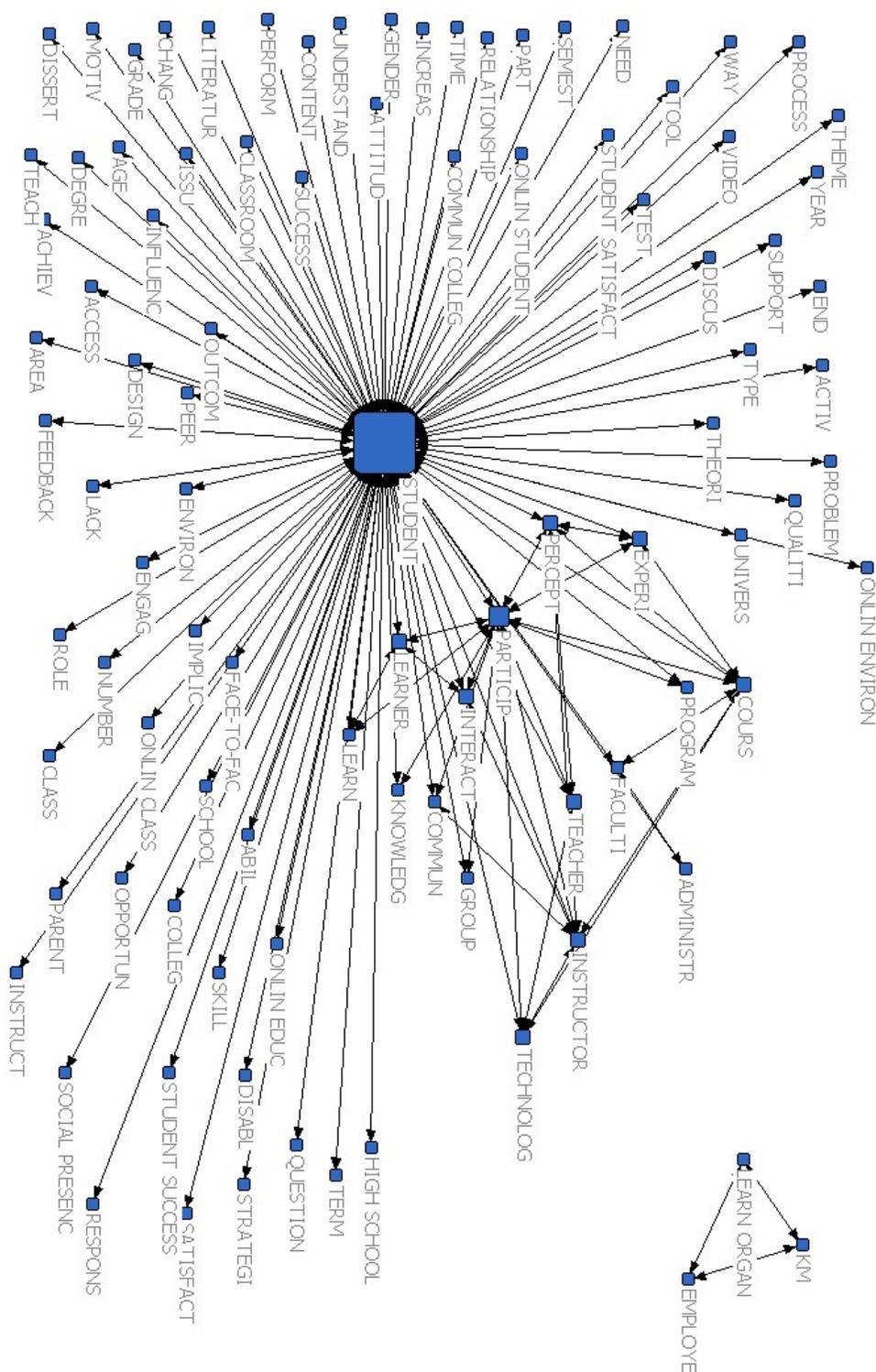
Figures 10. Emergent Keyword Co-Occurrence Network of pairs that appeared at least 250 times across dissertation abstracts between 2010-2014

The diagrams were all very similar across the three time periods. They were not broken into a number of cliques (or clusters of keywords). Rather the Emergent Keyword networks were all centralized around a small number of central topics. The most predominant feature, as noted in the frequency analysis of co-occurrence pairs, was that Student is the highly centralized node. A smaller cluster around Learner was also present in all three network visualizations. This strongly indicated that the abstracts of distance education dissertations across the time periods largely focused on student or learner issues.

Participation and Interaction were centrally related to a number of other nodes within the network; including Student, Instructor, Groups, and Learning. Interestingly both Perception and Experience were frequently connected to Participation, as well as Course and Instruction. In the later two periods Groups, Community, and Relationships also became related to Perception and Experience. This indicated issues around communicating with other learners and instructors were related to perceptions and experience of courses. These were all highly related to the Student node. Occurring at the latter period, this may also mean that distance education is now moving more toward designing collaboration among learners with the development of collaboration tools and learning management systems.

Contextual topics like Course and Program were related to Students, Administrators, and Instructors. Administration, Faculty and Perception all shared relationship ties. Course, Instructor, and Student were all related to Technology.

During the 2005-2009 time period a relationship between Student, Gender, and Age was shown. In the other time periods relationships between Student and Age and Student and Gender were present but not connected.

Also in the last time period a sub-grouping at the top right corner was seen and not related to the greater network. This small cluster of topics included Learning Organization, Knowledge Management, and Employee. This indicated that during the later time period a number of studies focused more on the organizational contexts that included professional development, knowledge management and training. It also indicated that the concepts brought forward in these abstracts did not frequently utilize language related to the academic context, such as Student or Instructor, enabling this cluster to remain separate.

Although the visualized networks enable us to see the structures of relationships intuitively, it is hard to uncover the unique features or patterns if the network has a large number of nodes and links. Network indices were calculated to identify the hidden patterns from a whole network by measuring the relational attributes quantitatively. In the next section, the EKCNs were analyzed with network indicators to investigate the undisclosed structural characteristics by visualization processes.

Network Measurements: Centrality, Betweeness, and Closeness

Centrality scores across all three show that Student and Participant have the highest number of direct linkages across all three time periods. This confirmed that dissertations abstracts place Student as the central topic covered across the three time periods. Also Participant was often used to refer to the research subject, so it may imply that in many of the abstracts, students were the research participants. Keywords with high

centrality fell into the general areas of focus previously identified during classification and keyword analysis. These include instructor focus (instructor and teacher), educational context (course and program), issues of management (administrator), design issues (group and community), and student oriented learning characteristics (interaction). Tables 32, 33, and 34 show the top twenty keywords with the highest centrality scores across the three time periods.

Degree Centrality		Betweeness Centrality		Closeness Centrality	
STUDENT	68	ABIL	135	STUDENT	2205.233154
PARTICIP	14	ACCESS	135	PARTICIP	26.73333168
COURS	9	ACHIEV	135	FACULTI	5.033333302
FACULTI	8	AGE	135	COURS	4.066666603
INSTRUCTOR	7	ATTITUD	135	INSTRUCTOR	1.950000048
TECHNOLOG	7	CHARACTERIST	135	TECHNOLOG	1.783333302
INTERACT	6	CLASS	135	INTERACT	1
PERCEPT	5	CLASSROOM	135	PERCEPT	0.200000003
EXPERI	4	COLLEG	135	ABIL	0
GROUP	3	COMPUT	135	ACCESS	0
LEARNER	3	CONTENT	135	ACHIEV	0
PROGRAM	3	DEGRE	135	AGE	0
TEACHER	3	DESIGN	135	ATTITUD	0
ADMINISTR	2	DISABL	135	CHARACTERIST	0
COMMUN	2	DISCUS	135	CLASS	0

LEARN	2	ENVIRON	135	CLASSROOM	0
TIME	2	FEEDBACK	135	COLLEG	0
ABIL	1	GENDER	135	COMPUT	0
ACCESS	1	INSTRUCT	135	CONTENT	0
ACHIEV	1	INTERNET	135	DEGRE	0

Table 32 Top Twenty Keywords with the highest Centrality Scores in the Emergent
Keyword Co-occurrence Network 2000-2004

Degree Centrality		Betweeness Centrality		Closeness Centrality	
STUDENT	90	FACULTI	265	STUDENT	3992.566406
PARTICIP	17	MEMBER		FACULTI	92.5
FACULTI	7	ABIL	182	PARTICIP	55.56666183
INSTRUCTOR	7	ACCESS	182	PERCEPT	2.983333111
PERCEPT	7	ACTIV	182	INSTRUCTOR	2
COURS	6	AREA	182	COURS	1.316666722
INTERACT	5	ATTITUD	182	INTERACT	0.666666687
COMMUN	4	BENEFIT	182	TECHNOLOG	0.400000006
TECHNOLOG	4	CASE STUDI	182	FACULTI	0
LEARNER	3	CHALLENG	182	MEMBER	
RELATIONSHIP	3	CHANG	182	ABIL	0
ADMINISTR	2	CLASS	182	ACCESS	0
AGE	2	CLASSROOM	182	ACTIV	0
EXPERI	2	COLLEG	182	AREA	0

GENDER	2	COMMUN	182	ATTITUD	0
GROUP	2	COLLEG		BENEFIT	0
KNOWLEDG	2	CONTENT	182	CASE STUDI	0
LEARN	2	CONTROL	182	CHALLENG	0
PROGRAM	2	DESIGN	182	CHANG	0
TEACHER	2	DISCUS	182	CLASS	0
		ENVIRON	182	CLASSROOM	0
		FLEXIBL	182		

Table 33. Top Twenty Keywords with the highest Centrality Scores in the Emergent
Keyword Co-occurrence Network 2005-2009

Degree Centrality		Betweeness Centrality		Closeness Centrality	
STUDENT	95	EMPLOYE	390	STUDENT	4492.617188
PARTICIP	14	KM	390	THEME	95
COURS	7	LEARN	390	PARTICIP	29.61666489
INSTRUCTOR	7	ORGAN	297	COURS	3.283333302
PERCEPT	6	UNIT	204	INSTRUCTOR	2.533333302
TECHNOLOG	5	ABIL	204	PERCEPT	1.083333254
EXPERI	4	ACCESS	204	FACULTI	0.5
INTERACT	4	ACHIEV	204	TECHNOLOG	0.5
LEARNER	4	ACTIV	204	INTERACT	0.333333343
TEACHER	4	AGE	204	LEARNER	0.333333343
COMMUN	3	AREA	204	TEACHER	0.200000003

FACULTI	3	ATTITUD	204	EMPLOYE	0
LEARN	3	CHANG	204	KM	0
ADMINISTR	2	CLASS	204	LEARN ORGAN	0
EMPLOYE	2	CLASSROOM	204	UNIT	0
GROUP	2	COLLEG	204	ABIL	0
KM	2	COMMUN		ACCESS	0
KNOWLEDG	2	COLLEG	204	ACHIEV	0
LEARN	2	CONTENT	204	ACTIV	0
ORGAN	2	CONTROL		AGE	0
PROGRAM		GROUP	204		
		DEGRE	204		
		DESIGN			

Table 34. Top Twenty Keywords with the highest Centrality Scores in the Emergent Keyword Co-occurrence Network 2010-2014

Since the majority of the nodes in the network were connected to Student. The shortest path between any other two nodes in a network was through the Student node.

The Professional development sub-group achieved a high level of betweeness among the other nodes in the sub-group. These high scores related to the subgroup remaining separate from the other connected keywords in the later period network.

Again, since Student had the most number of ties in the network the shortest path to any other node in the network was via Student. This was evidenced by its very large closeness measurement. Participant, Faculty, Course, Instructor, Interaction, and

Technology were all positioned in the network to have short paths to get to any other

node. Yet the most direct route was through Student.

Student Oriented Ties

Dissertation topic areas that emerged through keyword and classification analysis

were technological medium, educational context, issues of design, instructor issues,

issues of management, and learning characteristics. All of these areas are observed in the

emergent keywords even though they are clearly overshadowed by student-oriented

focus. It was anticipated that a level of specificity among the keywords in these topic

areas would emerge, but they did not.

The Student node was highly connected with topics that were shown to have a

high level of centrality to the network. Student was connected with nodes that indicated

focus on the education context, like Course, Program, University, Class, Semester,

College, and High School. Student was highly connected with instructor based emergent

keywords and shared strong ties with Instructor, Faculty, and Teacher. Management

oriented topics were represented by the tie with the Administrator node. Technology and

Internet nodes shared ties with Student and were about as specific as technological

medium nodes got within the network.

Student characteristics and attitudes were highly visible in the network.

Perception, Satisfactions, Attitude, Success, Motivation, and Engagement were

relationships that characterized the student experience. Interaction was shown to be

highly central to the network across all periods. Student was connected to other nodes

that also point towards this focus on interaction including Group, Community, and

Relationship**.** This connection indicated a focus on learner to learner interactions and

participation within learning communities. Student ties that connect Instructor –

Interaction – Student were also easily observed in the network visualizations. Finally ties

with Problems, Needs, and Access also indicated that improvement was a focus in

relation to students.

Emergent Keyword Co-Occurrence without the Student Node

In an effort to further understand the relationships between topics in the emergent

keyword networks, the nodes Student and Participant were removed. Further, the cut off

threshold for frequently co-occurring pairs was lowered to 150. The results are below.

