

A Bibliometric Mapping of the Structure of STEM Education using Co-Word Analysis

Shimelis G. Assefa

Department of Research Methods and Information Science, University of Denver, Ruffatto Hall, Room 244, 1999 E. Evans Ave, Denver, CO 80208. E-mail: sassefa@du.edu

Abebe Rorissa

Department of Information Studies, University at Albany, State University of New York, Draper Hall, Room 116, 135 Western Avenue, Albany, NY 12222. E-mail: arorissa@albany.edu

STEM, a set of fields that includes science, technology, engineering, and mathematics; allied disciplines ranging from environmental, agricultural, and earth sciences to life science and computer science; and education and training in these fields, is clearly at the top of the list of priority funding areas for governments, including the United States government. The U.S. has 11 federal agencies dedicated to supporting programs and providing funding for research and curriculum development. The domain of STEM education has significant implications in preparing the desired workforce with the requisite knowledge, developing appropriate curricula, providing teachers the necessary professional development, focusing research dollars on areas that have maximum impact, and developing national educational policy and standards. A complex undertaking such as STEM education, which attracts interest and valuable resources from a number of stakeholders needs to be well understood. In light of this, we attempt to describe the underlying structure of STEM education, its core areas, and their relationships through co-word analyses of the titles, keywords, and abstracts of the relevant literature using visualization and bibliometric mapping tools. Implications are drawn with respect to the nature of STEM education as well as curriculum and policy development.

Introduction

STEM is a commonly used acronym for “Science, Technology, Engineering, and Mathematics.” However, it is an umbrella term that encompasses several other fields ranging from environmental, agricultural, and earth sciences to life

science and computer science. Although grouping together several disciplines under the acronym “STEM” is not a problem on its own, it does present challenges when educators embark on curriculum development or when policymakers and lawmakers attempt to prioritize the limited funds available for research and development activities in STEM fields. Part of the challenge is coordinating activities among numerous stakeholders and interest groups. For instance, according to a recent report by the National Science and Technology Council Committee on STEM education (CoSTEM, 2011), 11 federal agencies in the U.S. invest in or support STEM-related activities and programs.

Because of the complex nature of STEM as an amalgam of fields, educators have difficulty addressing problems associated with preparing students in the various individual STEM fields. In an effort to understand and characterize STEM education and document an inventory of programs, activities, and funding priorities across 11 federal agencies, CoSTEM defined STEM education, admittedly in a narrow sense, as “formal or informal education that is primarily focused on physical and natural sciences, technology, engineering, and mathematics disciplines, topics, or issues (including environmental science education or environmental stewardship)” (CoSTEM, 2011, p. 79). In the same report, it was indicated that for the 2010 fiscal year, 252 investments that accounted for \$3.4 billion in federal spending across 11 U.S. federal agencies were made in STEM education. This accounts for about 0.3% of the nation’s \$1.1 trillion spent on education in that same year (CoSTEM, 2011, p. 37).

Although the acronym STEM seemingly includes four main fields, the broad landscape of the STEM set of fields is evident even when one singles out the “Science” component in STEM. Within the sciences alone, one can cite examples of subdisciplines that span not just school science courses in K–12, but also extend to graduate studies in microbiology. STEM is a complex field and has widely varying and

Received: October 21, 2012; revised December 13, 2012; accepted December 14, 2012

© 2013 ASIS&T • Published online 16 September 2013 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/asi.22917

interdisciplinary subfields. In the face of such complexity, understanding the domain of STEM, its knowledge, and STEM education has significant implications for preparing the workforce with the requisite knowledge, developing appropriate curricula, providing teachers with the necessary professional development, focusing research dollars on areas so that they have maximum impact, and developing national educational policy and standards.

Despite the tremendous amount of attention given to STEM education as a key lever for a nation's competitiveness or innovation capacity, there is not yet a clear understanding of what STEM education entails and how to support and/or implement programs and activities in the field (Bybee, 2010; Sanders, 2009). Understanding the complex web of relationships surrounding the field of STEM education is the first step to better planning of curricular activities, programs, and research. Understanding the development of a scientific domain, the main topics or the main research fields of the domain, and how these topics or fields relate to one another is usually the goal of bibliometric studies or bibliometric network analyses (Waltman, van Eck, & Noyons, 2010).

