March – 2010

An Investigation of Distance Education in North American Research Literature Using Co-word Analysis

Albert D. Ritzhaupt
University of North Carolina, Wilmington, USA

Michelle Stewart, **Patryce Smith**, and **Ann E. Barron**
University of South Florida, USA

Abstract

The field of distance education is composed of a multiplicity of topics leading to a vast array of research literature. However, the research does not provide a chronological picture of the topics it addresses, making it difficult to develop an overview of the evolution and trends in the literature. To address this issue, a co-word analysis was performed on the abstracts of research articles found in two prominent North American research journals (*N* = 517), the *American Journal of Distance Education* and the *Journal of Distance Education*, between 1987 and 2005. The analysis yielded underlying trends and themes for three different periods (pre-Web, emerging Web, and maturing Web). Additionally, similarity index analyses were conducted across time periods. The pre-Web era was characterized by the need for quality and development. The emerging Web era was characterized by the development of theory. The maturing Web era was characterized by interaction and the use of tools for communication. The results demonstrate that the North American distance education research literature is characterized by having few consistent and focused lines of inquiry. Conclusions are provided.

Keywords: Distance education; co-word analysis; similarity index analysis; themes; trends

Distance Education Research

In fewer than 20 years, the field of distance education has grown from correspondence courses to podcasts and virtual classrooms. The definition of distance education is a complicated phenomenon, drawing from many factors and perspectives. Keegan's (1996) definition of distance education is frequently cited in the research literature and involves the description of five qualities that distinguish distance education from other forms of instruction: a) the quasi-permanent separation of teacher and learner, b) the influence of an educational organization, c) the use of technical media, d) the provision of two-way communication, and e) the quasi-

permanent absence of learning groups. Related research literature is being produced at a very rapid rate, yet the empirical characterization of the field and its evolution is lacking. The purpose of this article is to provide a synopsis of previous meta-analytics studies conducted in the field of distance education and then to employ objective methods called *co-word analysis* and *similarity index analysis* to investigate trends and themes in the field from 1987 through 2005.

Distance Education Synthesis

Because the term distance education is overarching and encompasses a vast array of topics and literature, researchers in the field face a difficult challenge in attempting to generalize findings and inform practice (Salas, Kosarzycki, Burke, Fiore, & Stone, 2002). There have been several attempts to generalize findings in distance education literature using well-established techniques, which include quantitative meta-analysis, content analysis, and literature synthesis. These studies have attempted to capture various aspects of distance education. Table 1 summarizes the researchers, their methods, the sources of their literature, the number of articles, and the period under investigation. Each of these articles has its strengths and contributions.

Table 1

Previous Distance Education Synthesis Articles

Researchers	Method	Source	N	Period
Koble and Bunker, 1997	Mixed method using a content analysis technique, Porter's forum analysis, (1992), and synthesis, focusing on authors, readership, and classification using the International Centre for Distance Learning as classification system	*American Journal of Distance Education*	129 articles	1987-1995
Berge and Mrozowski, 2001	Content analysis technique using Sherry, Fulford and Zhang's (1998) classification system, focused on learner characteristics and needs, media influence on the instructional process, access issues and stakeholder roles, and research methodology (descriptive, co relational, experimental, case study)	*American Journal of Distance Education* *Distance Education* *Journal of Distance Education* *Open Learning* Dissertation Abstracts	1,419 articles and dissertations	1990-1999

Rourke and Szabo, 2002	Content analysis technique on article topic, research method, type, and biographical information about first authors.	*Journal of Distance Education*	235 articles	1986-2000
Lee, Driscoll, and Nelson 2004.	Content analysis of journal articles using Sherry (1995), Phipps, and Merisotis (1999) and Khan's (1997) classification for topics, methods, and citations	*American Journal of Distance Education Journal of Distance Education Distance Education Open Learning*	383 articles	1997 - 2002
Bernard et. al., 2004.	Quantitative meta analysis and synthesis of research comparing distance education to traditional classroom achievement, attitude, and retention	Included numerous journals and proceedings using comprehensive inclusion and exclusion criteria	232 articles and proceedings	1985-2002
Zhao, Lei, Yan, and Tan, 2005	Quantitative meta analysis and synthesis of research related to factors affecting "effective" distance education	Included numerous journals using comprehensive inclusion and exclusion criteria	51 articles	1987-1995
Zawacki-Richter, Bäcker, and Vogt 2009	Content analysis method on research areas, methods, and authorship patterns	*American Journal of Distance Education Journal of Distance Education Distance Education Open Learning International Review of Research in Open and Distance Learning*	695 articles	2000-2008

Koble and Bunker (1997) utilized a content analysis technique in concert with Porter's forum analysis (1992) to analyze the *American Journal of Distance Education* from 1987 to 1995 (129 articles). They chose the International Centre for Distance Learning's classification system (ICDL, 2006) to analyze the articles in this time period. Using this classification scheme, the articles were classified as the following: theory, policy and development (33%), media and delivery systems (31%), institution, staff, and management (16%), student psychology, motivation, and characteristics (15%), faculty participation and instructional process (11%), course design and curriculum development (10%), and student administration and support (3%) (Koble & Bunker, 1997). An analysis of the authorship showed that 70% of the contributors were from the United States, 20% from Canada, and the remaining were undisclosed (Koble & Bunker, 1997). Their results also concluded that many articles "continue to appear providing more evidence of the 'effectiveness' of distance education in different contexts" and that there is a

"move away from an early emphasis on correspondence study to one which includes telecommunications technologies and discussions of interaction …" (Koble & Bunker, 1997, p. 17). The researchers also demonstrated growth in the area of faculty issues, which they attributed to the changing roles in distance education environments (Koble & Bunker, 1997).

Berge and Mrozowski (2001) reviewed 1,419 articles from four major distance education journals and dissertation abstracts with a distance education classification. Their research used Sherry's (1998) classification system to categorize each of the articles on content and research method classification. Using a synthesis methodology, their research findings showed that the majority (76%) of the published articles in distance education used a descriptive research methodology, followed by case study (9%), correlational (8%), and experimental (7%) (Berge & Mrozowski, 2001). The researchers also concluded that most of the published literature addressed design issues, followed by learner characteristics, and strategies to increase learner interactivity and active learning. They also noted gaps in the research relating to an underemphasis on total academic program outcomes, limited explanations of high dropout rates, a focus on individual technologies rather than multiple technologies, and a lack of literature addressing the effectiveness of digital libraries (Berge & Mrozowski, 2001).

Rourke and Szabo (2002) used a traditional content analysis technique to classify articles from 1986 to 2000 from the *Journal of Distance Education* on topic, research method, type, and biographical information about first authors. Their findings showed that approximately 70% of the articles published in this period were empirical (data collection and inference), descriptive, publication reviews, and perspective papers. Unlike Koble and Bunker, Rourke and Szabo found that the representation of gender was nearly equal between male (45%) and female (43%), with the rest unidentified. The *Journal of Distance Education* is resident to Canada; 52% of the first authors were from Canada, and the remaining authorship was from primarily English-speaking countries. Seventy-two percent of the authors were affiliated with institutions of higher education, and the remaining authors were associated with the Canadian Association of Distance Education (Rourke & Szabo, 2002).