	Number of total elements		Frequently Occurred Pairs without Student or Participant Node	
Period	Nodes	Ties	Nodes	Ties
2000-2004	2,081	545,474	31	134 (N>150)
2005-2009	2,458	688,662	33	158 (N>150)
2010-2014	2,487	681,908	42	182 (N>150)

Table 35. Description of the EKCN Elements without Student or Participant Nodes

The EKCN without the Student or Participant nodes contained 31, 33, and 42

nodes across the three time periods. The network also contained 134, 158, 182 ties

respectively. The lowered cut off threshold of 150 ties minimum enabled the opportunity

of nodes not present in the previous networks to be included.

Tables 36, 37, and 38 detail the centrality measurements for the three EKWN's.

Results were similar across the networks. Course was among the most central nodes

during each time period. Faculty, Teacher, and Instructor were also among the most

central nodes. Experience and Perception are highly centralized across all three time

periods. Interaction is also shown to be a central topic in each of the networks.

Community was present across all three networks, but moves into the top five most

central keywords during the latter time period.

Degree Centrality		Betweeness Centrality		Closeness Centrality	
COURS	17	CLIENT	142	FACULTI	151.9999847
FACULTI	15	COUNSELOR	142	COURS	128.3095245
INTERACT	10	REACTION	142	INTERACT	37.55237961
EXPERI	9	PROFESSION	94	TEACHER	34.86190796
INSTRUCTOR	9	DEVELOP		PERCEPT	24.68571472
PERCEPT	9	TYPE	90	TECHNOLOG	20.80476189
TECHNOLOG	8	COLLEG	82	EXPERI	18.95714569
TEACHER	7	CONCERN	82	INSTRUCTOR	6.447619438
LEARN	6	DISTANC	82	TEACH	3.25
LEARNER	6	EDUC		LEARN	1.452380896
GROUP	5	TECHNOLOG	82	INSTRUCT	1.299999952
INSTRUCT	3	FACULTI		LEARNER	1.035714269
RELATIONSHIP	3	MEMBER	82	GROUP	0.342857152
ADMINISTR	2	NEED	78	CLIENT	0
CLIENT	2	INTERNET	78	COUNSELOR	0
COMMUN	2	NUMBER	76	REACTION	0
COUNSELOR	2	ADMINISTR	76	PROFESSION	0
PROGRAM	2	COMMUN	76	DEVELOP	
REACTION	2	UNIVERS	75	TYPE	0

TEACH	2	TEACH	74	COLLEG	0
TIME	2	RELATIONSHIP	73	CONCERN	0
UNIVERS	2	GROUP	72	DISTANC EDUC	0
COLLEG	1	LEARNER	70	TECHNOLOG	
CONCERN	1	LEARN	69	FACULTI	0
DISTANC	1	PROGRAM	69	MEMBER	0
EDUC		TIME	68	NEED	0
TECHNOLOG	1	TEACHER	68	INTERNET	0
FACULTI		INSTRUCT	67	NUMBER	0
MEMBER	1	INSTRUCTOR	64	ADMINISTR	0
INTERNET	1	INTERACT	61	COMMUN	0
NEED	1	EXPERI	61	UNIVERS	0
NUMBER	1	PERCEPT	61	RELATIONSHIP	0
PROFESSION		TECHNOLOG	56	PROGRAM	0
DEVELOP	1	FACULTI	52	TIME	
TYPE		COURS			

Table 36. Centrality Measurements of EKCN Nodes without Student or Participant Nodes 2000-2004

Degree Centrality		Betweeness Centrality		Closeness Centrality	
PERCEPT	18	LEARN STYLE	187	PERCEPT	152.2761841
COURS	14	PROCRASTIN	187	FACULTI	87.06666565
FACULTI	11	REFLECT	187	COURS	67.41667175

INSTRUCTOR	10	SCORE	187	INTERACT	30.71666527
EXPERI	9	STYLE	114	TEACHER	28.65000153
LEARN	9	INQUIRI	110	SATISFACT	27
LEARNER	9	FORMAT	103	TECHNOLOG	19.31904793
TECHNOLOG	9	TYPE	99	EXPERI	13.95476055
COMMUN	8	ADMINISTR	99	LEARN	11.37142849
INTERACT	8	BARRIER	99	INSTRUCTOR	11.01666641
PROGRAM	6	NEED	94	COMMUN	8.571428299
TEACHER	6	NUMBER	90	LEARNER	6.890476227
RELATIONSHIP	5	SOCIAL		TEACH	6.700000286
TEACH	5	PRESENC	89	PROGRAM	6.43333292
SATISFACT	4	AGE	89	FACULTI	1.616666675
UNIVERS	4	GENDER	89	MEMBER	
FACULTI	3	GROUP	88	LEARN STYLE	0
MEMBER		QUALITI	87	PROCRASTIN	0
AGE	2	TEACHER	85	REFLECT	0
GENDER	2	TEACH	83	SCORE	0
GROUP	2	SATISFACT	83	STYLE	0
QUALITI	2	UNIVERS	81	INQUIRI	0
ADMINISTR	1	RELATIONSHIP	81	FORMAT	0
BARRIER	1	FACULTI		TYPE	0
FORMAT	1	MEMBER	77	ADMINISTR	0
INQUIRI	1	PROGRAM	76	BARRIER	0

LEARN STYLE	1	LEARNER	76	NEED	0
NEED	1	COMMUN	76	NUMBER	0
NUMBER	1	INTERACT	75	SOCIAL	0
PROCRASTIN	1	INSTRUCTOR	75	PRESENC	
REFLECT	1	LEARN	73	AGE	0
SCORE	1	EXPERI	73	GENDER	0
SOCIAL	1	TECHNOLOG	72	GROUP	0
PRESENC		FACULTI	67	QUALITI	0
STYLE	1	COURS	63	UNIVERS	0
TYPE	1	PERCEPT		RELATIONSHIP	0

Table 37. Centrality Measurements of EKCN Nodes without Student or Participant Nodes 2005-2009

Degree Centrality		Betweeness Centrality		Closeness Centrality	
COURS	16	DEGRE	241	PERCEPT	233.5333252
PERCEPT	16	STUDENT	241	TEACHER	214.4916687
TEACHER	16	SATISFACT		COURS	118.9083557
EXPERI	11	TOOL	241	EXPERI	75.2666626
COMMUN	10	TRAINE	241	FACULTI	73.10832214
LEARN	10	TUTOR	241	EMPLOYE	65.5
FACULTI	9	WEB	241	THEME	34
INSTRUCTOR	9	UNIT	164	KM	31.5
INTERACT	9	COMPANI	151	TECHNOLOG	17.31666565

LEARNER	9	ORGAN	150	PROGRAM	9.866667747
TECHNOLOG	9	ORGANIZ	150	LEARN	9.275000572
EMPLOYE	6	CLIMAT		COMMUN	9.208333969
PROGRAM	6	COLLEG	141	INSTRUCTOR	6.824999809
KM	5	FACULTI	141	LEARNER	6.824999809
ADMINISTR	3	MEMBER		INTERACT	3.374999762
LEARN ORGAN	3	THEME	130	DEGRE	0
RELATIONSHIP	3	CLASSROOM	130	STUDENT	0
SATISFACT	3	NEED	130	SATISFACT	0
GROUP	2	PROFESSION	130	TOOL	0
KNOWLEDG	2	DEVELOP		TRAINE	0
ORGAN	2	ROLE	130	TUTOR	0
ORGANIZ	2	SCHOOL	130	WEB	0
CLIMAT	2	TEACH	130	UNIT	0
THEME	1	DESIGN	128	COMPANI	0
CLASSROOM	1	TIME	128	ORGAN	0
COLLEG	1	SATISFACT	126	ORGANIZ	0
COMPANI	1	QUALITI	125	CLIMAT	
DEGRE	1	KNOWLEDG	124	COLLEG	0
DESIGN	1	ADMINISTR	120	FACULTI	0
FACULTI	1	LEARN ORGAN	120	MEMBER	
MEMBER		RELATIONSHIP	119	CLASSROOM	0
NEED	1	GROUP	119	NEED	0

PROFESSION	1	KM	118	PROFESSION	0
DEVELOP		EMPLOYE	117	DEVELOP	
QUALITI	1	TECHNOLOG	110	ROLE	0
ROLE	1	INSTRUCTOR	109	SCHOOL	0
SCHOOL	1	LEARNER	109	TEACH	0
STUDENT	1	FACULTI	107	DESIGN	0
SATISFACT		PROGRAM	104	TIME	0
TEACH	1	INTERACT	103	SATISFACT	0
TIME	1	COMMUN	102	QUALITI	0
TOOL	1	LEARN	100	KNOWLEDG	0
TRAINE	1	EXPERI	98	ADMINISTR	0
TUTOR	1	TEACHER	96	LEARN ORGAN	0
UNIT	1	COURS	94	RELATIONSHIP	0
WEB		PERCEPT	91	GROUP	

Table 38. Centrality Measurements of EKCN Nodes without Student or Participant Nodes 2010-2014

Figures 11, 12, and 13 detail the network visualizations across the three time periods. The Faculty, Instructor, and Teacher nodes had similar relationships across the networks. Faculty was related to University and Administration. Instructor was closer to other highly central nodes like Learners, Groups, and Technology. With the removal of the Student and Participant nodes, the Instructor/Teachers became a central actor in the networks. With the individual student relationship removed, the strong relationship

between the Instructor with Groups and Communities can be seen in each of the networks.

The Learn, Learner, Interaction, Relationship, and Community nodes are clustered closely in the networks. In this grouping Interact, Learner, and Group acted as hubs. Interaction was a highly central node that connected with Instructor, Technology, Learner, Group, Community, and Course among others. This along with its high centrality measurement indicated that Interaction was a popular research area in dissertations.
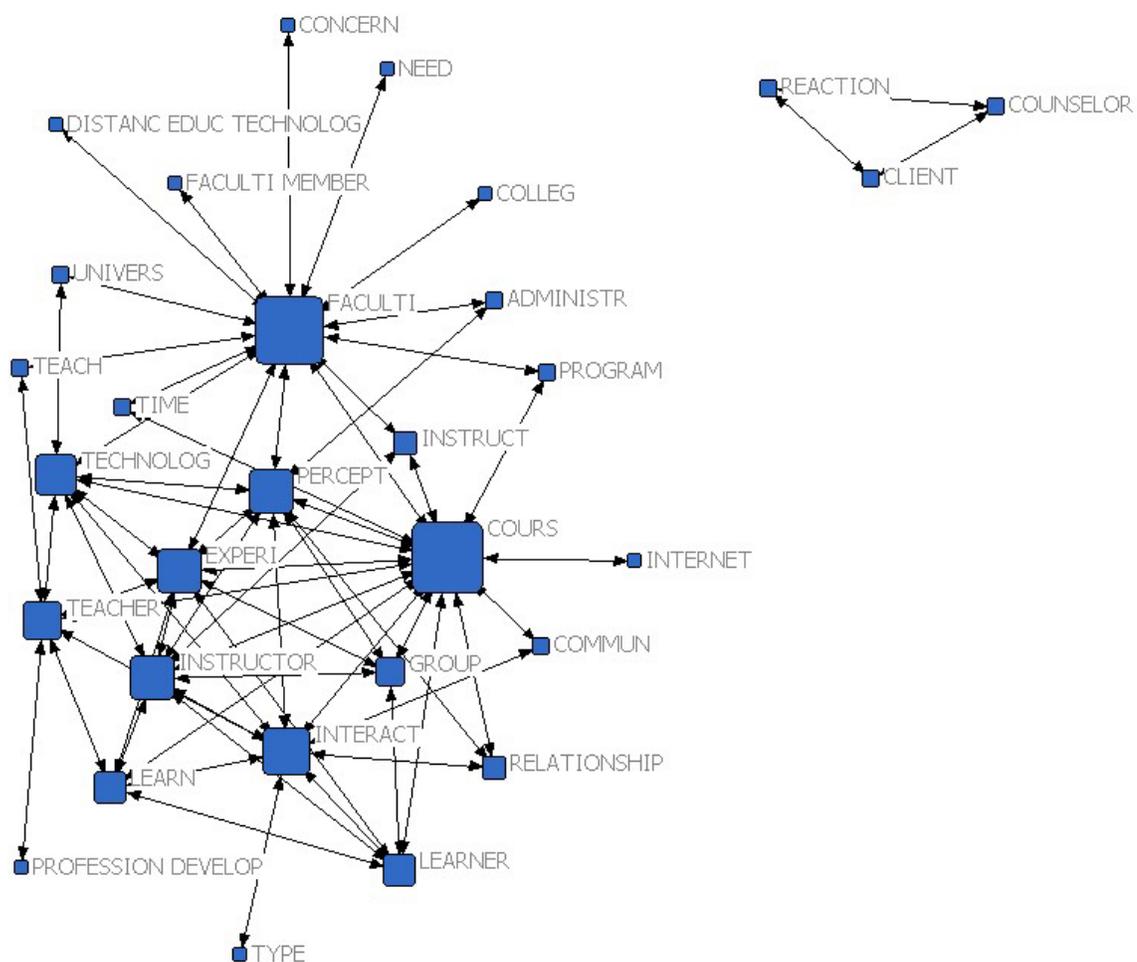
Figure 11. Emergent Keyword Co-Occurrence Network of pairs that appeared at least 100 times without Student or Participant Nodes across between 2000-2004