Generally, in the sociology of science and scientometrics, different approaches have been used to characterize and understand a given domain area. These methods are variously known as co-word analysis (Callon, Law, & Rip, 1986; Whittaker, 1989), citation mapping (Healey, Rothman, & Hoch, 1986; Small, 1999), and bibliometric analysis, bibliometric modeling, or bibliometric mapping (Crouch, Irvine, & Martin, 1986; Franklin & Johnston, 1988; King, 1987; Noyons, 2001, 2005). In addition to the usefulness of these methods for evaluation and measurement of science, they allow for the creation of visual maps that represent the scholarly activities and the relationships that exist in a given domain.

This study uses a co-word analysis method to analyze the field of STEM education to identify insights that will have policy and curricular development implications. By analyzing titles, author-provided keywords, and abstracts of tens of thousands of scholarly works from two major databases (Web of Science and ERIC), we aim to identify core knowledge areas that characterize the domain of STEM education. A secondary goal is to investigate whether visualization and maps produced through co-word analysis can provide useful insights to support the work of policymakers and educators in curriculum and professional development as they shape the future of STEM and STEM education, and achieve better dividends in terms of a well educated workforce for the 21st-century workplace.

In sum, this is knowledge diffusion research. The goal is to map scholarly communication in the field of STEM education. By analyzing the artifacts of the domain, i.e., journal articles, books, and other publications in the various STEM fields as well as in STEM education, we aim to describe the domain of STEM and STEM education, or the topology of the field. We attempt to answer the following research questions:

- RQ1: What are the core knowledge areas and their relationships in STEM and STEM education?
- RQ2: What can be learned from the visualizations and maps for the purpose of curriculum development?
- RQ3: What knowledge can be gained from the visualization maps for policymakers and other stakeholders in research as well as resource management and professional development activities?

Background and Previous Research

STEM and STEM Education

A previous study that looked into STEM education showed the existence of a wide range of interpretations and conceptions by different stakeholders, ranging from K–12 school districts where it is seen as traditional disciplinary coursework composed of the subjects that make up STEM, to policy views most popularized by the National Science Foundation (NSF), to current understanding of the integration of the various disciplines as used in solving real-world problems (Breiner, Harkness, Johnson, & Koehler, 2012). The NSF, whose mission includes support for research and education in a number of STEM fields (except medical sciences), defines STEM broadly. In the NSF definition, STEM includes common categories such as mathematics, natural sciences, engineering, and computer and information sciences, as well as fields in the social and behavioral sciences such as psychology, economics, sociology, and political science (Green, 2007).

Notable activities around STEM date back to the early 1980s, when the American Association for the Advancement of Science (AAAS) created *Project 2061* (American Association for the Advancement of Science, 1989) in response to an earlier report by the National Commission on Excellence in Education (1983) entitled *A nation at risk: The imperative for educational reform*. In its most current form, the focus on STEM education dates from the early 1990s, first under the umbrella term SMET (Sanders, 2009), then changed to STEM in 2001 when used by the NSF director at the time (Teaching Institute for Excellence in STEM [TIES], 2010). An unprecedented level of activity and attention was given to STEM in the last decade. According to the federal STEM education portfolio report (CoSTEM, 2011), several U.S. government departments and agencies, including the National Science Foundation (NSF), National Aeronautics and Space Administration (NASA), US Department of Agriculture (USDA), Environmental Protection Agency (EPA), Department of Education (DoEd), Health and Human Services (HHS), Department of the Interior (DOI), Nuclear Regulatory Commission (NRC), Department of Transportation (DOT), Department of Defense (DOD), and Department of Energy (DOE), are involved in STEM-related efforts. What was once primarily an initiative of the government has now attracted private enterprise and nonprofit foundations, universities, and national advocacy groups. Under the UTeach Institute that started at the University of Texas at Austin in 1997 and is now replicated in 29

universities nationwide, several multifaceted stakeholders are mobilized under the program dubbed “100Kin10” to prepare 100,000 STEM teachers in 10 years (UTeach Institute, 2012).