A comprehensive meta-analysis of 232 articles and proceedings between 1985 and 2002 compared distance education with traditional classroom environments on independent achievement, attitude, and retention (Bernard, Abrami, Lou, Borokhovski, Wade, Wozney, Wallet, Fiset, & Huang, 2004). They concluded that much distance education research was of low quality, "particularly in terms of internal validity" (Bernard et al., 2004, p. 416). "Overall results indicated effect sizes of essentially zero on all three measures and wide variability. This suggests that many applications of distance education outperform their classroom counterparts and that many perform more poorly" (Bernard et al., 2004, p. 379).

Lee, Driscoll, and Nelson (2004) examined 383 journal articles published in the *American Journal of Distance Education*, *Distance Education*, *Journal of Distance Education,* and *Open Learning* from 1997 to 2002. Results from the analysis of the research methods revealed that "78 (20%) of the 383 articles published in the four journals described theoretical inquiry, 47 (12%) were experimental research, 138 (36%) were case studies, 23 (6%) were evaluation research, 26

(7%) were developmental research, 48 (13%) were survey research, and 23 (6%) were implemented with combined inquiries" (Lee, Driscoll, & Nelson, 2004, p. 231). They concluded that distance education has progressed as a field because of an emergence of new media technologies and suggest that "the research methodology and the research paradigm for distance education are important because they lead empirical investigations of a theory to draw on other theory that may be more suitable for a better explanation of distance education approaches" (Lee, Driscoll, & Nelson, 2004, p. 239)

A study by Zhao, Lei, Yan, and Tan (2005) sought to discover factors that affect "effective" distance education by applying meta-analysis to 51 articles meeting their selection criteria. The empirical findings varied greatly and were not found to be significantly different from traditional instructional methods. However, their conclusions were that not all implementations of distance education were "created equal," which may be attributed to the small effect sizes. Their findings also provide evidence that not all content is suitable for distance education, human interaction is critical, and some learners may not be able to benefit from the delivery method (Zhao et al., 2005).

A more recent study conducted by Zawacki-Richter, Bäcker, and Vogt (2009) examined five major journals, including the *American Journal of Distance Education*, *Journal of Distance Education*, *Distance Education*, *Open Learning*, and the *International Review of Research in Open and Distance Learning*. In total, 695 unique journal articles were examined. Their findings showed that the most frequently studied research areas included interaction and communication in online learning communities (17.6%), instructional design (17.4%), and learner characteristics (16.3%). In terms of research methods, they coded quantitative, qualitative, triangulation, and other and found that in 2007 and 2008 "other" (mixed methods) was the most frequently used methodology, followed by quantitative.

Purpose of the Study

Previous research studies have attempted to generalize findings and inform practice, but distance education research falls behind because of rapid innovations in technology and instructional practice (Lee, Driscoll, & Nelson, 2004). Consequently, many unanswered questions remain. How have the major tenets changed over time? Most previous investigations have focused on a few aspects of the field (e.g., research methodology) or domains (e.g., peer collaboration), using a priori sets of themes and subjective interpretation. In classifying research articles to a specific category or theme, a tremendous amount of information may have been lost. The purpose of this research was to employ an objective investigation, using co-word analysis. Specific research questions include the following:

1. What themes can be identified over the years in the abstracts published in major North American distance education research journals?
2. What are the relationships among the themes identified in abstracts published in major North American distance education research journals?

3. How have the themes identified in the abstracts published in major North American distance education research journals changed over time?

Co-Word Analysis

This research study attempts to examine distance education research by implementing a co-word analysis methodology to identify themes, trends, and structural characteristics in North American distance education literature. Co-word analysis is an automated content analysis technique that is effective in mapping the strength of relationships among textual data. It employs a graphical modeling technique that is similar to association analysis (Edwards, 1995; Kaufman & Rousseeuw, 1990). Co-word analysis identifies co-occurrence strengths of terms and creates a set of lexical graphs that effectively illustrate the strongest associations between various terms (Coulter, Monarch, & Konda, 1998; Whittaker, 1989). "In contrast to most other types of statistical graphics, the graphs do not display data, but rather an interpretation of the data, in the form of a model" (Edwards, 1995, p. 146).

As noted by Ding, Chowdhury, and Foo (2001), co-word analysis reduces data into detailed visual representations with the essential information contained in the data. It is based on the notion that words are the important carriers of scientific concepts, ideas, and knowledge (van Raan & Tijssen, 1993). Unlike previous meta-analytic methods, co-word analysis allows the primary themes to emerge from the research literature. Thus, the research literature allows the words of a discipline to describe the themes relevant to a domain.

In previous studies, this method has been applied to software engineering, polymer chemistry research and technology literature, information retrieval research literature, scientometrics, neural networks, and information systems education literature (Callon, Courtial, & Laville, 1991; van Raan & Tijssen, 1993; Courtial, 1994; Coulter et al., 1998; Ding, Chowdhury, & Foo, 2001; Ritzhaupt, 2003). The results of previous studies vary, but co-word analysis has proven to be a useful method for identifying research themes and trends. Carnegie Mellon University's Software Engineering Institute developed the software, Co-Word Analysis Information Retrieval, or CAIR, which was employed in this study.

Time Periods and Corpus

To obtain a comprehensive corpus of textual information, two of the leading journals in the North American field of distance education were selected for inclusion in this study: the *American Journal of Distance Education* (AJDE) and *Journal of Distance Education* (JDE). The *American Journal of Distance Education* is "the internationally recognized journal of research and scholarship in the field of American distance education" (p.1, 2008). Available since 1987, AJDE is currently published by Rogers and Taylor. The *Journal of Distance Education* is "an international publication of the Canadian Network for Innovation in Education" (p. 1, 2008). JDE has been published since 1986. Both journals were selected because they had similar publication inception dates (1987 and 1986), are published in North America, and are leading publication venues in the field of distance education. North American journals were selected

because the co-word method relies heavily on the consistency of the language.

Though somewhat arbitrary, a retrospective look at the last 20 years of distance education, from the inception of the World Wide Web (Web) until the present, reveals three, relatively distinct time periods (pre-Web, emerging Web, and maturing Web). The period from 1987 until 1993 marks the era in which the Web was not yet significant to distance education research; the first graphical user interface (Mosaic) was not developed until 1993 (NCSA, 2008). The emergence of the Web and its initial impact on distance education occurred from approximately 1994 to 1999 (Reiser, Dear, & Edge, 2001). It was during this time period that tools such as WebCT and FrontPage became available (Chan, 2005). The maturation of distance education and the Web since 2000 marks another significant period (until 2005, the endpoint of this study's look at the era).

Term Extraction

Prior to using co-word analysis, the data (text retrieved from the abstracts of the articles) required preparation. Some articles were written in French (JDE is resident to Canada) and had to be translated before being included in the analysis. The abstracts from the journal articles were used in the analysis, each serving as a separate unit of analysis. Extraneous terms, such as *this paper*, were removed from the analysis, and others were modified to accommodate synonyms, variations in language, and acronyms. For example, it is common in distance education literature to refer to *computer-mediated communication* as CMC.

The corpus was organized by placing the abstracts for each period in separate text files. Table 2 summarizes the number of documents by journal and time period, the average number of words per unit, and the number of terms extracted for each time period by abstract. A total of 517 articles are included in the analysis.

Table 2

Articles by Journal and Time Period, Average Words per Unit, and Number of Terms Extracted by Unit and Period

	AJDE	JDE	*Abstracts	#Abstracts
1987-1993	108	115	84.97	231
1994-1999	74	85	136.39	242
2000-2005	75	60	115.56	205

*Average number of words per unit.