Figure 12. Emergent Keyword Co-Occurrence Network of pairs that appeared at least
100 times without Student or Participant Nodes across between 2005-2009
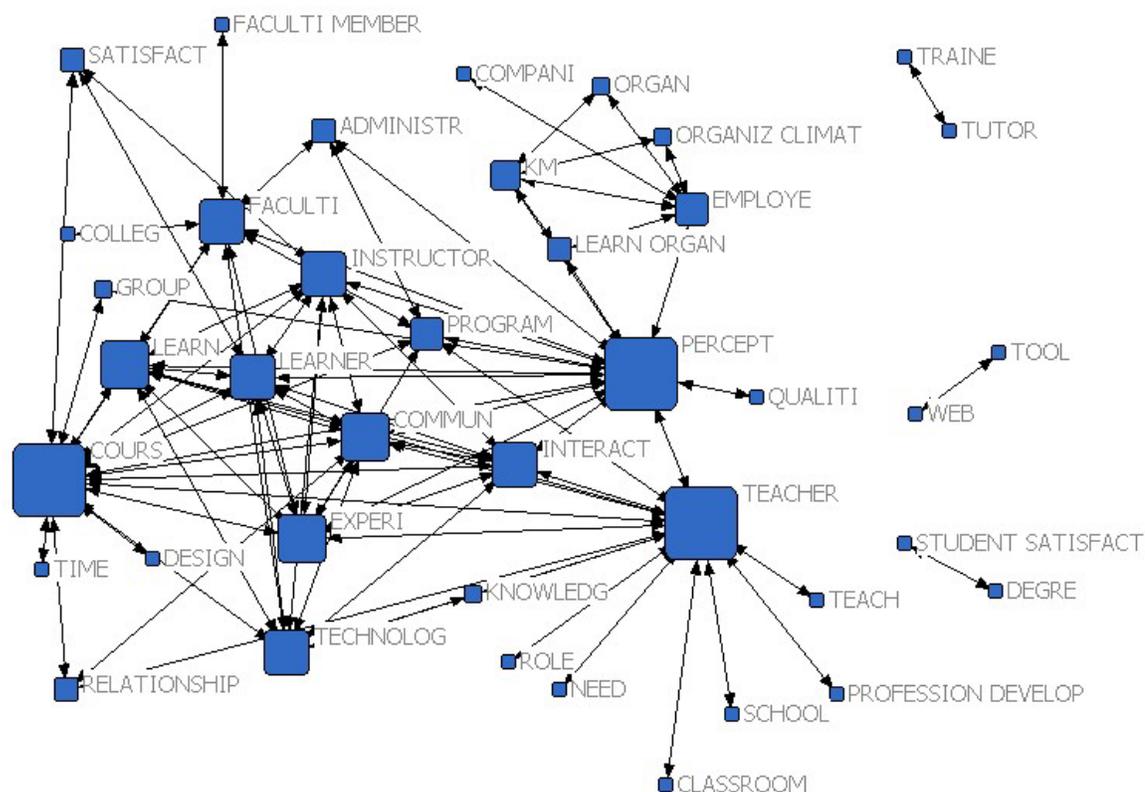
Figure 13. Emergent Keyword Co-Occurrence Network of pairs that appeared at least 100 times without Student or Participant Nodes across between 2010-2014

Perceptions and Experience were also highly central nodes. These were each highly related to Course, Learner, Technology, Groups, Communities, and Instructor. Perceptions and Experience were topics that may have indicated points of analysis in dissertations. Along with quality and Satisfaction, these nodes may indicate the measurement points doctoral students used in dissertations on distance education.

Results and Discussion

The general research questions asked, "What research topics can be identified in doctoral dissertation research on distance education published in North America in English between 2000-2014?" And how have the themes identified changed over time, specifically 200-2004, 2005, 2009, 2010-2014?

When graduate researchers submit their dissertations to the PQDT database there are three main kinds of information they are asked to attach to their study that characterizes topic and focus of the research. These three areas are Classification (chosen from the PQDT existing framework of subjects), Keywords (self selected words or phrases that indicate subject), and Abstracts (a paragraph length synopsis of the topics of the research). Classification and keywords analysis utilized a bibliometric analysis approach and measured the frequencies of topics as they occurred across the three time periods. Algorithmic language processing was used to pull emergent keywords and phrases out of dissertation abstracts. These key words were then analyzed using a Emergent Keyword Co-Occurrence Network approach that measured frequently co-occurring pairs and measured their centrality in relation to the rest of the network. These three analysis points looked at three different sets of topic indicators generated by graduate researchers. They offer a triangulated view into the topics distance education dissertations covered over the past fifteen years. The researcher analyzed the results of the three analyses and found that seven topics were pervasive across the three time periods.

Table 39 details the emergent framework of topics that graduate researchers addressed in dissertations on distance learning from 2000-2014.

Student
Instructor
Interaction
Administration and Management
Design
Educational Context
Technological Medium

Table 39. Topics Addressed in Doctoral Dissertation on Distance Education 2000-2014

Student is the only topic in the list that did not appear in classification or keyword analysis, yet characteristics of learning did emerge and point toward student as a focus of research. Student was the dominant node in the co-occurrence network, indicating that issues relating to learners were highly central in the characterization of research found in dissertations abstracts. Student was highly related to characteristics of learning in which the experiences, attitudes, and successes of distance learning students were the focus of dissertation research.

Issues specific to instructors, instructional design, and educational context all pull the focus of dissertation research towards the shaping and delivery of distance education. The instructor was positioned in the middle of this. We also see that administration and management are related to educational context and course functioning. This is of regular focus by the graduate researchers across the three time periods.

Interaction was not found in classification or keyword analysis. Yet it was found to be highly central to emergent keyword networks. Interaction between learners, instructors, groups, and communities were present across all three networks.

Technological medium was not expanded on very much, yet online technology was a central characterization point for these dissertations and should not be left out of the list of predominant topics in the sample.

How have these topics changed over time? Little change was observed over time. Triangulating between classifications, keywords, and emergent keywords we see indications that these seven topic areas were predominant across all three time periods, 2000-2004, 2005-2009, and 2010-2014. While subtle changes in bibliometric ordering and frequency as well as difference in centrality measurements were observed across time periods, the chief trend observed was that these research areas were all found across time.

This may by due to the fact that the topic areas in the emergent framework are quite broad. A level of specificity or granularity within the topic areas was not found. The researcher hypothesized that emergent keywords from abstracts would enable a specificity of topic beyond what was found. Instead the most frequently co-occurring emergent keywords pointed toward broad research domains, not specific topics within these domains. It may be concluded that this form of bibliometric and social network analysis of topics from dissertation database data enabled a broad picture of the topics covered in dissertation research. But it was limited by not being able to retrieve more specific sub-topics within each of the topic areas listed in the emergent framework found in Table 39.

Comparison to Zawacki-Richter et al.'s (2009) Framework

In 2009 Zawacki-Richter, 2009 performed an international Delphi study to develop the *Classification of Research Areas in Distance Education (CRADE)*. Based on analysis of 25 experts from 11 countries, three broad meta-levels of distance education research were derived (Table 2).

The seven broad dissertation topic areas that emerged in this study were Student, Instructor, Interaction, Design, Educational Context, Administration and Management, and Technological Medium. Most of these areas are represented in Zawacki-Richter CRADE framework (2009a; Zawacki-Richter & Anderson, 2014).

Interaction based nodes were found to connect learners, instructors, communities, and groups. This was represented in the CRADE research area: Interaction and communication in learning communities.

The node Student in the network analysis was found to be related to characteristics of learning in which the experiences, attitudes, and successes of distance learning students are the focus of dissertation research. This is found in the CRADE research area: Learner characteristics.

Instructor as topic of focus in this study may be indicated by the CRADE Area of Research: Professional development and faculty support. Yet Professional development was considered in the topic area Educational Context in this study. Also Instructor based topics were found to be linked to keywords like Pedagogy and Teaching. Indicating that instruction is the focus of this topic area, not support.

Instructional Design was found in both the topic areas that emerged from this study as well as in the CRADE framework. This was also true for Management and

Administration in this study and Management and organization in the CRADE framework. Technological Medium and Educational Technology are comparative between the two frameworks as well.

Finally Educational Context is characterized as location of distance learning. It involves audience and whether a study is academic or organizationally based. Educational context includes University, Higher Education, Community College, Graduate, Under Graduate, High School, and Professional Development. Each of these locate the distance learning design in a broadly different context. Many dissertations characterized their work with the Educational Context topic. There is some overlap with the CRADE research area Distance teaching systems and institutions.

The remaining eight CRADE research areas are not addressed in this study's topic framework. Bozkurt et al. (2015) found similar results. They reported that the majority of studies investigated in their study fell within only a few areas of the CRADE: Interaction and communication in learning communities, Learner characteristics, Instructional design, and Educational technology.

The CRADE research areas not found represented in the dissertation sample included Research methods in DE and knowledge transfer, Globalization of education and cross-cultural aspects, Innovation and change, Costs and benefits, Theories and models, Learner support services, Access, equity, and ethics, Quality assurance. This indicated that dissertations might not have focused research in these areas. Yet as discussed previously, specificity of topic areas was not found in this study due largely to data sample: keywords, classification, and abstracts. While these areas of focus may

have been present in dissertations, they did not emerge as topic areas in this study due to lack of keywords and classifications indicating these topics.

This indicates a weakness in this study approach. Since analysis depends on keywords and classifications a higher level of specificity in keywords and language used in abstracts is necessary for topics to emerge. While broad topic areas were shown to emerge more specific areas may not have been reported because of lack of keywords present in the sample. Conflicting language used to indicate the same topics in different studies may be unrepresented by this approach. Finally, pair co-occurrence of keywords may not be sufficient to indicate research topics.

Relationships Between Research Topics in Distance Education Dissertations

The study method was able to identify seven research topic areas frequently addressed in dissertations over the past fifteen years. There is a high level of similarity between the framework that emerged from this study and the CRADE framework. This helped to confirm the accuracy of the broad topic areas that were found using this method. Yet the CRADE framework was unable to illustrate the relationships between the research topic areas. The social network analysis approach used in the study allowed for investigation of relationships between topics. The relationships between nodes indicating each of the seven emergent topic areas found in this study were analyzed. Figure 14 details the relationships between topics.
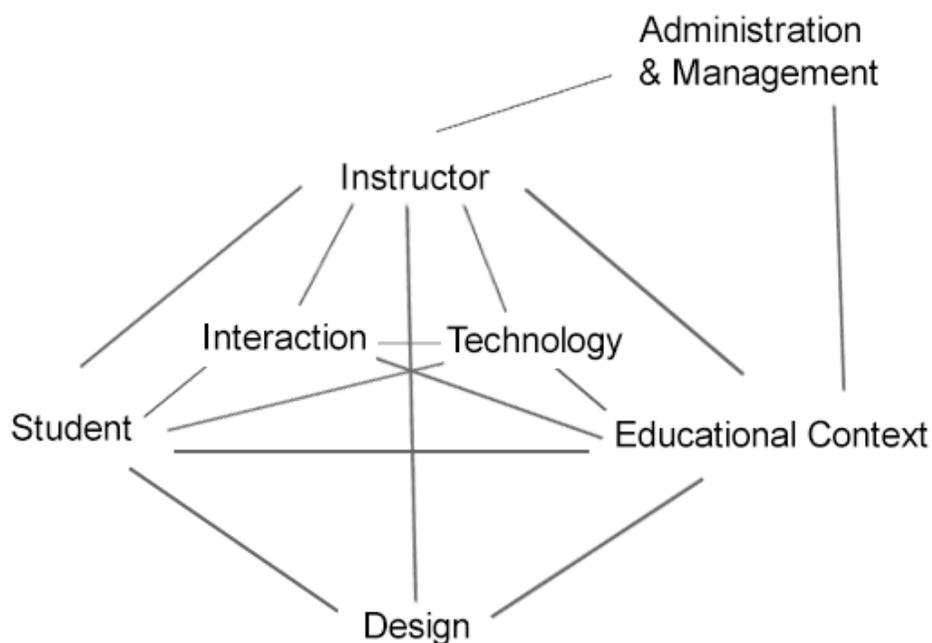
Figure 14. Relationships between the Topics Addressed in Doctoral Dissertation on Distance Education 2000-2014

The Instructor and Educational Context topics were related to all other research found topics in the dissertation sample. This does not mean that every dissertation that addressed instructor oriented topics also addressed all other topics. Rather combinations of research topics were observed and the instructor topic was found related to some other topics such as Student and Interaction. Access to the Administration and Management topic was the feature that enabled Instructor and Educational Context access to all topics in the framework. Conversely it was found that Administration and Management nodes in the social networks were predominantly tied to Instructor and Educational Context oriented nodes. It was found, for example, that studies that addressed Administration and Management did not also include research topics such as Design or Interaction.

The Student topic had access to five of the possible six other topics. It was found that Administration and Management was not directly related to the Student topic.

Rather Instructor or Educational Context topics were intermediary topics between Student and Administration.

Research that focused on design-oriented issues were found to be most highly related to Instructor, Student, and Educational Context. While instructional design was found highly represented in classification and keyword data, it was found to be secondary and not highly represented in emergent keyword networks. Particularly, design oriented nodes were not found in two of the three networks that did not include the student node. This indicated that studies addressing instructional design were more highly related to studies with a Student focus, than Instructor or Educational Context.

The remaining topic areas Interaction, and Technology each had access to four of the other topics. These topics were directly related to design and Administration and Management.

While not generalizable beyond the sample of dissertations found in this study, the bibliometric and social network analysis approach utilized enabled the emergence of seven general research topics. Further, a map detailing the relationships between dissertation research topics was forwarded.

Summary

This chapter detailed the findings of bibliometric sub-questions indicating characteristics of graduate researchers addressing distance education in their dissertations, as well as illustrating the reference base they work from. Keyword analysis, classification analysis, and Emergent Keyword Co-Occurrence Network analysis each provided a different look into the topics graduate researchers use to characterize the topics of their dissertations. From these three analyses a set of seven

broad topic areas were found : Student, Instructor, Interaction, Design, Educational

Context, Administration and Management, and Technological Medium. No change was

observed in these over the three five year time periods investigated in the study, which

indicated a trend that these broad topics continue to be of interest to researchers across

the years sampled. The next chapter will detail conclusions and implications for further

study.

CHAPTER V

SUMMARY, CONCLUSIONS, AND RECOMMENDATIONS

Overview

The primary purpose of this study was to investigate the research topics found in distance education dissertations from 2000-2014. This purpose was addressed through two modes of investigation, bibliometrics and social network analysis. Bibliometric analysis enabled a picture of the doctoral researcher sample to emerge. Frequency ranking detailed the universities and departments research came from as well as the most commonly investigated topics, and sources the researchers cited. SNA enabled investigation of the relationships between research topics that emerged from dissertation abstracts.

Research Questions

1. What research topics can be identified in doctoral dissertation research on distance education published in North America in English between 2000-2014?

2. How have the research topics changed over time, specifically from 2000-2004, 2005-2009, and 2010-2014?

3. How do the research topics compare to the research areas forwarded in Zawacki-Richter (2009a) CRADE framework?

Methodology

A sample of dissertation records, n=3954, that address distance education and range between the years 2000 – 2014 were gathered from the ProQuest Dissertations & Theses A&I (PQDT) database. Distance education related dissertations were

characterized as having the following keywords in the title, abstract, or subject fields of the database record: "distance education", "distance learning", "online learning", "electronic learning", "e-learning", "network learning", "distributed learning" or "web-based learning". The collected sample was exported from the PQDT database and imported into databases created specifically for this study. Custom algorithms and natural language processing techniques were used to organize and standardize data within the databases. Custom algorithms were created to analyze the databases and return answers to the specific study sub-questions. Bibliometric statistics were analyzed using a custom web-interface designed for the study. SNA was performed using Ucinet and networks were visualized using NetDraw.

Each of the 3954 dissertations listed university information and 45% (1375) of these included department information. From the data research sub-questions addressing where dissertation authors studied were analyzed. A small sub-sample, 338 dissertations, included references sections. From these reference sections 41,408 references were analyzed to address the citation based research sub-questions. Each of the dissertations in the sample included author selected classifications, as well as author generated keywords, and abstract. The resulting 13,921 classifications, 25,707 keywords, and 271,905 abstract based emergent keywords were analyzed using bibliometric and SNA approaches. This data was used to address the dissertation research topic sub-questions. To address change over time research sub-questions were investigated across three time periods 2000-2004, 2005-2009, and 2010-2014.

Study Technology

The current study contributed an approach to SNA that is highly automated both in the data gathering and data analysis phases of the investigation. The algorithms developed to import data out of the PQDT database demonstrated a technique where the heavy lifting of both data gathering, data formatting, and import into the study databases were mechanized. This acted to save hundreds of hours at the front end of the investigation. As compared with other studies, this approach enabled much larger sample sizes to be used. Future researchers may benefit from similar approaches, when they work with large data sets or big data.. The algorithms developed for this study were relatively rigid, meaning that they were specific to pulling data from the PQDT and formatting the data into the specific database structure used in this study. Further development of the study algorithms could make it so that a researcher may query across multiple databases of their own choosing for gathering records and to determine their own output database structure.

This approach may also benefit researchers across discourses. This is not a Distance Education specific technology and therefore the algorithmic supported SNA approach contributed in the current study could be used to investigate research topics and citation data across numerous diverse fields.

Study Technology Validity and Reliability

Test samples of 25 dissertation records and 10 citation reference sections were scraped, imported, and formatted into the database over five consecutive tests. Resulting data was identical across each of the tests. This indicated high reliability of the algorithms to complete import and formatting tasks.

Output from the tests was examined and it was found that data was imported, segmented, and correctly formatted across all records for each test. This continuity of results across the tests indicated that the software was reliable. The caveat to this reliability and validity was that these algorithms are currently designed to work only with data from the PQDT. Using the software with data from other scholarly databases would not produce reliable results. Therefore, future researchers need to modify this approach for other databases.

Social Network Analysis as a Research Method

Researchers using SNA are provided with a convenient representation and summary of relationships between investigated nodes. This study was able to utilize this approach to demonstrate relationship of research topics utilizing co-occurrence networks and relationships between reference section citations using co-citation networks.

Similar to cluster analysis, both objective and subjective roles are required for SNA. Romesburg (2004) notes that the only objective parts of the research process are the data processing methods. Interpretation of these results is by necessity a subjective practice. Therefore the quantitative reporting of frequencies and centrality found in this study should be considered the objective portion of the study. If different researchers were given the same dataset and utilized the same software analysis packages, the researchers would be able to obtain the same results as reported in the current study. On the other hand the resultant discussion of emergent topics and relationships between citations required subjective analysis. Thorough discussions of the choices made by the researcher explaining the reasoning behind the reported conclusions are therefore critical when using the SNA approach found in the current study.

The benefit of using SNA was that the most frequently related nodes in a network could be easily identified, which enabled the identification of common topics and citations within the networks. The weakness of this method in this case was that information about the relationship ties themselves was not available. For example, Social Construction of Knowledge and Cognitive Presence were found to be highly related topics in both of the co-citation networks. Yet the data did not make available the nature of this relationship. Did these topics support each other? Were they adversarial? Further study utilizing a different research methodology, such as qualitative content analysis, would be necessary to address the nature of the relationships found in the SNA network ties.

<div align="center">Summary and Conclusions</div>

Graduate Researchers: Universities and Departments

The study included 3954 dissertations. There were 1110 distance education dissertations produced between 2000-2004. There was an increase to 1440 dissertations between 2005-2009 and a slight decline to 1404 dissertations between 2010-2014. The year with the most dissertations published was 2008 with 344.

During 2007 the PQDT began to allow graduate students to add department to their dissertation listing in the database. No department data was listed for 65% of the dissertations in the sample. Analysis of the remaining 45% of dissertations with departments led to a taxonomy of department types (Table 10).

Results showed that a large majority, 78%, of dissertations came from education oriented department types including general Education departments (such as colleges of education and teacher training departments), Educational Leadership, Educational

Technology, and Educational Psychology. Non-education oriented department types where distance education dissertations were consistently published each year from 2007-2014 were Business, Psychology, Computer Science, and Nursing. These results indicated that disciplines outside of education have not only begun to utilize but also research distance education in their fields.

Each dissertation in the sample included a university affiliation. The researcher segmented universities into three types: Public, Private, and For Profit. Table 11 illustrated the number of dissertations published by time and university type.

Distance education based dissertations came predominately, 63%, from Public universities over the three time periods.   Private universities accounted for 12% and For Profit Universities accounted for 25% of study sample. Tables 12, 13, and 14 detailed the top dissertation granting universities across university types and the three time periods.

Results showed that Capella University produced 6 times more dissertations than the next most frequent degree granting institution from 2005-2009 and 3 times as many dissertations as the next most frequent degree granting institution from 2010-2014.   It was also the top degree granting university across the Education, Miscellaneous, Educational Psychology, and Business department types, as well as dissertations that did not include a department. The data indicated that Capella University largely shapes the distance education graduate researcher population, since one out of every four of these researchers come from the school.

These results also indicated that a large number of doctoral students wanted to complete coursework through an online degree program. This is further supported by the result that from 2010-2014 68% of the top 10 universities for each university type offered

an online PhD or EdD in at least one of the department types included in the study, see

Table14.

Dissertation Reference Sections

Reference sections were not available for dissertations listed in the PQDT before

2008. Of this study's n=3,954 sample, 338 (9%) of the dissertation listings included

reference section citations. Therefore results represented only a small sub-sample of the

total study sample. Although in 2014 more than 50% of the dissertations listed in the

PQDT included reference sections, indicating future studies of reference sections may

enjoy larger sample sizes.

Results showed that journal articles had a shorter shelf life and there was much

variation between the top ranked journal articles between the 2005-2009 and 2010-2014

time periods, Tables 17 & 18. The most commonly cited articles were predominately

distance education oriented. The top ranking journals, which were consistently distance

education oriented across both time periods, further supported this.

Conversely, books were utilized to ground dissertations in statistics, methods, and

theory, Tables 21 & 22. Little variation was seen between the top cited books between

the two periods. This high citation rate across both time periods allowed topics of

longevity to be identified. These included learning communities, distance education

foundations, diffusion of innovation, and adult learning.

Relationships Between Topics in Dissertation Reference Section Co-Citation Networks

Co-citation analysis allowed for the relationships between citations to be

investigated, Figures 6 & 7. Figure 15 builds off of Figure 6 to illustrate the relationships

of topics observed in the co-citation network from 2005-2009.
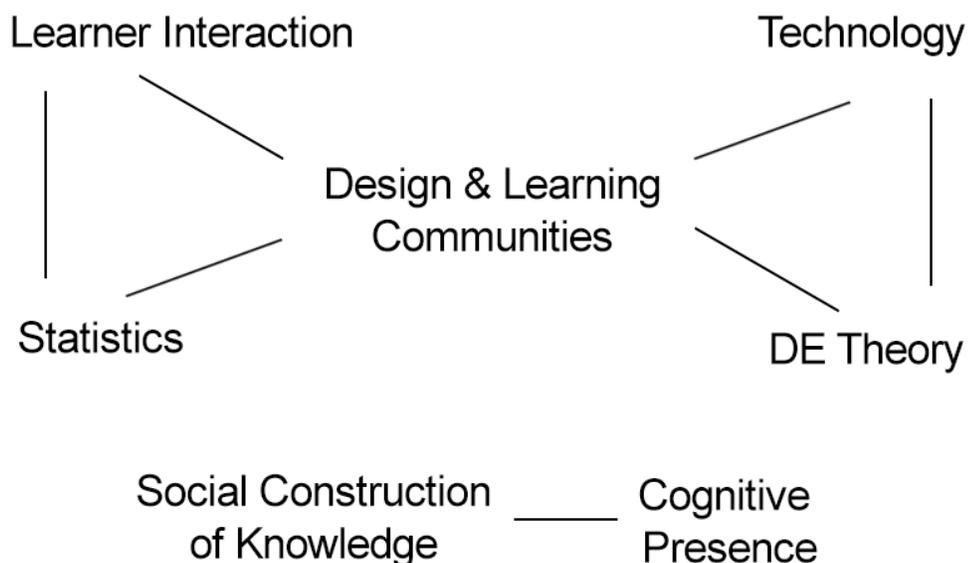
Figure 15. Relationships of Topics found in Co-Citation Network 2005 - 2009

Design and Learning Communities was the most central node and the topic with the most relationships to other topics in the network. It is related to Learner Interaction, Technology, Statistics, and Distance Education Theory. This suggested that during this time period graduate researchers commonly utilized citations addressing Design and Learning Communities in tandem with these other topics to support their research. The topics Learner Interaction and Statistics sources, as well as Technology and DE Theory were also cited together during the 2005-2009 time period.

Figure 15 also illustrates that the co-citation network for 2005-2009 had a sub-group between Social Construction of Knowledge and Cognitive Presence that was not related to the main cluster. This indicated a strong relationship between these topics and this relationship occurred most frequently independent of relationships to the other commonly cited sources in the network.

Figure 16 shows that Qualitative Methods was the topic most frequently co-cited during the 2010-2014 time period. This shift of the central node from Design and Learning Communities during the first time period to Qualitative Methods during the second time period suggested that graduate researchers increasingly ground their research in methods based literature. It was observed that in the main cluster co-cited topics were related specifically to methods based citations and not to each other.
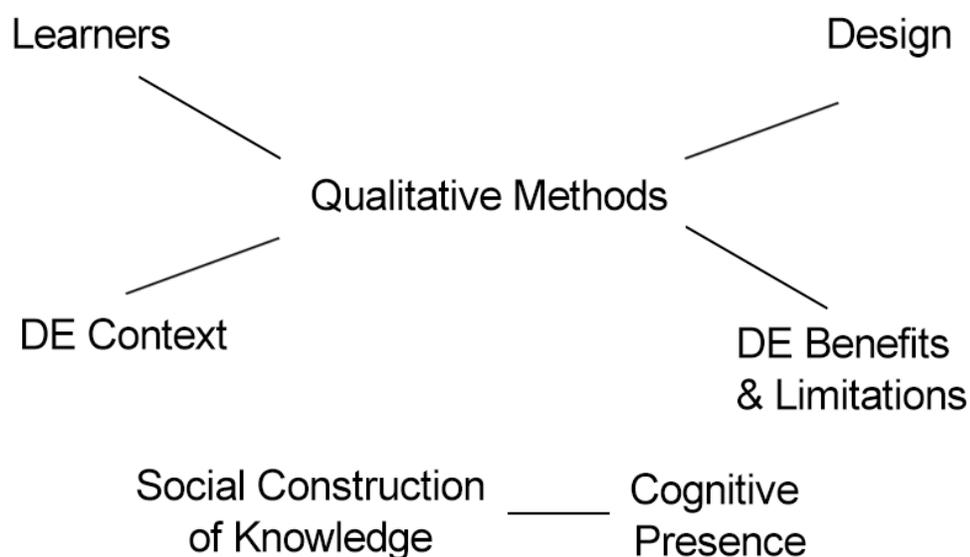

Figure 16. Relationships of Topics found in Co-Citation Network 2010 - 2014

The social construction of knowledge and cognitive presence sub-group was also present during the second time period. This indicated that these topics continued to be of interest to graduate researchers during the second time period, and remained largely related to each other and not commonly cited with the other sources found in the co-citation network.

Research Topics In Distance Education Dissertations

Classifications, keywords, and emergent keywords algorithmically derived from dissertation abstracts were all analyzed as research topic indicators for dissertations.