#Terms extracted with occurrence level of 5 for abstracts

Co-Word Analysis Algorithm

The co-word analysis algorithm uses *Strength (S_{ij})* to measure the association between the terms *i* and *j,* with the following equation (Coulter et al., 1998, p. 1209):

$$S_{ij} = \frac{c_{ij}^2}{c_i \bullet c_j}, 0 \le S \le 1$$

Let c_i be the number of times the term *i* occurs in a corpus and c_j be the number of times the term *j* occurs in another corpus then let c_{ij} be the number of co-occurrences of the terms *i* and *j*. The result is S_{ij} between the range of zero to one, inclusive, which can be used to measure the strength of the association of two terms co-occurring across a number of units.

The co-word algorithm incorporates two passes on the corpus to produce pair-wise connection strengths between related terms. The pass-1 terms are the primary associations. The outcome is a set of lexical maps with interacting terms based on the strengths of relationships. The first pass uses a breadth-first approach: the ensuing terms are considered internal terms, and the links are considered internal links (Ritzhaupt, 2003). The pass-1 terms are depicted by bolded rectangles, and the pass-1 links are depicted by bolded links (see Figure 1).

The second pass is concerned with the relationships between lexical maps generated in pass-1. To be a candidate for inclusion in pass-2, both terms must be in some pass-1 lexical map (Coulter, 1998). Thus, all the terms used in pass-1 are reused in pass-2. A pass-2 link connects a pass-1 term occurring in one lexical map to another pass-1 term in another lexical map (Ritzhaupt, 2003). Thus, the second pass generates the pervasive relationships. Pass-2 items are represented with non-bolded lines and rectangles (see Figure 1).

The co-word algorithm allows the researcher to select two parameters: the number of pass-1 terms that can exist in a single lexical map and the minimum co-occurrence. "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). Therefore, it was determined through trial and error, the terms must co-occur in at least five separate abstracts to yield the most comprehensive results for these data.

Results and Discussion

Co-word analysis provides several different results, including the following: (a) lexical maps, which highlight the co-occurrences among terms; (b) cohesion and coupling graphs, which display the internal strength and interactivity of maps; and (c) super networks, which visualize pair-wise relationships between lexical maps.

Lexical Maps (Themes)

The analysis of the 1987 to 1993, 1994 to 1999, and 2000 to 2005 abstracts generated 9, 11, and 6 lexical maps, respectively. The lexical map names assigned to the abstracts by each time period are shown in Table 3. The maps in Table 3 also illustrate the number of terms (N) and links (L) within each lexical map. This analysis of the lexical maps addresses the first research question, "What themes can be identified over the years in the abstracts published in major North American distance education research journals?"

Substantially fewer maps were generated in the subsequent two time periods. The marked decrease can be explained in that the average number of words per abstract and the number of articles during 1987-1993 was substantially higher than the following time periods. The growth in distance education research over the past 20 years has spawned more specific publication venues, such as the *Journal of Asynchronous Learning Networks* and the *Journal of Online Learning and Teaching*. In addition, the AJDE and JDE have raised their quality requirements since the first time period.

Table 3

Lexical Maps Generated for Abstracts by Time Period

1987-1993 Abstracts	N	L	1994-1999 Abstracts	N	L	2000-2005 Abstracts	N	L
87-93-Abs1-Distance Education-Course-Student	20	40	94-99-Abs1-Study-Distance Education-Learning	20	45	00-05-Abs1- Study-Distance-Student	20	39
87-93-Abs2-Need-Development-Quality	20	25	94-99-Abs2-Interview-Method-Process	20	26	00-05-Abs2-Activity- Research-Practice	20	23
87-93-Abs3-Program-Method-University	20	38	94-99-Abs3-Need-Education-Role	19	27	00-05-Abs3-Interaction-Communication-Tool	20	27
87-93-Abs4-Relationship-Model-Process	20	26	94-99-Abs4-Development-Theory-Information	20	38	00-05-Abs4-Teacher-Experience-Results	20	33
87-93-Abs5-Concept-Term	10	13	94-99-Abs5-Goal-Value	12	15	00-05-Abs5-Community-Development-Evaluation	20	27

1987-1993 Abstracts	N	L	1994-1999 Abstracts	N	L	2000-2005 Abstracts	N	L
87-93-Abs6-Difference-Level-Instruction	18	51	94-99-Abs6-Content-Time-Computer	20	24	00-05-Abs6-Way-Type-Information	20	22
87-93-Abs7-Teaching-Education-Factor	20	24	94-99-Abs7-Research-Change-Purpose	20	25			
87-93-Abs8-Role-Field	12	15	94-99-Abs8-Resource-Teaching-School	20	33			
87-93-Abs9-Evaluation-Problem-Design	20	30	94-99-Abs9-Program-Knowledge-Skill	20	27			
			94-99-Abs10-Way-Instruction-Interaction	20	29			
			94-99-Abs11-Environment-Author-Model	20	33			

*N= Number of terms, L=Number of links

Each lexical map was named using a systematic naming algorithm involving the lexical map's central terms. The maps' names consist of the time period (e.g., "87-93") followed by the corpus and map number (e.g., "Abs2" being the second map generated from the abstract corpus), followed by the three primary terms from the lexical map. The primary terms were identified by the pass-1 term with the maximum number of pass-1 connections appearing first, followed by the second, and the third. In the event of a tie, the number of pass-2 links connecting the terms was used to select the term. If a lexical map only had two pass-1 terms, these were the only terms used in the naming convention. Thus, the lexical map names should be viewed as independent themes that emerged from the research literature. In the following sections, the ability of co-word analyses to show fading, evolving, and emerging themes in the lexical maps is illustrated.

Teleconference: A fading communication technology.

One powerful aspect of co-word analysis is its inherent ability to highlight themes and terms that have faded from one time period to the next. "Distance education that emerged in the United States in the 1980's was based on the technologies of teleconferencing…" (Moore & Kearsley, 2005, p. 38). The term *teleconference* occurred in the 1987 to 1993 time period but failed to emerge as a term as part of a theme in the subsequent time periods. This is a relatively clear indication that the use of *teleconferencing* is fading from research literature in distance education. As shown in Figure 1, *teleconference* is a pass-1 term 87-93-Abs4-Relationship-Model-Process

lexical map. Keegan's (1986) definition of distance education includes these qualities: the use of technical media and the provision of two-way communication. The map illustrates *communication* as a pass-1 term and *teleconferencing* as a medium. Though teleconferencing may be fading from research literature, this does not indicate whether it is fading as a modality in practice.
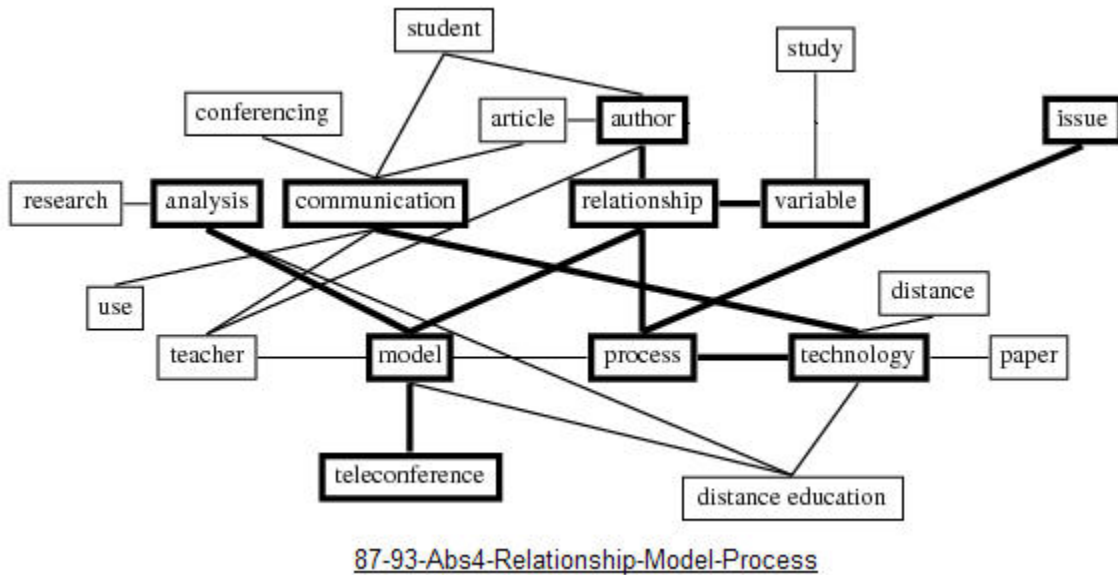