PQDT Classifications (see appendix 6) were chosen by doctoral students to characterize their dissertation. Analysis indicated that across the three time periods researchers regularly chose classifications in the areas of technological medium, educational context, issues of design, issues around instructor, and issues of management. Doctoral student were free to choose any keywords to characterize their research. The researcher of this study hypothesized a high level of specificity in keywords. Yet results showed that doctoral students used broad term keywords to characterize their dissertations. Results indicated that similar to classification across the three time periods keywords focused on areas of educational context, design, instructor, and management.

Across the three time periods, Student was the most central node in emergent keyword co-occurrence networks, Figures 8, 9, & 10. This indicated that Student was a predominant topic found in dissertation abstracts. Participation and Interaction were highly connected with Groups, Community, and Relationships. Administration, Faculty and Perception all shared relationship ties. Course, Instructor, and Student were all related to Technology. Participant, Faculty, Course, Instructor, Interaction, and Technology were all positioned in the network to have short paths to get to any other node, yet structurally the Student node was so central it was hard to determine the relationships between the other nodes in the networks.

To counter for this an analysis of the networks without the Student and Participant nodes was undertaken, Figures 11, 12, &13. Results showed that Instructor/Teachers became a central node in the networks. Learn, Learner, Interaction, Relationship, and Community nodes were clustered closely in the networks. Interaction was also a highly

central node that connected with Instructor, Technology, Learner, Group, Community, and Course among others.

The three indicators classification, keyword, and emergent keywords offered a triangulated view into the topics distance education dissertations covered over the past fifteen years. The researcher analyzed the results of the three analyses and found that seven topics were pervasive across the three time periods; Student, Instructor, Interaction, Administration and Management, Design, Educational Context, Technological Medium, see Table 39.

Little change was observed over the three time periods, rather these predominant topics were observed in classifications, keywords, and emergent keywords during the time periods 2000-2004, 2005-2009, and 2010-2014. While subtle changes in bibliometric ordering and frequency as well as difference in centrality measurements were observed across time periods, the chief trend observed was that these seven research areas were all found across time.

The topic areas that emerged from analysis are quite broad. This may be one explanation for observing little change in topics across the three time periods. The researcher hypothesized a high level of specificity of keywords to emerge, particularly from the emergent keyword analysis. This was not found. This may indicate that keywords and keywords derived from abstracts are not adequate enough to provide research topic information with a high degree of specificity. Future researchers should take note of this finding.

Comparison to Zawacki-Richter (2009) CRADE Framework

The seven dissertation topic areas that emerged in this study were Student,

Instructor, Interaction, Design, Educational Context, Administration and Management,

and Technological Medium. Table 40 illustrates that these co-related to seven of the

research areas forwarded by Zawacki-Richter (2009a)

Topics Addressed in Doctoral Dissertation on Distance Education 2000-2014	CRADE Zawacki-Richter (2009a)
Student	Learner characteristics
Instructor	Professional development and faculty support
Interaction	Interaction and communication in learning communities
Administration and Management	Management and organization
Design	Instructional Design
Educational Context	Distance teaching systems and institutions
Technological Medium	Educational Technology

Table 40. Comparison of Topics Addressed in Doctoral Dissertation on Distance
Education 2000-2014 and the CRADE Framework

The remaining eight CRADE research areas were not found consistently in this

study's results. Bozkurt et al. (2015) found similar results. In their citation analysis that

utilized the CRADE they reported that the majority of studies investigated in their study

fell within only a few areas of the framework: Interaction and Communication in

Learning Communities, Learner Characteristics, Instructional Design, and Educational

Technology.

Relationships Between Research Topics in Distance Education Dissertations

The seven research topics found in distance education dissertations were similar to seven of the fifteen research areas found in Zawacki-Richter (2009a) framework. While this helps to confirm the accuracy of research topics forwarded by this study, the methodology used here enabled the researcher to take the contribution of this study one step further. The SNA approach used in the study allowed for investigation of relationships between topics. The relationships between nodes indicating each of the seven emergent topic areas found in this study were analyzed. Figure 14 details the relationships between topics.

Figure 14 shows that in dissertations Administration and Management topics only frequently co-occurred with Instructor or Educational Context topics. Dissertations that addressed Design oriented topics only frequently co-occurred with Student, Instructor, or Educational Context. The combinations of the remainder of the topics frequently co-occurred with each other. This was important to note. What was found in this study was that dissertation research topics were not located in the single domain of one research topic. Rather, since an SNA approach that focuses on relationships was utilized, it was found that dissertations utilized combinations of the seven topics Student, Instructor, Interaction, Administration and Management, Design, Educational Context, and Technological Medium in their studies.

<center>Limitations</center>

The sample used in this study was gathered from a single database source, the PQDT. While Davies (2010) noted that 97.2% of dissertations from the US and 87.2% of dissertations from Canada are indexed in the PQDT, dissertations not listed in the PQDT database are not represented in the study. The sample is also limited by the search terms

used to indicate that dissertations were about distance learning. These included "distance education", "distance learning", "online learning", e-learning", electronic learning", "network learning", "distributed learning", or "web-based learning". Dissertations that did not use one of these terms in the title, classification, keywords, or abstract were not included in this study. This sample was further limited in regards to the reference section analysis. Only 16% of the 2064 dissertations since 2008 included reference sections. The reference section analysis included here only represents the findings from this small sub-sample and cannot be generalized to the larger study sample.

Results were also limited to analysis by the study researcher. Categorization of department types, analysis, and interpretation of results were performed and reported by this single researcher. Subjectivity was not countered by enlisting multiple raters to interpret the results.

<div align="center">Significance</div>

The first section of the current study, *Graduate Researchers: Universities and Departments*, contributes a cross disciplinary picture of the graduate researcher population writing distance education dissertations. Previous studies often look at distance education in the context of a single discipline. The current study found that distance education dissertations came from 17 different department types. The majority of dissertations came from education-oriented department types like Educational Leadership, Educational Psychology, and Educational technology. Yet it was also observed distance education dissertations regularly came from four non-education oriented department types: Business, Psychology, Computer Science, and Nursing.

These 17 department types can help future researchers who are interested in a more holistic and cross-disciplinary perspective in distance education. They may use the department types found in this study as starting points to look deeper in these discourses' research body for distance education research that could contribute back out to the larger distance education research community.

Another important contribution made in the *Graduate Researchers: Universities and Departments* section is that one out of four graduate students in the study sample graduated from an online degree program at Capella University. Furthermore from 2010-2013 more than two thirds of the top degree granting universities in the public, private, and for profit sectors all published a distance education dissertation from a fully online doctoral program. These findings may be helpful to researchers investigating design of online degree programs. This study found that graduate students doing dissertations in distance education frequently completed online degrees. Many universities are considering issues of quality in fully online programs. This study shows a preference for online degrees, at least by the population of students researching online learning.

The *Dissertation Reference Sections* portion of the study illustrated the sources that graduate researchers used to ground their distance education dissertations in the literature. One contribution here was that the co-citation networks showed the relationships of citation topics between the 2005-2009 and 2010-2014 time periods. One observation was that learning communities and instructional design were at the center of the relationships during the early time period, yet this changed to qualitative methods being the more central topic during the latter time period. This approach for mapping

research topics using co-citation networks may help other researchers to map knowledge structures and topics in their own field.

The *Research Topics In Distance Education Dissertations* section of the study contributed seven frequently addressed research topics found in dissertations. These were Student, Instructor, Interaction, Design, Educational Context, Administration and Management, and Technological Medium. These may act as a framework of research topics for future research and are novel because they were derived from a cross disciplinary sample that moves beyond a singular department type or discourse.

Another contribution was that these seven dissertation topics confirm areas in Zawacki-Richter (2009a) and Zawacki-Richter & Anderson (2011) CRADE model. Zawacki-Richter, et al. (2009b) and Bozkurt et al (2015) utilized the CRADE as a framework for content analysis reviews of research. In both articles a study was categorized to a single research topic area of the CRADE. The current study further contributes a picture of the relationships found between research topics in dissertations, Figure 14. It was found that doctoral students addressed not a single research topic but a combination of the seven research topics found in this study. Therefore moving forward, when performing a content analysis of research studies, using the framework presented here or with the CRADE, it may not be appropriate to categorize a study into a single research topic area, but rather be open to categorization in a combination of areas.

Finally, this study contributed a highly automated SNA approach. The algorithms developed for this study automated the import and formatting of data. The data could then be taken into analysis software. Use of abstracts and natural language processing enabled a much higher n size (n=3954) to be investigated than in comparison with the

only other study to analyze distance education dissertations Davies et al. (2010) where

n=100. The heavy lifting could then be dedicated to the interpretation of the results. This

can help future researchers because this approach is not distance education specific,

rather this SNA approach to understanding the relationships between topics and reference

citations may be applied to any discourse.

Further Research: Graduate Researchers: Universities and Departments

In the area of *Graduate Researchers: Universities and Departments* it is

recommended that future research investigates non-education-oriented department types

to see what the state of distance education research is within these fields. Particularly,

researchers should investigate the four discourses this study found to commonly grant

distance education dissertations: Business, Psychology, Computer Science, and Nursing.

Do these fields have contributions in any of the seven topic areas previously not found in

the larger distance education research conversation? Now that distance education has

become common in these areas, are there methodological approaches that may be novel

to the field? Finally, how can researchers forward a more holistic and cross-disciplinary

discussion about distance education?

It is also recommended that research into the "Capella Phenomenon" be

undertaken. With one out of every four dissertations in the field being granted by

Capella, further research to see how this entity affected the results of the current study

should be done. If the large for profit schools, including Capella, are removed from the

study sample are the results to the current study the same? A comparison should be done.

Also it is of interest to follow up on the current study to see what graduate researchers in

the distance education pursue after graduation. Do they pursue academics, professional careers, or alternative routes like consulting?

Further Research: Research Topics In Distance Education Dissertations

It was found that dissertations utilized combinations of the seven topics Student, Instructor, Interaction, Administration and Management, Design, Educational Context, and Technological Medium in their studies. The relationships between these topics were shown in Figure 14. This figure shows the trends of topic combinations found in doctoral research over the past fifteen years, as evidenced by the results of this study. Of note, two topic combination sets have not been largely addressed by dissertations in distance education: Interaction and Instructional Design, and Technology and Instructional Design. Future research should investigate these two combinations of topics as they are not covered in the research body and provide an opportunity for future research to fill these gaps. Likewise the Administration and Management topic was only related to Instructor and Education Context. Future researchers may find that combining Administration and Management with the remaining four topics (Student, Design, Interaction, and Technology) may result in findings not addressed in previous dissertation research. In this way the results on an SNA investigation into research topics and the subsequent relationship map, Figure 14, can 1) illustrate what research topic combinations have been previously addressed in a field and 2) therefore act as a map for future researchers in finding topic combinations that need investigation.

Further Research: SNA Methods

The SNA co-occurrence approach to topics indicated seven commonly addressed research topics. Yet these were found to not have a high level of specificity. Keyword

co-occurrence limits analysis to co-occurring pairs.   Natural language processing techniques enable the investigation of co-occurrence beyond a single pair. Rather an N-Gram approach enables the measurement of N co-occurrences. For example, with this approach research topics that only co-occur 3 or 4 times together would be pulled for analysis. It is recommended that a study similar to the current study design utilize a N-Gram approach to co-occurrence to see if emergent dissertation research topics with a higher level of specificity than found in this study occur.

Future Research: International Dissertations in English

The current study used the PQDT to retrieve dissertation records from the US and Canada. The number of distance education dissertations in English available from other countries listed in the PQDT was too small to be representative of the scholarship in those countries. Further research into international distance education dissertations published in English is recommended. A number of databases listing international dissertations and theses are available.

First the ProQuest Dissertations & Theses: Global database may be the first step. It is the largest single repository of graduate dissertations and theses in the world, is international in scope, and gets deposits from universities in 88 countries (http://www.proquest.com/products-services/pqdtglobal.html). Since it is a PQDT product many of the algorithms used in the current study may work to import and format data from this database. The TROVE: The Australasian Digital Theses Program includes theses at all levels from Australian and New Zealand universities (http://trove.nla.gov.au).   The British Library provides Ethos, the E-these online service, and lists doctoral theses from 131 participating UK institutions (http://ethos.bl.uk/). The

DART-Europe E-theses Portal lists doctoral theses from 108 institutions in participating European Union countries, including England, Spain, Catalonia, France, Germany, Italy, Magyar, Sweden, Greece, Romania, Armenia, Poland, Norway, and Serbia (http://www.dart-europe.eu). Shodhganga, the Indian Electronic Theses and Dissertation Repository, provides access to theses and dissertations from 226 contributing universities (http://shodhganga.inflibnet.ac.in). Therefore, the time is ripe to engage in international reviews of distance education research.

Concluding Comments

Algorithmic research approaches do not remove the researcher from the equation. Analysis and interpretations remain solely based in the researcher's subjective hands. What algorithmic approaches, such as found in this study, afford is the possibility to address samples previously found unapproachable. Davies (2010,44-45) notes "the time and effort involved in reading and categorizing a decade's worth of available (graduate) research manuscripts presented a daunting task". This study forwarded a methodology that lets this largely uninvestigated body of research be analyzed. As database indexing becomes more complete and algorithmic techniques improve, these methodological approaches may do some of the heavy lifting for the researcher facing similar daunting tasks. It is not a matter of John Henry versus the steam machine, rather a matter of us using the tools we have available to help us further our understanding of our field and to understand the field from another perspective.

REFERENCES

Adams, J. (2014). Bibliometrics: The citation game. Nature, *510*(7506), 470-471.