87-93-Abs4-Relationship-Model-Process

Figure 1. 87-93-Abs4-Relationship-Model-Process lexical map.

Interaction: An emerging and evolving theme in distance education.

Co-word analysis also helps illustrate themes that emerge and swiftly evolve over time. *Interaction* first emerged as a theme (94-99-Abs10-Way-Instruction-Interaction) in the second time period. During this time period, there is a strong connection between *interaction* and *instruction* and a slightly weaker connection between *interaction* and *communication,* as shown in Figure 2.

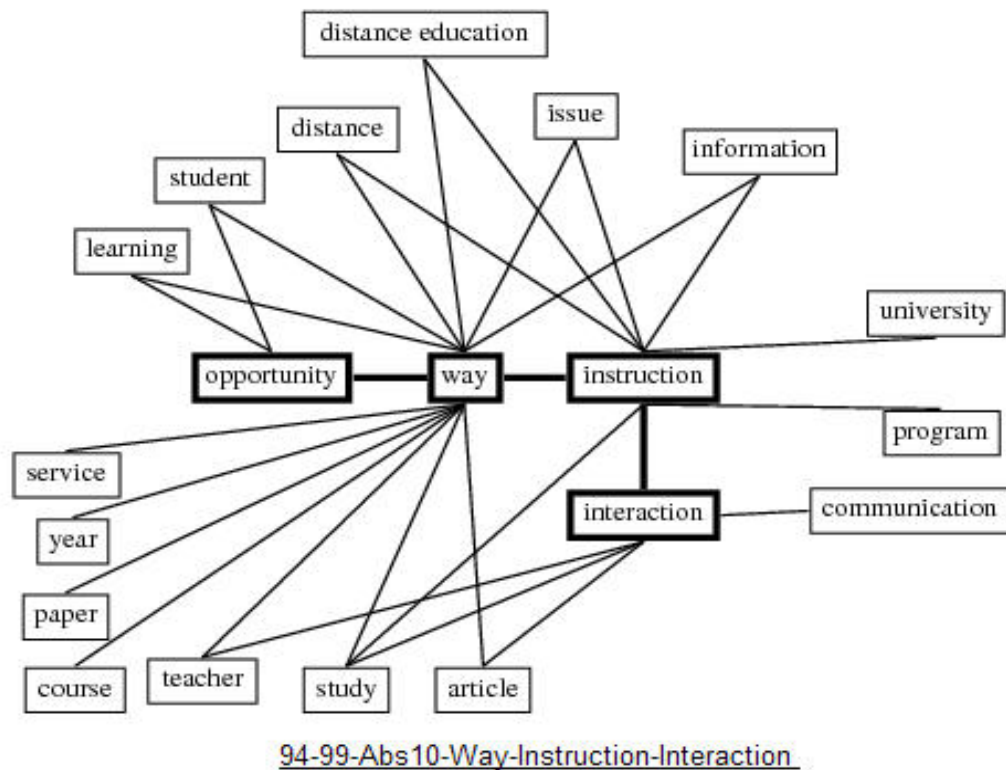94-99-Abs10-Way-Instruction-Interaction

Figure 2. 94-99-Abs11-Way-Instruction-Interaction lexical map.

In the subsequent time period, the lexical map 00-05-Abs3-Interaction-Communication-Tool shows that the connection between *interaction* and *communication* strengthens (see Figure 3), where both *communication* and *interaction* are pass-1 terms. The 2000 to 2005 time period shows that *interaction* also interacts with other relevant terms: *strategy*, *learning environment*, and *tool*. In this time period, *interaction* appears to have escalated as a key term in distance education research literature and has clearly evolved between time periods.
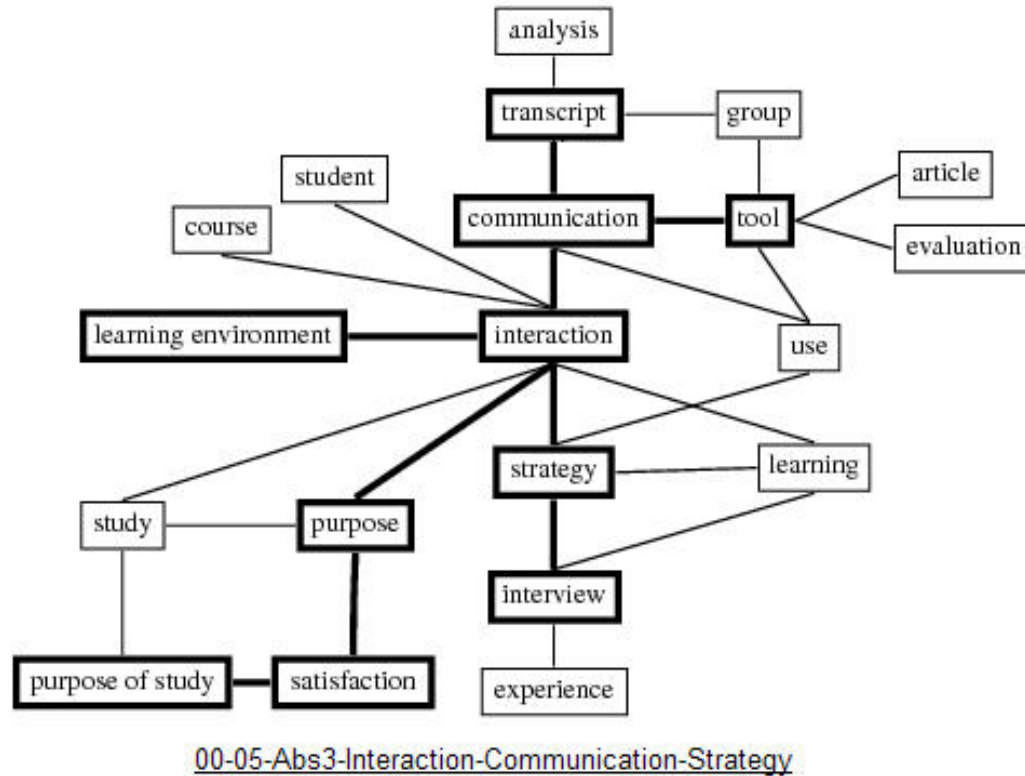
Figure 3. 00-05-Abs3-Interaction-Communication-Tool lexical map.

Emergence of computers.

During the second time period, the term *computer* emerges as shown in Figure 4 (94-99-Abs6-Content-Time-Computer). *Computer* had not appeared in the previous time period. *Computer* connects to *tool*, *access*, and *content* as pass-1 terms, illustrating perhaps the use of computers as the tool for providing access and delivering content emerging in the research literature. During this time period, *computer access* was a major concern for rural, low income, and disabled populations. However, neither *computer* nor *access* appears as a pass-1 term in 2000 to 2005. The availability of advanced telecommunications capability in the United States has increased dramatically, making access of decreased concern (FCC, 2004).
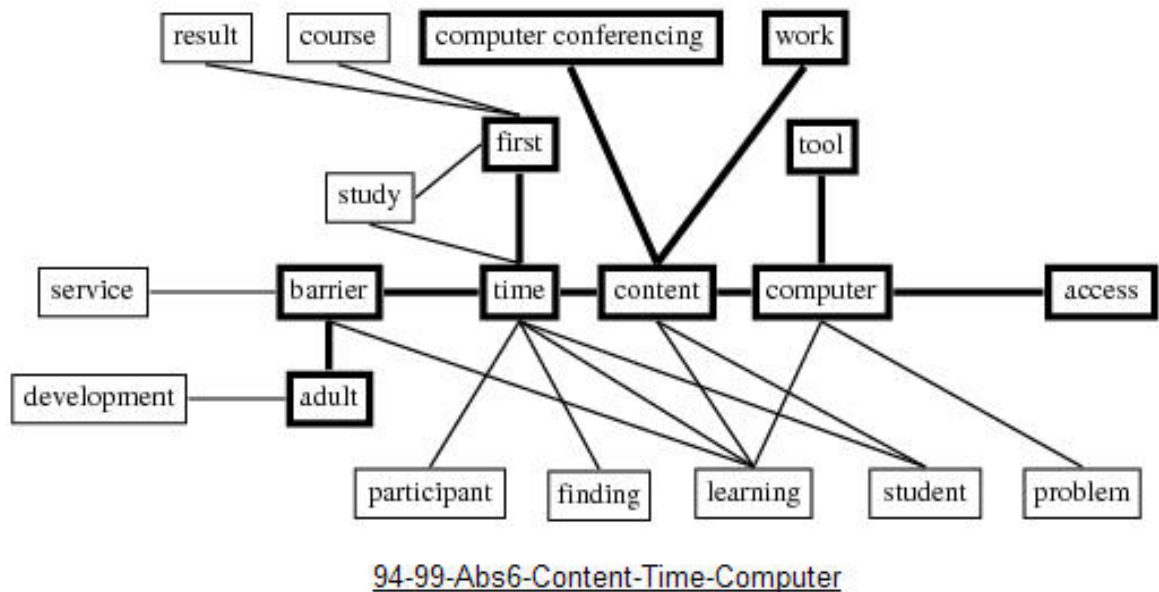
94-99-Abs6-Content-Time-Computer

Figure 4. 94-99-Abs6-Content-Time-Computer.

Cohesion and Coupling Graphs

Two variables, cohesion and coupling, are used to measure the internal strength and interactivity of maps in a time period. Cohesion is defined as the mean of pass-1 strengths of a network, and coupling is defined as the square root of the sum of the squares of the pass-2 strengths of a network (Ritzhaupt, 2003). Cohesion is used to measure the internal strength of a lexical map, while coupling represents a lexical map's position in strength of interaction with other lexical maps. Those lexical maps with high cohesion have strongly related terms; whereas, those with high coupling have stronger relationships with other lexical maps. A Cartesian graph is constructed with cohesion along the vertical axis, coupling along the horizontal axis, and the origin formed at their median values. An interesting result from the cohesion and coupling graphs (CCG) across the three time periods is that relatively few of the lexical maps exhibit both higher than average cohesion and coupling (those appearing in the first quadrant of the Cartesian plane). This attests to distance education covering a broad spectrum of research literature with a relatively broad focus of topics and terminology. Across the three time periods, the 1987 to 1993 time period has only one map in the upper right quadrant; both subsequent time periods have two.

Table 3 displays the lexical map names of the diagrams in the following section. The CCG for the 1987 to 1993 time periods is shown in Figure 5, with the median coupling and cohesion shown at each axis and each node, with *X* representing lexical map 87-93-Abs*X-theme*. The diagram illustrates that lexical map one (87-93-Abs1-Distance Education-Course-Student) has the highest degree of cohesion and coupling, which would indicate the pervasiveness of *distance education, course and student* as a theme. In addition, the theme *distance education, course and student* exhibits a strong internal relationship and has the highest interactivity with other themes in the

50

time period. The theme 87-93-Abs2-Need-Development-Quality exhibits the third-highest degree of cohesion and a slightly below-average coupling. The theme 87-93-Abs8-Role-Field exhibits an average degree of cohesion and the second-lowest degree of coupling, indicating this theme does not appear to interact with other themes during this time period.
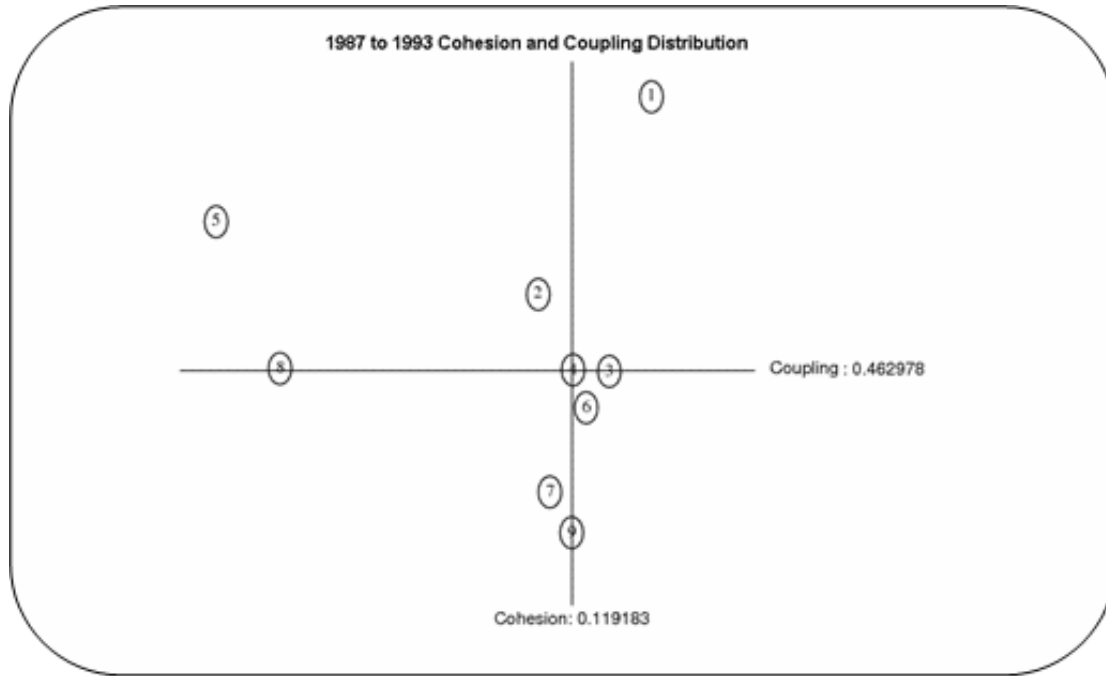

Figure 5. Cohesion and coupling graph (1987 to 1993).