Alencar, A. (2012) ISI (*.isi). Retrieved from
http://wiki.cns.iu.edu/pages/viewpage.action?pageId=1933374

Anglin, G. J., Morrison, G. R. (2000). An analysis of distance education research: Implications for the instructional technologist. *The Quarterly Review of Distance Education, 1*(3), 198-194.

Backhaus, K., Lügger, K., & Koch, M. (2011). The structure and evolution of business-to-business marketing: a citation and co-citation analysis. *Industrial Marketing Management, 40*(6), 940-951.

Barnes, J.A. (1972) *Social Networks.* Reading, MA: Addison-Wesley.

Beile, P. M., Boote, D. N., & Killingsworth, E. K. (2004). A microscope or a mirror?: A question of study validity regarding the use of dissertation citation analysis for evaluating research collections. *The Journal of Academic Librarianship, 30*(5), 347-353.

Benckendorff, P. (2009). Themes and trends in Australian and New Zealand tourism research: A social network analysis of citations in two leading journals (1994-2007). *Journal of Hospitality and Tourism Management,* 16, 1-15.

Berge, Z., & Mrozowski, S. (2001). Review of research in distance education, 1990 to 1999. *The American Journal of Distance Education, 15*(3), 5-19.

Bird, S, Loper, E., & Klein, E. (2009), Natural Language Processing with Python. O'Reilly Media Inc.

Borgatti, S., Mehra, A., Brass, D. and Labianca, G. (2009) Network analysis in the social sciences. *Science, 323*(5916):892–895.

Borgatti, S.P., Everett, M.G., & Freeman, L.C. (2002). Ucinet for Windows: Software for Social Network Analysis. Harvard, MA: Analytic Technologies.

Borgatti, Stephen P., Everett, M.G., Johnson, J.C. (2013) Analyzing social networks. London: Sage

Borgman, C. L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information and Technology, 36*(1), 2-72.

Bozkurt, A., Akgun-Ozbek, E., Yilmazel, S., Erdogi, E., Ucar, H., Guler, E., Sezgin, S.,

Karadeniz, A., Sen-Ersoy, N., Goksel-Canbek, N., Dincer, D. G., Ari, Suleyman, & Aydin, C.H. (2015) Trends in distance education research: A content analysis of journals 2009-2013. *International Review of Research in Open and Distance Learning, 16*(1). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1953

Brass, D. J., Brurkhardt, M. E. (1992).Centrality and power in organizations. In N. Nohria & R. G. Eccles (Eds.), *Networks and Organizations: Structure, Form, and Action* (pp. 191-215). Boston, MA: Harvard Business School Press.

Broadus, R. N. (1987). Toward a definition of "bibliometrics". *Scientometrics 12*(5), 373-379.

Burt, R. S., Kilduff, M., & Tasselli, S. (2013). Social network analysis: Foundations and frontiers on advantage. *Annual review of psychology, 64*, 527-547.

Butts, C.T. (2008). Social network analysis with sna. *Journal of Statistical Software, 24*(6), 1-51.

Catanese, S. A., De Meo, P., Ferrara, E., Fiumara, G., & Provetti, A. (2011). Crawling facebook for social network analysis purposes. In Proceedings of the international conference on web intelligence, mining and semantics. p. 52.

Cho, Yonjoo, Park, S., Jo, S.J. &Suh, S. (2013) The landscape of educational technology viewed from the ETR&D Journal. *British Journal of Educational Technology 44*(5). 677-694.

Coulter, N., Monarch, I., & Konda, S. (1998). Software engineering as seen through its research literature: A study in co-word analysis. *Journal of the American Society for Information Science, 49*(13), 1206-1223.

Culnan, M. (1986). The intellectual development of management information systems. *Management Science*, 32, 156-172

Davies, R. S., Howell, S. L., & Petrie, J. A. (2010). A review of trends in distance education scholarship at research universities in North America, 1998-2007. *The International Review of Research in Open and Distance Learning, 11*(3), 42-56. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/876/1602

De Bellis, N. (2009). Bibliometrics and citation analysis: from the science citation index to cybermetrics. Scarecrow Press.

Ding, Y. (2011). Scientific collaboration and endorsement: Network analysis of coauthorship and citation networks. *Journal of informetrics, 5*(1), 187-203.

Diodato, V. (1994). *Dictionary of bibliometrics*. New York: Haworth Press.

Evans, T. D., & Green, R. (2013). Doctorates for professionals through distance education. In M. Moore (Ed), *Handbook of distance education*, 654-669.

Everett, M. G., Borgatti, S. P. (2005). Extending Centrality. In P. J. Carrington, J. Scott & S. Wasserman (Eds.), *Models and Methods in Social Network Analysis* (pp. 57-76). New York, NY: Cambridge University Press.

Freeman, L. C. (1979). Centrality in social networks: Conceptual clarification. *Social Networks, 1*(3), 215-239.

Freeman, L. C. (2011). The development of social network analysis: With an emphasis on recent events. In J. Scott & P. J. Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 26-39). Thousand Oaks, CA: SAGE Publications.

Fudell, D., & Hardy, D. (1998). Distance education: A primer. [Online]. Available: http://www.utexas.edu//cc/cit/de/deprimer/ [2013, March 3].

Garfield, E. (1977). *Citation indexing: Its theory and application in science.* Philadelphia, PA: ISI Press.

Garfield, E. (1988). Announcing the SCI Compact Disc Edition: CD-ROM gigabyte storage technology, novel software, and bibliographic coupling make desktop research and discovery a reality. *Current Comments, 22*, 160-170

Garrison, D. (1985). Three generations of technological innovations in distance education. *Distance Education, 6*(2), 235-241.

Hanneman, R. A., Riddle, M. (2005). Introduction to social network methods. Riverside, CA: University of California, Riverside. (published in digital form at http://faculty.ucr.edu/~hanneman/).

Hanneman, R. A., & Riddle, M. (2011). A brief introduction to analyzing social network data. In J. Scott & P. J. Carrington (Eds.), *The SAGE Handbook of Social Network Analysis* (pp. 331-339). Thousand Oaks, CA: SAGE Publications.

Ho, Y. S. (2012). Top-cited articles in chemical engineering in Science Citation Index Expanded: A bibliometric analysis. *Chinese Journal of Chemical Engineering, 20(*3), 478-488.

Kadushin, C. (2011). Understanding Social Networks: Theories, Concepts, and Findings. Oxford University Press.

Khan, B. H. 1997. Web-based instruction: What is it and why is it? In Web-based instruction, ed. B. H. Khan, 5–18. Englewood Cliffs, NJ: Educational Technology Publications.

Kim, A. C. H. (2012). Knowledge structure in sport management: bibliometric and social network analyses. (Unpublished doctoral dissertation). Ohio State University, Columbus, OH.

Koble, M. A., Bunker, E. L. (1997). Trends in research and practice: An examination of the American Journal of Distance Education 1987 – 1995. *The American Journal of Distance Education, 11*(2), 19-37.

Lee, Y., Driscoll, M. P., & Nelson, D. W. (2004). The past, present, and future of research in distance education: Results of a content analysis. *American Journal of Distance Education, 18*(4), 225-241.

Lee, P.-C., & Su, H.-N. (2010). Investigating the structure of regional innovation system research through keyword co-occurrence and social network analysis. *Innovation: Management, Policy, & Practice, 12*(1), 26-40.

Leong, S. M. (1989). A citation analysis of the Journal of Consumer Research. *Journal of Consumer Research*, 15, 492-497.

Lewis, D.D., Yang, Y., Rose, T.G., Li, F. (2004). RCV1: A new benchmark collection for text categorization research. Journal of Machine Learning Research 5. 361-397.

Marin, A., & Wellman, B. (2011). Social network analysis: an introduction. In J. Scott & P. J. Carrington (Eds.), *The SAGE Handbook of Social Network Analysis*. Thousand Oaks, CA: SAGE Publication.

MIT Libraries (2012). Stop words. Retrieved from http://libguides.mit.edu/content.php?pid=36863&sid=271410

Moed, H. F. (2005). Citation Analysis in Research Evaluation. Dordrecht, Netherlands: Springer.

Moore, M. G. (1990). Background and overview of contemporary American distance education. *Contemporary issues in American distance education* (pp. xii−xxvi). New York: Pergamon Press.

Moore, M. G. (2003). Preface. In M. G. Moore & W. G. Anderson (Eds.), *Handbook of distance education* (pp. ix-xii). Mahwah, NJ: Lawrence Erlbaum.

Moore, M., & Kearsley, G. (2005). *Distance education: A systems view*. (2nd ed.). Belmont, CA: Wadsworth Publishing.

Natural Language ToolKit. (2012a) NLTK 2.0 documentation: stem Package. Retrieved from http://nltk.org/api/nltk.stem.html

Natural Language ToolKit. (2012b) NLTK 2.0 documentation: chunk Package. Retrieved from http://nltk.org/api/nltk.chunk.html?highlight=phrase

Onyancha, O. B., & Ocholla, D. N. (2004). A comparative study of the literature on HIV/AIDS in Kenya and Uganda: A bibliometric study. *Library & Information Science Research,* 26, 434-447.

Osareh, F. (1996). Bibliometrics, citation analysis and co-citation analysis: A review of literature I. *Libri, 46*, 149-158.

Pagel, P. S., & Hudetz, J. A. (2011). An analysis of scholarly productivity in United States academic anaesthesiologists by citation bibliometrics. *Anaesthesia, 66*(10), 873-878.

Palmer, A. L., Sese, A., & Montano, J. J. (2005). Tourism and statistics: Bibliometric study 1998-2002. *Annals of Tourism Research, 32*(1), 167-178.

Phipps, R. Merisotis, J. (1999). What's the difference? A review of contemporary research on the effectiveness of distance education in higher education. Institute for Higher Education Policy. Available at http://www.iher.com/difference.pdf

Pilkington, A., & Meredith, J. (2009). The evolution of the intellectual structure of operations management - 1980-2006: A citation/co-citation analysis. *Journal of Operations Management,* 27, 185-202.

Reference Manager. (2011). RIS Format Specifications. Retrieved from http://www.refman.com/support/risformat_intro.asp

Reffay, C., & Martínez-Monés, A. (2013). Basic concepts and techniques in social network analysis. In Edited by R. Luckin, S. Puntambekar, P.Goodyear, B. L Grabowski, J. Underwood, N. Winters (Eds), *Handbook of Design in Educational Technology*. Routledge. 448-456.

Ritzhaupt, A. D., Stewart, M., Smith, P., & Barron, A. E. (2010). An investigation of distance education in North American research literature using co-word analysis. *The International Review of Research in Open and Distance Learning, 11*(1), 37-60. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/763/1483

Rourke, L., & Szabo, M. (2002). A content analysis of the Journal of Distance Education 1986-2001. *Journal of Distance Education, 17*(1), 63-74.

Salton, G. (1971). The SMART Retrieval System: Experiments in Automatic Document Processing. Prentice-Hall.

Schildt, H. (2004). SITKIS: Software for bibliometric data management and analysis version 1.0 manual. Retrieved from http://users.tkk.fi/hschildt/sitkis/

Schlosser, L.A., & Simonson, M. (2009). Distance education: Definition and glossary of terms (3rd ed.). Charlotte, NC: Information Age.

Sharplin, A., & Mabry, R. (2007). The relative importance of journals used in management research: An alternative ranking. *Human Relations*, 38, 139-149.

Sherry, K. (1996). Issues in distance learning. *International Journal of Educational Telecommunications, 1*(4), 337-65.

Simonson, M., Schlosser, C., Orellana, A. (2011) Distance education research: a review of the literature. *Journal of Computing in Higher Education 23*. 124-142.

Smith, L. (1981). Citation analysis. *Library Trends, Summer*, 83-106.

Shimp, U. R. (2007). Evaluation of the distance education literature: a content analysis using the institute for higher education policy benchmarks and selected bibliometric methods. (Unpublished doctoral dissertation). Oklahoma State University, Stillwater, OK.

Waugh, C. K., & Ruppel, M. (2004). Citation analysis of dissertation, thesis, and research paper references in workforce education and development. *The Journal of Academic Librarianship, 30*(4), 276-284.

White, Howard, Wellman, Barry and Nazer, Nancy (2004) 'Does citation reflect social structure?' *Journal of the American Society for Information Science and Technology, 55*(2): 111- 126.

Wiberley, S. E., Jr. (2003). A methodological approach to developing bibliometric models of types of humanities scholarship. *Library Quarterly, 73*(2), 121-159.

Zawacki-Richter, O. (2009a). Research areas in distance education: A delphi study. *International Review of Research in Open and Distance Learning, 10*(3), 1-17. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/674/1260

Zawacki-Richter, O., Bäcker, E. M., & Vogt, S. (2009b). Review of distance education research (2000 to 2008): Analysis of research areas, methods, publication, and authorship patterns. *International Review of Research in Open and Distance Learning, 10*(6), 21-50. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/741/1433

Zawacki-Richter, O., Anderson, T., Tuncay, N. (2010) The growing impact of open access distance education journals: a bibliometric analysis. *Journal of Distance Education 24*(3).

Zawacki-Richter, O., Anderson, T. (2011) The geography of distance education –
bibliographic characteristics of a journal network. *Distance Education 32*(3). 441-
456.