The CCG for the 1994 to 1999 time period is shown in Figure 6. The lexical maps 94-99-Abs1-Study-Distance Education-Learning and 94-99-Abs4-Development-Theory-Information exhibit the highest coupling with other themes. This indicates a time period with greater emphasis on the development of sound *theory* and systematic *study* in *distance education*. The lexical map 94-99-Abs2-Interview-Method-Process has a below-average degree of coupling and the third-highest degree of cohesion, indicating that *interviews* may have been a predominant *method* utilized during this time period.
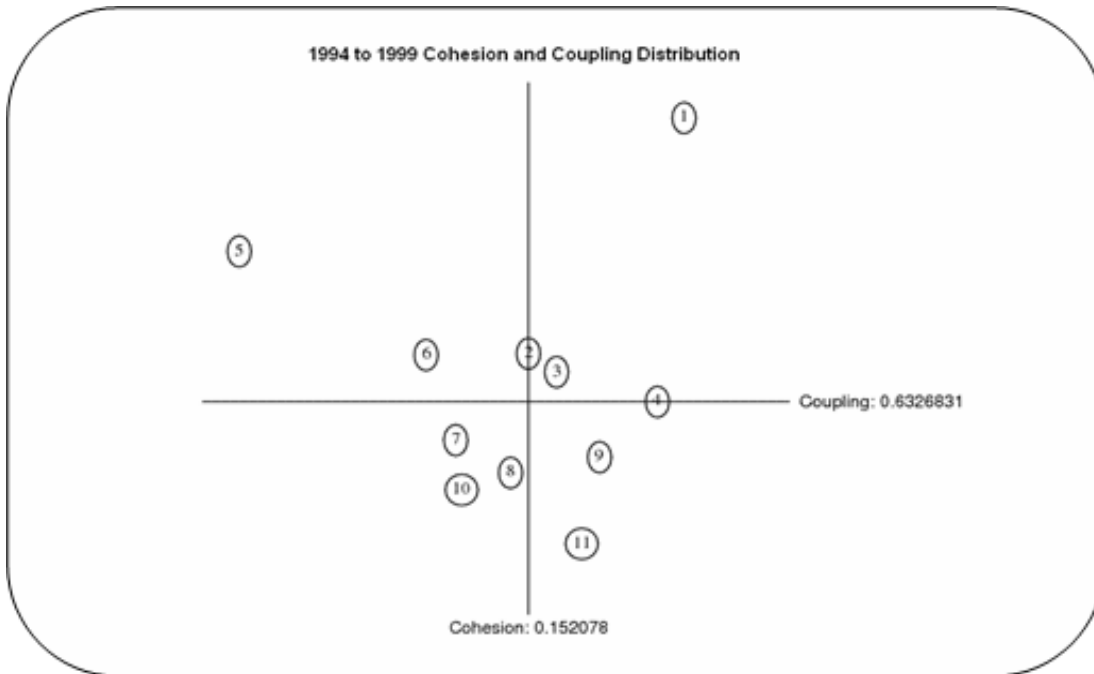
Figure 6. Cohesion and coupling graph (1994 to 1999).

The CCG for the 2000 to 2005 time period is shown in Figure 7. The lexical maps 00-05-Abs1-Study-Distance-Student-Learning and 00-05-Abs4-Teacher-Experience-Results exhibit a high degree of coupling with other themes and also strong internal strength (cohesion). The lexical map 00-05-Abs3-Interaction-Communication-Tool exhibits the second-highest degree of interactivity within this time period. The lexical map 00-05-Abs2-Activity-Research-Practice has a modest degree of cohesion and the poorest degree of interaction with other themes.
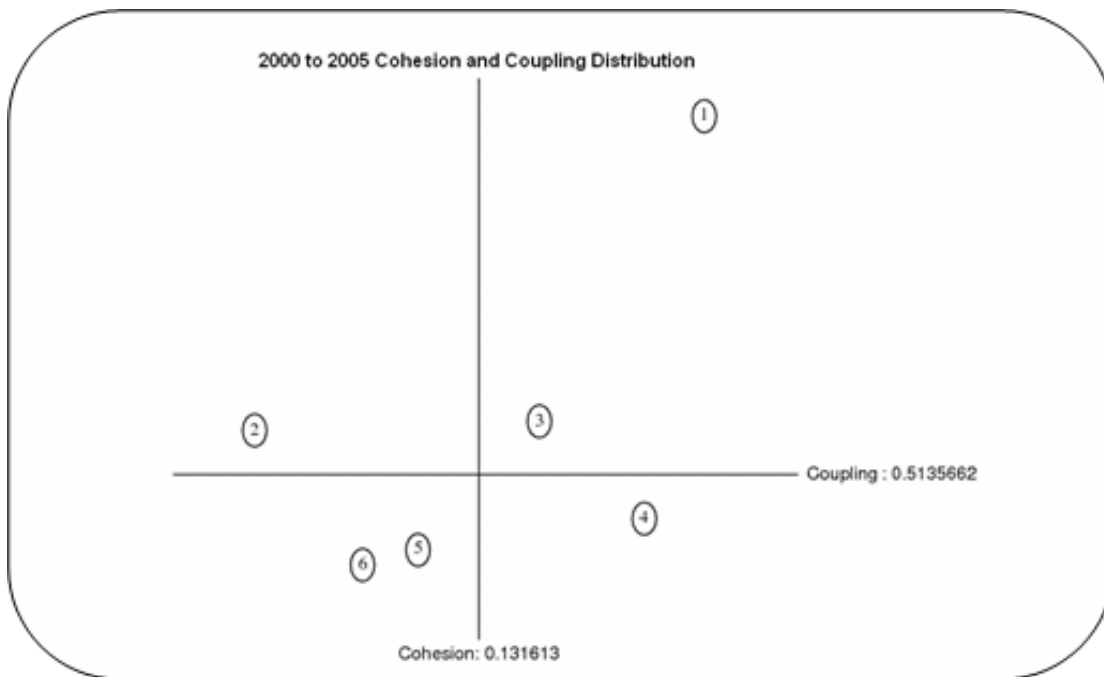
Figure 7. Cohesion and coupling graph (2000 to 2005).

Super Networks

There are essentially three types of lexical maps: principal, secondary, and isolated (Coulter, 1998). Principal maps are connected to one or more (secondary) maps. Secondary maps are linked to principal maps through a high number of pass-2 terms, and isolated networks are those with a very low number of pass-2 terms interacting with other maps (Coulter et al., 1998). The analysis of the CCGs and the super networks addresses the second research question, "What are the relationships among the themes identified in abstracts published in major North American distance education research journals?"