APPENDIX 1

REQUESTS FOR USE OF FIGURES

Request for reuse of frameworks and tables has been made and is pending approval for

inclusion in this study:

Zawacki-Richter et al., 2009b, p. 26 – CRADE Framework

APPENDIX 2

DISSERTATIONS: STUDY DATABASE FIELDS

```
CREATE TABLE IF NOT EXISTS `studyDB_full_2015_final` (

  `id` int(11) NOT NULL AUTO_INCREMENT,

  `importId` varchar(250) NOT NULL,

  `pqdtDocNum` varchar(250) NOT NULL,

  `title` varchar(250) NOT NULL,

  `author` varchar(250) NOT NULL,

  `abstract` blob NOT NULL,

  `advisor` varchar(250) NOT NULL,

  `classification` varchar(250) NOT NULL,

  `committeeMember` varchar(250) NOT NULL,

  `copyright` varchar(250) NOT NULL,

  `country` varchar(250) NOT NULL,

  `db` varchar(250) NOT NULL,

  `degree` varchar(250) NOT NULL,

  `degreeDate` varchar(250) NOT NULL,

  `department` varchar(250) NOT NULL,

  `departmentType` varchar(250) NOT NULL,

  `dissertationThesisNumber` varchar(250) NOT NULL,

  `documentType` varchar(250) NOT NULL,

  `isbn` varchar(250) NOT NULL,

  `keyword` varchar(250) NOT NULL,
```

```
`language` varchar(250) NOT NULL,

`pages` varchar(250) NOT NULL,

`placePub` varchar(250) NOT NULL,

`pqdtDocID` varchar(250) NOT NULL,

`pubInfo` varchar(250) NOT NULL,

`pubYear` varchar(250) NOT NULL,

`schoolCode` varchar(250) NOT NULL,

`source` varchar(250) NOT NULL,

`sourceType` varchar(250) NOT NULL,

`subject` varchar(250) NOT NULL,

`universityLocation` varchar(250) NOT NULL,

`university` varchar(250) NOT NULL,

`universityType` varchar(250) NOT NULL,

PRIMARY KEY (`id`)

) ENGINE=MyISAM  DEFAULT CHARSET=utf8 AUTO_INCREMENT=3954 ;
```

APPENDIX 3

CITATIONS: STUDY DATABASE FIELDS

```
CREATE TABLE IF NOT EXISTS `studyDB_2015_references` (

 `id` int(11) NOT NULL AUTO_INCREMENT,

 `dissId` int(11) NOT NULL,

 `degreeDate` varchar(250) NOT NULL,

 `formatted` int(11) NOT NULL,

 `refId` int(11) NOT NULL,

 `refType` varchar(250) NOT NULL,

 `refYear` varchar(250) NOT NULL,

 `refTitle` varchar(250) NOT NULL,

 `refPublication` varchar(250) NOT NULL,

 `refJournalIssue` varchar(250) NOT NULL,

 `refOtherDetails` varchar(250) NOT NULL,

 `refAuthorRaw` varchar(250) NOT NULL,

 `notFormattedRecord` varchar(250) NOT NULL,

 `refRaw` blob,

 `departmentType` varchar(250) NOT NULL,

 PRIMARY KEY (`id`)

) ENGINE=MyISAM  DEFAULT CHARSET=utf8 AUTO_INCREMENT=41409;
```

APPENDIX 4

ISI FILE EXPORT FIELDS

PT   Publication type (e.g., book, journal, book in series)

AU   Author(s)

TI   Article title

SO   Full source title

DE   Author keywords

ID   KeyWords Plus

AB   Abstract

CR   Cited references

LA   Language

RP   Reprint address

J9   29-character source title abbreviation

JI   ISO source title abbreviation

VL   Volume

IS   Issue

PY   Publication year

TC   Times cited

C1   Research addresses

ER   End of record

APPENDIX 5

DEPARTMENTS ASSIGNED TO THE 17 DEPARTMENT TYPES TYPOLOGY

Nursing
$arr_nursing = array('Nursing',
'Public Health',
'Nursing and Allied Health Professions',
'College of Nursing',
'Nursing Practice',
'Rehabilitation',
'Department of Health Education',
'Health Education',
'Health Professions',
'Lifespan Development and Educational Sciences',
'Health Services',
'Marybelle and S. Paul Musco School of Nursing and Health Professions',
'Rehabilitation Science',
'Human Factors and Ergonomics');

Education
$arr_education = array('Education',
'Curriculum, Teaching and Learning',
'College of Education',
'Curriculum and Instruction',
'Teaching and Learning/Educational Specialities',
'School of Educational Studies',
'Education - Ph.D',
'Education 0249',
'Education, Joint Program Cal Poly SLO',
'Educational Studies and Research',
'Department of Education',
'School of Education',
'Curriculum &amp; Instruction',
'Department of Educational Studies',
'Educational Studies',
'Teaching &amp; Learning',
'Department of Secondary Education',
'Curriculum and Teaching',
'Science and Mathematics Education (SMED)',
'Lynch School of Education',
'Education, Curriculum and Instruction',
'Teaching and Learning',
'Teaching, Curriculum, &amp; Learning',
'Graduate School of Education',
'TC: Comparative and International Education',

'Educational Theory and Practice-Curriculum and Instruction',
'Education - Curriculum &amp; Instruction (Ed. D.)',
'Education - Curriculum &amp; Instruction (Ed.D.)',
'Curriculum &amp; Instruction (PhD)',
'School of Education: Curriculum and Instruction',
'College and Graduate School of Education, Health and Human Services / School of Teaching',
'Education (all programs)',
'Professional Studies in Education',
'Education and Human Services',
'Secondary Education',
'Curriculum &amp; Instruction/Literacy Studies',
'Education, Curriculum &amp; Instruction',
'College of Education and Human Services',
'Curriculum, Instruction, and Assessment',
'Curriculum, Instruction, and Teacher Education',
'Educational Research',
'Instructional Design, Development and Evaluation',
'Learning and Instruction',
'Learning and Teaching',
'San Francisco, HSOE',
'School of Education - Ed.D',
'TC: Economics and Education',
'Teacher Education',
'Education (School of)',
'Mathematics and Science Education Program');

Special Education
$arr_speced = array('Special Education',
'Special Education and Rehabilitation',
'Deaf Education/Deaf Studies',
'Education, Health, and Human Services',
'Curriculum, Instruction and Special Education');

Higher Education
$arr_highed = array('Higher Education',
'Higher and Postsecondary Education',
'Higher Ed/Community College Ed',
'School of Education: Teacher Education and Higher Education',
'Community College Education',
'Adult and Community College Education',
'Adult Education',
'Adult, Career and Higher Education',
'Higher Education Management',
'Higher Education Program',
'Supervision, Curriculum, and Instruction-Higher Education',

'Higher, Adult, and Lifelong Education - Doctor of Philosophy',
'Community College Leadership Program');

Education Technology
$arr_edtech = array('Educational Technology',
'School of Information Systems and Technology',
'Information and Learning Technologies',
'Information Systems Technology',
'Computing Technology in Education (MCTE, DCTE)',
'Information Science (DISC)',
'Information Systems (DISS)',
'Educational Technology, Research and Assessment',
'Instructional Technology',
'Information Science and Learning Technologies',
'School of Information Science and Learning Technology',
'Information Technology',
'Information &amp; Library Science',
'Information Technology (PhD)',
'Curriculum and Instruction: Educational Media',
'Instructional Technology (EdDIT)',
'Information Sciences',
'Information Studies',
'Learning Technologies',
'Curriculum and Instruction - Instructional Design and Technology',
'Adult Learning &amp; Technology',
'Information, LTEC, ECMP',
'E-Learning',
'Adult Learning and Technology',
'Educational Techology',
'Instructional Systems',
'Library & Information Science',
'Teaching, Learning, and Technology',
'The School of Information Studies',
'Library &amp; Information Science',
'Curriculum, Instruction, and Media Technology');

Organizational
$arr_org = array('Organizational Learning and Instructional Technology',
'Human and Organizational Learning',
'Organizational Studies',
'Workforce Development Education',
'Workforce Education and Development',
'Organization Development',
'Organizational Leadership',
'The School of Human and Organization Development',
'Work-Based Learning Leadership');

Business
$arr_business = array('Technology Management',
'Management',
'Doctor of Management Program',
'School of Business',
'School of Business and Technology',
'Applied Management and Decision Sciences',
'Business Administration',
'School of Business &amp; Economics: Information Systems and Operations Management',
'Nance College of Business Administration',
'Business - Foster School of',
'College of Business',
'School of Business and Technology Management',
'Business and Technology Management');

Ed Lead
$arr_edlead = array('Educational Leadership',
'Educational Leadership, Policy, and Technology Studies',
'Higher Education Administration',
'Instructional Leadership',
'The School of Educational Leadership and Change',
'Educational Administration and Leadership (CSU/UCI Joint) - Ed.D',
'EducLeadrshp (JtEdD-SDSU&amp;CSUSM)',
'Education(Leadership)',
'Education(Policy, Planning and Administration)',
'Educational Leadership and Administration',
'Educational Leadership &amp; Policy Studies',
'Educational Leadership, Administration, and Foundations',
'Department of Educational Leadership',
'Educational Leadership and Policy Studies',
'Education Leadership',
'Advanced Studies, Leadership, and Policy',
'Education Policy, and Leadership',
'Education/Higher Education Administration',
'Educational Policy and Administration',
'Instructional Systems, Leadership, and Workforce Development',
'Educational Leadership and Policy Analysis',
'Educational Administration',
'Educational Leadership &amp; Higher Education',
'Leadership',
'Educational Leadership, Management, and Policy',
'Education, Leadership &amp; Policy',
'Educational Leadership (EDD)',
'School of Education: Educational Leadership and Cultural Foundations',

'Educational Leadership &amp; Foundations',
'Department of Counseling &amp; Higher Education',
'ED Teaching and Learning (Columbus campus)',
'Educational Administration and Supervision',
'Department of Educational Leadership and Policy Studies',
'Higher Education Leadership and Management',
'Higher Education Administration, Community College Leadership Program',
'Ed. Leadership &amp; Admin',
'Community College Leadership',
'Educational Policy &amp; Leadership',
'Policy and Leadership',
'Administration, Leadership, and Technology',
'Administrative and Policy Studies',
'College of Education - Educational Leadership',
'EdD Educational Administration',
'Education Policy Studies',
'Educational Administration and Foundations: Educational Administration',
'Educational Leadership and Higher Education',
'Educational Leadership and Management',
'Educational Leadership and Management (School of Education)',
'Educational Leadership and Policy',
'Educational Leadership Studies',
'Educational Policy Studies',
'Educational Policy, Organization and Leadership',
'Educational Psychology and Educational Technology - Doctor of Philosophy',
'EducLeadrshp (JtEdD-SDSU & CSUSM)',
'Gary Cook Graduate School of Leadership',
'Leadership and Innovation',
'Theory and Policy Studies in Education',
'EducLeadrshp (JtEdD-SDSU &amp; CSUSM)',
'Education - Ed.D. Leadership',
'Educational Leadership',
'Education (Leadership)',
'Educational Leadership, Policy and Law',
'The School of Educational Leadership and Change',
'Educational Administration and Leadership',
'Education - Ed.D. Leadership 0659',
'Educational Leadership and Counseling Department',
'Department of Instructional Leadership and Academic Curriculum');

Ed Pysch
$arr_edpsych = array('Educational Psychology',
'Department of Educational and Counselling Psychology',
'Education(Psychology and Technology)',
'Education and Human Development',
'Educational Psychology &amp; Learning Systems',

'Counseling, Adult and Higher Education',
'Counseling and Personnel Services',
'Leadership and Counseling',
'Educational Psychology and Educational Technology',
'School of Psychology',
'Counseling and Educational Psychology',
'School Psychology',
'Counseling, School and Educational Psychology',
'Department of Psychology',
'Counselor Education and Supervision (ExCES)',
'Educational Psychology',
'Counselor Education',
'Instructional Psychology and Technology',
'Counselor Education');

Psych
$arr_psych = array('Psychology',
'The School of Psychology',
'Global Psychology with a concentration in Transpersonal Psychology',
'Psychology',
'Clinical Psychology',
'Cognitive Science',
'Graduate School of Applied and Professional Psychology',
'Psychological & Quantitative Foundations',
'Psychological &amp; Quantitative Foundations',
'Cognitive &amp; Neural Systems GRS');

Computer Science
$arr_compsci = array('Computer Science',
'Computing Science',
'Electrical and Computer Engineering',
'Electrical Engineering 0303',
'Computer Science and Engineering',
'Electrical Engineering (Signal and Image Proc)',
'Electrical Engineering',
'Engineering Sciences (Mechanical Engineering)',
'Electrical Engineering: Doctor of Philosophy',
'Engineering Mgt and Systems Engineering',
'Department of Computer Science',
'Industrial Engineering',
'Electrical Engineering &amp; Computer Science',
'Computer Engineering',
'Engineering',
'Computer Science and Information Systems',
'Electrical Engineering',
'School of Electrical and Computer Engineering',

'Department of Computer and Information Science',
'Computer and Information Science',
'School of Computing',
'Engineering and Technology Education',
'Computation and Neural Systems',
'Informatics-Information Science',
'Information Science',
'Information Systems',
'Informatique / Computer Science',
'Robotics',
'Computational Sciences and Informatics',
'Applied Mathematics and Computational Science');

Ag
$arr_ag = array('Agricultural and Biosystems Engineering',
'Agricultural Education and Studies',
'Agricultural Information Science and Education',
'Agricultural Economics',
'Agricultural Education');

Language
$arr_language = array('Linguistics',
'English',
'Language and Literacy Education',
'Rhetoric and Writing',
'Rhetoric and Scientific and Technical Communication');