Coupling is taken to the next level with super networks. Super networks are a more focused coupling visualization that describes a pair-wise relationship between lexical maps within a time period. Super networks are defined with principal and secondary maps as follows: If Map-A has internal terms that are pass-2 terms in x links of Maps-B, and each of these x links has a pass-2 strength value that surpasses the minimum pass-1 strength value of Map-B, then Map-A is a secondary network of Map-B (Coulter et. al., 1998). The super networks for the abstract corpus by time period are shown in Figure 8.
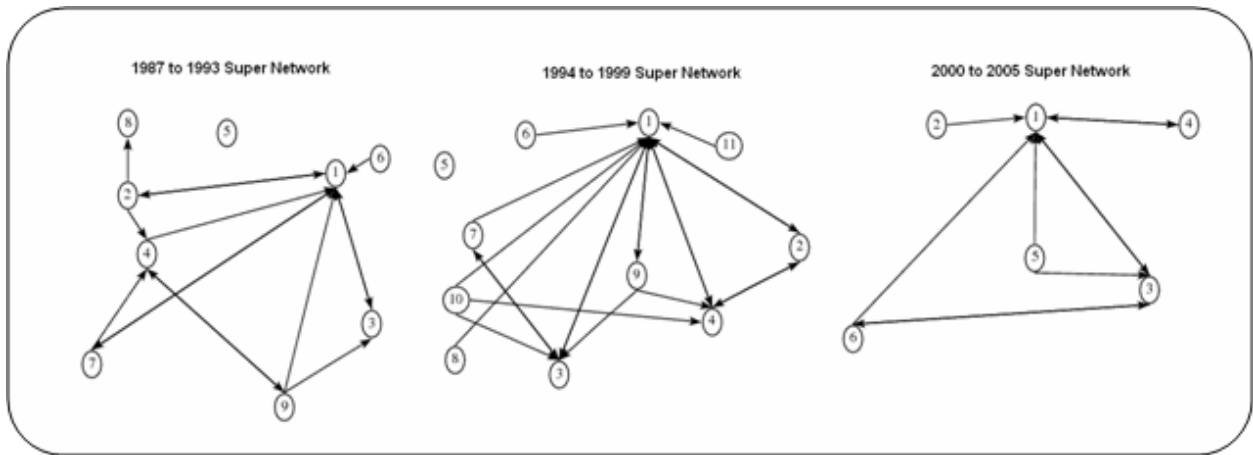
Figure 8. Super networks for abstracts by time period.

The 1987 to 1993 time period shows 87-93-Abs1-Distance Education-Course-Student as a central theme for this time period because it serves as the primary map for six of the nine lexical maps resulting in the time period. Also of importance are the themes 87-93-Abs4-Relationship-Model-Process and 87-93-Abs3-Program-Method-University. The 1994 to 1999 time period indicates 94-99-Abs1-Study-Distance Education-Learning as the central theme followed by 94-99-Abs4-Development-Theory-Information and 94-99-Abs3-Need-Education-Role. The 2000 to 2005 time period demonstrates the 00-05-Abs-Study-Distance-Student lexical map followed by the 00-05-Abs3-Interaction-Communication-Tool lexical map as the central themes in the time period.

The 87-93-Abs5-Concept-Term is an isolated map in the 1987 to 1993 time period, and 94-99-Abs5-Goal-Value is an isolated map in the 1994 to 1999 time period, while the 2000 to 2005 time period did not generate any isolated maps. The isolated lexical maps indicate these themes (lexical maps) do not interact with other themes within the time period and may be interpreted as a fading theme. Few themes appear to be completely isolated from the other themes in each time period.

Analysis across Time Periods

The transformations of lexical maps and their intersections with other lexical maps across time periods have the potential to provide insights into the emergence of distance education research themes. To quantify this type of analysis, a similarity index (SI) approach can be adopted (Callon, Courtial, & Laville, 1991). SI measures the int   ersection of descriptors in two lexical maps but does not directly include corresponding links. Since all terms in lexical maps are at least indirectly linked, this measure captures some portion of similarity. Consider two lexical maps N_i and N_j, and let w_i be the number of terms in N_i and let w_j be the number of terms in N_j. Finally, let w_{ij} be the number of descriptors N_i and N_j then (Coulter et al., 1998, p. 1218),

$$SI(w_i, w_j, w_{ij}) = 2 \times (\frac{w_{ij}}{w_i + w_j}), 0 < SI \leq 1$$

.

Similarity index analysis provides an answer to the third research question, "How have the themes identified in the abstracts published in major North American distance education research journals changed over time?" For example, the emergence of *interaction* in distance education research literature is an important one. The term *interaction* does not appear in the first time period then appears in the 94-99-Abs10-Way-Instruction-Interaction and 00-05-Abs3-Interaction-Communication-Tool as a pass-1 node. In particular, SI can be used to trace the emergence history of the term *interaction*. Figure 9 illustrates the relationship of 94-99-Abs10-Way-Instruction-Interaction with the lexical maps in the previous time period (only those *SI* > .4 are shown). Hence, 94-99-Abs10-Way-Instruction-Interaction incorporates several terms from the 1987 to 1993 time period. The strongest relationship is with the 87-93-Abs6-Difference-Level-Instruction theme, indicating the term *instruction* has a direct link with the emergence history of the term *interaction* in distance education research literature.

N_1	N_2	w_1	w_2	w_{12}	*SI*
94-99-Abs10-Way-Instruction-Interaction	87-93-Abs4-Relationship-Model-Process	20	20	9	0.45
	87-93-Abs5-Concept-Term	20	10	7	0.47
	87-93-Abs6-Difference-Level-Instruction	20	18	11	0.58
	87-93-Abs7-Teaching-Education-Factor	20	20	8	0.40
	87-93-Abs8-Role-Field	20	12	7	0.44
	87-93-Abs9-Evaluation-Problem-Design	20	20	9	0.45

Figure 9. Similarity index for 94-99-Abs10-Way-Instruction-Interaction.

The 94-99-Abs6-Content-Time-Computer lexical map exhibits the emergence of the term *computer* in distance education research, which did not emerge in the previous time period. The emergence of the term *computer* in distance education is undoubtedly a function of the Internet and the pervasiveness of lower-cost, personal computers. However, its emergence may be interconnected with other facets of distance education research. Figure 10 illustrates the relationship of 94-99-Abs6-Content-Time-Computer with the lexical maps in the previous time period (Only those *SI* > .25 are shown). The 94-99-Abs6-Content-Time-Computer theme does not intersect with as many or as strongly as the previous example. However, the theme does mildly intersect with four themes from the previous time period.

N_1	N_2	w_1	w_2	w_{12}	SI
94-99-Abs6-Content-Time-Computer	87_93-Abs1-Distance Education-Course-Student	20	20	6	0.30
	87-93-Abs3-Program-Method-University	20	20	5	0.25
	87-93-Abs6-Difference-Level-Instruction	20	18	6	0.32
	87-93-Abs9-Evaluation-Problem-Design	20	20	5	0.25

Figure 10. Similarity index for 94-99-Abs6-Content-Time-Computer.

Conclusions

What can be concluded from the analysis of research publications on distance education? First, each time period can be characterized by a few central themes illustrated in the super networks and CCGs. The central theme of the first time period or pre-Web era, 87-93-Abs1-Distance Education-Course-Student and 87-93-Abs2-Need-Development-Quality, indicates the need for developing quality distance education courses and institutional policy, as supported by Moore and Kearsley, who state, "A far bigger problem [than Internet access] is the quality of media produced for distribution via the technology" (2005, p.8).

The 1994 to 1999 time period (emerging Web era) places emphasis on the study of distance education (94-99-Abs1-Study-Distance Education-Learning) and the development of sound theory (94-99-Abs4-Development-Theory-Information) to guide research efforts. As pointed out by Moore and Kearsley (2005),

> Research is ineffective when it is not set in a theoretical framework. Researchers are not able to build on the work of others, they are less likely to identify the really significant questions, and their results are of limited generalizability. (p. 256)

This was also echoed in Lee, Driscoll, and Nelson's (2004) work. What is unique about this finding is that the emerging Web period is characterized by a heavier emphasis in theory development.

The central theme in the third time period, maturing Web era, appears to emphasize the study of distance education (00-05-Abs1- Study-Distance-Student), and there is a greater emphasis on strategies for communication and interaction (00-05-Abs3-Interaction-Communication-Tool). This triangulates the recent findings of Zawacki-Richter, Bäcker, and Vogt (2009) in that interaction and communication among learning communities were the most frequently studied areas.