Sciences
$arr_science = array('Neuroscience',
'Applied Cognition and Neuroscience',
'Department of Biology',
'Chemistry',
'Large Animal Clinical Sciences',
'Engineering and Applied Sciences',
'Evaluative Clinical Sciences',
'Physics - Doctor of Philosophy',
'Plant and Soil Sciences',
'Chemistry and Biochemistry');

Communication
$arr_communication = array('Comm. Media',
'Communication Science and Disorders',
'Communication',
'Communication &amp; Information Sciences',
'Communication & Information Sciences',
'Communication and Cognitive Science',

'Communication and Information Sciences');

Music
$arr_music = array('Music Education',
'Music',
'Music Education (Music)',
'School of Music, Theatre, and Dance: Music',
'Music Division');

Misc
$arr_misc = array('Aviation &amp; Space Science',
'Aeronautics and Astronautics',
'Mathematical Sciences',
'Mathematics',
'Engineering Education',
'Environmental Science &amp; Management',
'Forestry',
'Geography',
'Economics',
'Human Services',
'School of Public Service Leadership',
'Doctor of Ministry Department',
'Public Service Leadership',
'Department of Individual and Family Studies',
'Family Studies',
'Nutritional Sciences',
'Human Development',
'Harold Abel School of Social and Behavioral Sciences',
'Sociology',
'Clinical Research',
'Art',
'Germanic Studies',
'Arts and Humanities',
'Modern Languages',
'Literacy and Culture',
'International and Transcultural Studies',
'Liberal Studies',
'Apparel, Educational Studies and Hospitality Management',
'Department of Hospitality Management and Dietetics',
'Atlas',
'Graduate School - New Brunswick',
'Interdisciplinary',
'Advanced Studies',
'College of Doctoral Studies',
'Doctoral Studies',
'Graduate School',

'Doctoral Leadership Studies',
'Interdisciplinary Studies',
'School Improvement',
'Research and Advanced Studies',
'Technology',
'School of Church Ministries',
'School of Missions and Evangelism');

APPENDIX 6

PQDT CLASSIFICATION FRAMEWORK

Arts, Business, Education, Humanities, and Social Sciences

AREA, ETHNIC, AND GENDER STUDIES
African American studies 0296 African studies 0293 American studies 0323 Asian American studies 0343 Asian studies 0342 Baltic studies 0361 Black studies 0325 Canadian studies 0385 Caribbean studies 0432 Classical studies 0434 East European studies 0437 Ethnic studies 0631 European studies 0440 French Canadian culture 0482 Gender studies 0733 GLBT studies 0492 Hispanic American studies 0737 Holocaust studies 0507 Islamic culture 0512 Judaic studies 0751 Latin American studies 0550 Middle Eastern studies 0555 Native American studies 0740 Near Eastern studies 0559 North African studies 0560 Pacific Rim studies 0561 Regional studies 0604 Scandinavian studies 0613 Slavic studies 0614 South African studies 0654 South Asian studies 0638 Sub Saharan Africa studies 0639 Women's studies 0453

BUSINESS
Accounting 0272 Arts management 0424 Banking 0770 Business 0310 Entrepreneurship 0429 Finance 0508 Management 0454 Marketing 0338 Sports management 0430

COMMUNICATIONS AND INFORMATION SCIENCES Communication 0459 Information science 0723 Journalism 0391 Library science 0399 Mass communication 0708 Technical communication 0643 Web studies 0646

FINE AND PERFORMING ARTS
Art criticism 0365 Art history 0377 Cinematography 0435 Dance 0378 Design 0389 Film studies 0900 Fine arts 0357 Music 0413 Performing arts 0641 Theater 0465 Theater history 0644

EDUCATION
Adult education 0516 Art education 0273 Bilingual education 0282 Business education 0688 Community college education 0275 Continuing education 0651 Curriculum development 0727 Early childhood education 0518 Education 0515 Education finance 0277 Education policy 0458 Educational administration 0514 Educational evaluation 0443 Educational leadership 0449 Educational psychology 0525 Educational technology 0710 Educational tests & measurements 0288 Elementary education 0524 English as a second language 0441 Foreign language instruction 0444 Gifted education 0445 Health education 0680 Higher education 0745 Higher education administration 0446 History of education 0520 Home economics education 0278 Industrial arts education 0521 Instructional design 0447 Language arts 0279 Mathematics education 0280 Middle school education 0450 Multicultural education 0455 Music education 0522 Pedagogy 0456 Performing arts education 0457 Philosophy of education 0998 Physical education 0523 Reading instruction 0535 Religious education 0527 School counseling 0519 Science education 0714 Secondary education 0533 Social sciences education 0534

Sociology of education 0340 Special education 0529 Teacher education 0530 Vocational education 0747

HUMANITIES HISTORY
African history 0331 American history 0337 Ancient history 0579 Asian history 0332 Black history 0328 Canadian history 0334 European history 0335 History 0578 History of Oceania 0504 History of science 0585 Latin American history 0336 Medieval history 0581 Middle Eastern history 0333 Military history 0772 Modern history 0582 Russian history 0724 World history 0506

LANGUAGE & LITERATURE
African literature 0316 American literature 0591 Ancient languages 0289 Asian literature 0305 British and Irish literature 0593 Canadian literature 0352 Caribbean literature 0360 Classical literature 0294 Comparative literature 0295 English literature 0593 French Canadian literature 0355 Germanic literature 0311 Icelandic & Scandinavian literature 0362 Language 0679 Latin American literature 0312 Linguistics 0290 Literature 0401 Literature of Oceania 0356 Medieval literature 0297 Middle Eastern literature 0315 Modern language 0291 Modern literature 0298 Rhetoric 0681 Romance literature 0313 Slavic literature 0314

PHILOSOPHY AND RELIGION
Aesthetics 0650 Biblical studies 0321 Canon law 0375 Clerical studies 0319 Comparative religion 0618 Divinity 0376 Epistemology 0393 Ethics 0394 Logic 0395 Metaphysics 0396 Pastoral counseling 0397 Philosophy 0422 Philosophy of Religion 0322 Philosophy of science 0402 Religion 0318 Religious history 0320 Spirituality 0647 Theology 0469

LAW AND LEGAL STUDIES
Alternative dispute resolution 0649 Intellectual property 0513 International law 0616 Law 0398 Patent law 0562

SOCIAL SCIENCES
Archaeology 0324 Area planning and development 0341 Criminology 0627 Cultural anthropology 0326 Demography 0938 Economic history 0509 Economic theory 0511 Economics 0501 Economics, Commerce-Business 0505 Economics, Labor 0510 Folklore 0358 Forensic anthropology 0339 Geography 0366 Individual & family studies 0628 International relations 0601 Labor relations 0629 Military studies 0750 Organization theory 0635 Organizational behavior 0703 Peace studies 0563 Physical anthropology 0327 Political Science 0615 Public administration 0617 Public policy 0630 Recreation and tourism 0814 Social research 0344 Social structure 0700 Social work 0452 Sociolinguistics 0636 Sociology 0626 Transportation planning 0709 Urban planning 0999

INTERDISCIPLINARY
Alternative energy 0363 Biographies 0304 Climate change 0404 Cultural resources management 0436 Energy 0791 Food science 0359 Home economics 0386 Information

technology 0489 Multimedia 0558 Museum studies 0730 Sustainability 0640 Textile research 0994 Wood sciences 0746

Behavioral, Natural, and Physical Sciences

AGRICULTURE
Agriculture 0473 Agronomy 0285 Animal diseases 0476 Animal sciences 0475 Fisheries and aquatic sciences 0792 Forestry 0478 Horticulture 0471 Plant pathology 0480 Plant sciences 0479 Range management 0777 Soil sciences 0481 Urban forestry 0281 Wildlife management 0286

ARCHITECTURE
Architecture 0729 Architectural engineering 0462 Landscape architecture 0390

BEHAVIORAL SCIENCES
Animal behavior 0602 Behavioral sciences 0384 Clinical psychology 0622 Cognitive psychology 0633 Counseling psychology 0603 Developmental psychology 0620 Experimental psychology 0623 Occupational psychology 0624 Personality psychology 0625 Physiological psychology 0989 Psychobiology 0349 Psychology 0621 Quantitative psychology and psychometrics 0632 Social psychology 0451

BIOLOGICAL SCIENCES
Biochemistry 0487 Bioinformatics 0715 Biology 0306 Biomechanics 0648 Biophysics 0786 Biostatistics 0308 Cellular biology 0379 Developmental biology 0758 Endocrinology 0409 Entomology 0353 Evolution & development 0412 Genetics 0369 Histology 0414 Limnology 0793 Microbiology 0410 Molecular biology 0307 Morphology 0287 Neurosciences 0317 Parasitology 0718 Physiology 0719 Plant biology 0309 Systematic biology 0423 Virology 0720 Zoology 0472

ECOSYSTEM SCIENCES
Ecology 0329 Macroecology 0420 Paleoecology 0426

ENGINEERING
Aerospace engineering 0538 Artificial intelligence 0800 Automotive engineering 0540 Biomedical engineering 0541 Chemical engineering 0542 Civil engineering 0543 Computer engineering 0464 Computer science 0984 Electrical engineering 0544 Engineering 0537 Geological engineering 0466 Geophysical engineering 0467 Geotechnology 0428 Industrial engineering 0546 Mechanical engineering 0548 Mining engineering 0551 Naval engineering 0468 Nanotechnology 0652 Nuclear engineering 0552 Ocean engineering 0547 Operations research 0796 Packaging 0549 Petroleum engineering 0765 Plastics 0795 Robotics 0771 System science 0790

ENVIRONMENTAL SCIENCES
Conservation biology 0408 Environmental economics 0438 Environmental education 0442 Environmental engineering 0775 Environmental geology 0407 Environmental health 0470 Environmental justice 0619 Environmental law 0439 Environmental

management 0474 Environmental philosophy 0392 Environmental science 0768
Environmental studies 0477 Land use planning 0536 Natural resource management 0528
Water resources management 0595 Wildlife conservation 0284

GEOSCIENCES
Aeronomy 0367 Atmospheric chemistry 0371 Atmospheric sciences 0725
Biogeochemistry 0425 Biological oceanography 0416 Chemical oceanography 0403
Continental dynamics 0406 Geobiology 0483 Geochemistry 0996 Geographic
information science
and geodesy 0370 Geology 0372 Geomorphology 0484 Geophysics 0373 Hydrologic
sciences 0388 Marine geology 0556 Meteorology 0557 Mineralogy 0411 Paleoclimate
science 0653 Paleontology 0418 Petroleum geology 0583 Petrology 0584 Physical
geography 0368 Physical oceanography 0415 Planetology 0590 Plate tectonics 0592
Remote sensing 0799 Sedimentary geology 0594

HEALTH AND MEDICAL SCIENCES
Aging 0493 Alternative medicine 0496 Audiology 0300 Dentistry 0567 Epidemiology
0766 Gerontology 0351 Health care management 0769 Health sciences 0566
Immunology 0982 Kinesiology 0575 Medical ethics 0497 Medical imaging and radiology
0574 Medicine 0564 Mental health 0347 Nursing 0569 Nutrition 0570 Obstetrics and
gynecology 0380 Occupational health 0354 Occupational therapy 0498 Oncology 0992
Ophthalmology 0381 Osteopathic medicine 0499 Pathology 0571 Pharmaceutical
sciences 0572 Pharmacology 0419 Physical therapy 0382 Public health 0573 Public
health occupations
education 0500 Speech therapy 0460 Surgery 0576 Toxicology 0383 Veterinary medicine
0778

MATHEMATICAL AND PHYSICAL SCIENCES
Acoustics 0986 Analytical chemistry 0486 Applied mathematics 0364 Astronomy 0606
Astrophysics 0596 Atomic physics 0748 Chemistry 0485 Condensed matter physics 0611
Electromagnetics 0607 High temperature physics 0597 Inorganic chemistry 0488 Low
temperature physics 0598 Materials science 0794 Mathematics 0405 Mechanics 0346
Molecular chemistry 0431 Molecular physics 0609 Nanoscience 0565 Nuclear chemistry
0738 Nuclear physics 0756 Optics 0752 Organic chemistry 0490 Particle physics 0798
Physical chemistry 0494 Physics 0605 Plasma physics 0759 Polymer chemistry 0495
Quantum physics 0599 Statistics 0463 Theoretical mathematics 0642 Theoretical physics
0753

APPENDIX 7

EMERGENT KEYWORD NETWORK: REMOVED WORDS

```
$arr = array('DISTANC EDUC',        'DEVELOP',
'ONLIN COURS',                       'INSTITUT',
'DISTANC',                           'ANALYSI',
'DISTANC EDUC PROGRAM',              'ADDIT',
'ONLIN',                             'FACTOR',
'EDUC',                              'IMPLEMENT',
'DISTANC LEARN',                     'METHOD',
'ONLIN LEARN',                       'RESEARCH QUESTION',
'ONLIN LEARN ENVIRON',               'RECOMMEND',
                                     'QUESTIONNAIR',
'STUDI',                             'IMPACT',
'RESEARCH',                          'CONCLUS',
'RESULT',                            'ORDER',
'DATA',                              'IMPORT',
'PURPOS',                            'MEAN',
'USE',                               'CONCEPT',
'FIND',                              'SAMPL',
'EFFECT',                            'ASS',
'SIGNIFIC DIFFER',                   'GOAL',
'DIFFER',                            'SURVEY',
'DEVELOP',                           'SUBJECT',
'LEVEL',                             'POPUL',
'INFORM',                            'FUTUR RESEARCH',
'INTERVIEW',                         'DATA ANALYSI',
'MODEL',                             'MEASUR',
'TRAIN',                             'EL',
'PRACTIC',                           'SEN',
                                     'FURTHERMOR');
```