Second, some emerging, fading, and evolving themes across time periods of distance education research have been identified. Teleconferencing as a theme seems to have faded from the research literature, beginning in the second time period. Meanwhile, the theme *interaction* and the tool *computer* have emerged. Subsequently, the theme of interaction swiftly evolved into a central role, as can be seen by examining the super network and CCGs from the 1994-1999 and 2000-2005 time periods. In the second time period, 94-99-Abs10-Way-Instruction-Interaction was a secondary term for many of the other themes, while in the third time period 00-05-Abs3-Interaction-Communication-Tool is a principal theme to three other themes during this period. The SI revealed *instruction* may be a direct link to the emergence of *interaction* in distance education research literature.

Third, it can be concluded that distance education research is broad in scope and can be characterized as having few consistent and focused lines of inquiry in the research literature. This can be deduced by examining the relatively small number of themes (lexical maps) that exhibit both a high degree of cohesion and coupling across the three time periods. This research confirms that there are many facets and contours of distance education. Co-word analysis overcomes many of the limitations of a priori research methodology, which potentially loses a tremendous amount of information. However, as of yet, co-word analysis is still an imperfect science because natural language, in this case English, contains many idiosyncrasies. The use of specific jargon in the field of distance education can make it difficult to ascertain the precise meaning of a lexical map or relationship.

At minimum, this research has demonstrated the capability of the co-word analysis method and its extensions on a large corpus of distance education literature. Other disciplines, such as software engineering (Coulter et al., 1998), employ strict taxonomies for their publication databases. This research should encourage and aide researchers and practitioners in the field of distance education to investigate alternate methods to precisely define the contours of this multifaceted discipline.

Acknowledgements

References

American Journal of Distance Education. (2008). Retrieved from http://www.ajde.com/index.htm

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

Bernard, R.M., Abrami, P.C., & Lou, Y., Borokhovski, E., Wade, A., Wozney, L., Wallet, P.A., Fiset, M., & Huang, B. (2004). How does distance education compare with classroom instruction? A meta-analysis of the empirical literature. *Review of Educational Research, 74*(3), 379-439.

Callon, M., Courtial, J.P., & Laville, F. (1991). Co-word analysis as a tool for describing the network of interactions between basic and technological research: The case of polymer chemistry. *Scientometrics, 22*(1), 153–205.

Callon, M., Law, J., & Rip, A. (1986). *Mapping of the dynamics of science and technology.* London: Macmillan.

Chan, L. (2005). WebCT revolutionized e-learning. *UBC Reports*, *51*(7). Retrieved from http://www.publicaffairs.ubc.ca/ubcreports/2005/05jul07/webct.html

Chen, T. (2003). Recommendations for creating and maintaining effective networked learning communities: A review of the literature. *International Journal of Instructional Media*, *30*(1), 35-44.

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.

Courtial, J. P. (1994). A coword analysis of scientometrics. *Scientometrics, 31*(3), 251-260.

Ding, Y., Chowdhury C. G., & Foo, S. (2001). Bibliometric cartography of information retrieval research by using co-word analysis*. Information Processing and Management*, *37*(6), 817-842.

Edwards, D. (1995). *Recent advances in descriptive multivariate analysis*. Royal Statistical Society Lecture Note Series, Chapter 7, Graphical Modeling, 135-156. Clarendon Press.

Federal Communications Commission (FCC) (2004). *Availability of advanced telecommunications capability in the United States*. FCC 04-208, GN Docket No. 04-54 Fourth Report to Congress.

Garrison, D., & Shale, D. (1987). Mapping boundaries of distance education: Problems in defining the field. *American Journal of Distance Education*, *15*(1), 7-23.

International Centre for Distance Learning. Retrieved from the Institute of Educational Technology and Library of the Open University: http://www-icdl.open.ac.uk/.

Journal of Distance Education. (2008). Retrieved from http://www.jofde.ca/index.php/jde/index

Kaufman L., & Rousseeuw, P.J. (1990). *Finding groups in data*. John Wiley & Sons, Inc., Wiley Series in Probability and Mathematical Statistics. http://www3.interscience.wiley.com/cgi-bin/bookhome/119054821

Keegan, D. (1996). *Foundations of distance education* (3rd ed.). London: Routledge.

Khan, B. H. (1997). Web-based instruction: What is it and why is it? In B.H. Khan (Ed.), *Web-based instruction* (pp. 5–18). Englewood Cliffs, NJ: Educational Technology Publications.

Koble, M. A., & Bunker, E. L. (1997). Trends in research and practice: An examination of the American Journal of Distance Education 1987 to 1995. *American Journal of Distance Education, 11*(2), pp. 19-38.

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.

Moore, M.G. (1973). Towards a theory of independent learning and teaching. *Journal of Higher Education, 44*, 661-679.

Moore, M., & Kearsley, G. (1996). *Distance education: A systems view*. Belmont, CA: Wadsworth Publishing.

Moore, M., & Kearsley, G. (2005). *Distance education: A systems view*. (2nd ed.). Belmont, CA: Wadsworth Publishing.

NCSA (2008). *About NCSA Mosaic*. Retrieved from http://www.ncsa.uiuc.edu/Projects/mosaic.html

Phipps, R., & Merisotis, J. (1999). *What's the difference? A review of contemporary research on the effectiveness of distance learning in higher education.* Washington, DC: The Institute for Higher Education Policy. Retrieved from http://www.ihep.org/Publications/publications-detail.cfm?id=88

Porter, J. (1997). *Audience and rhetoric*. Engelwood Cliffs, NJ. Prentice Hall.

Reisman, S., Dear, R. G., & Edge, D. (2001). Evolution of web-based distance learning strategies. *International Journal of Educational Management*, *15*(5), 245-251.

Ritzhaupt, A. D. (2003, April). An exploration of themes and trends in information-related education: Using computer-aided content analysis. *Proceedings of the Annual College of Computing, Engineering, and Construction Symposium and Career Showcase*, Jacksonville, FL.

Rourke, L., & Szabo, M. (2002). A content analysis of the Journal of Distance Education 1986-2001. *Journal of Distance Education, 17*(1), 63-74.

Salas, E., Kosarzycki, M. P., Burke, C. S., Fiore, S. M., & Stone, D. L. (2002). Emerging themes in distance learning research and practice: Some food for thought. *International Journal of Management Reviews*, *4*(2), 135-153.

Sherry, L. (1995). Issues in distance learning. *International Journal of Educational Telecommunications, 1*(4), 337–365.

Sherry, A., Fulford, C, & Zhang, S. (1998). Assessing distance learners' satisfaction with instruction: A quantitative and a qualitative measure. *American Journal of Distance Education, 12*(3), 4-28.

van Raan, A.F.J., & Tijssen, R.J.W. (1993). The neural net of neural network research. *Scientometrics, 26*(1), 169–192.

Verduin, J.R., & Clark, T. A. (1991). *Distance education: The foundations of effective practice*. San Francisco: Josey-Bass.

Whittaker, J. (1989). *Creativity and conformity in science: Titles, keywords, and co-word analysis*. London: SAGE.

Zhao, Y., Lei, J., Yan, B., & Tan, S. (2005). What makes the difference? A practical analysis of research on the effectiveness of distance education. *Teachers College Record*, *107*(8), 1836-1884.

Zawacki-Richter, O., Bäcker, E. M., & Vogt, S. (2009). 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). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/741/1433

Athabasca University