The number of reports, initiatives, and legislation launched is also an indicator of the scope of activities in STEM disciplines. Notable in this regard include Race to the Top (a \$4.35 billion initiative by the Obama administration), Educate to Innovate (The White House, 2009), Investing in Innovation fund (i3) (\$650 million), U.S. STEM Coordination Act, American Competitiveness Initiative, and *Rising above the gathering storm* (National Academy of Sciences, 2007).

Despite a wide array of initiatives and efforts, there are not yet substantial U.S. advances in student achievement in STEM subjects. According to the results of the 2009 Program for International Student Assessment (PISA) covering 65 OECD countries, the U.S. ranked 31st in mathematics (487 with a 496 PISA average) and 23rd in science (502 with a 501 PISA average) (OECD, 2011). The 2007 Trends in International Mathematics and Science Study (TIMSS) report portrayed similar but better scores for mathematics and science subjects at the fourth- and eighth-grade levels. The average scores for U.S. fourth-graders were 529 and 539, and for eighth-graders were 508 and 520 higher than the TIMSS scale average (500 at both grades for mathematics and science, respectively) (Gonzales, Williams, Jocelyn, Roey, Kastberg, & Brenwald, 2008). The average U.S. fourth-grade score was higher than those of students in 23 of the 35 (for mathematics) and 25 of the 35 (for science) countries involved in the study; for the eighth grade, the average U.S. score was higher than those of students in 37 of the 47 (mathematics) and 35 of the 47 (for science) other countries (Gonzales et al., 2008). The complexity of STEM fields is one important factor to be investigated in order to launch an effective program to prepare students for jobs in a 21st-century workforce.

The Study of Scholarly Communication and Research Output in a Given Domain

Existing knowledge output in a given domain is a result of years of scholarly communication between and among communities of researchers and scientists who identify themselves as belonging to a particular discipline. In any scientific endeavor, the output of the scholarly communication is recorded in large part in journal articles, books, conference proceedings, editorials, conference presentations, etc. Authors, scholars, and researchers in the field communicate their findings through published works building upon their own or peers' work, expanding on, refuting, or confirming theories, hypotheses, applications, and assumptions. That is how continuity and inquiry are maintained. For decades, treating this accumulated knowledge as an object of empirical investigation has been a focus in the sociology of science in general (Callon et al., 1986) and bibliometrics and citation analysis in particular (Borgman & Furner, 2002; King, 1987).

By studying the research output in a given domain, researchers reveal the development of trends and diffusion of ideas within a particular field or discipline. More specifically, by studying bibliographic elements such as the author, title, abstract, and keywords of scientific publications in a given domain, researchers establish the structure of a field or map of the discipline. The structure reveals a web of relationships between cited and citing documents, the themes that emerge in the scientific discourse, and the landscape of the field under investigation in general. The structure of the research activities is often depicted as a “funnel of interest” (Callon et al., 1986) or as a continuous looping flow (Borgman & Furner, 2002).

Traditionally, bibliometric approaches have been widely used to determine the scientific and academic merit of a particular journal (Waltman et al., 2010; Althouse, West, Bergstrom, & Bergstrom, 2009) and authors' productivity (Borrmann, Mutz, & Daniel, 2009; McCain, 1990), and for creating thematic maps through co-word or co-occurrence analysis (Callon et al., 1986; Courtial & Law, 1989; He, 1999).

Co-word Analysis, Mapping, and Visualization

Based on the co-occurrence of pairs of words, co-word analysis seeks to extract themes and detect the linkages among these themes directly from the subject content of texts. It does not rely on an *a priori* definition of research themes in science. This enables us to follow actors objectively and detect the dynamics of science without reducing them to the extremes of either internalism or externalism (Callon et al., 1986). Co-word analysis is often accompanied by mappings and visualizations of significant words and terms and/or phrases.

Mapping and clustering techniques are heavily used to understand the structure of a bibliometric network, including the identification of major topics in the domain, the relationships between the topics, and the diffusion of ideas over time (Waltman et al., 2010). In view of this continuous and interconnected flow of interests, scientific and research activities are often considered as self-organizing communities (An, Janssen, & Milios, 2002). The nature of the relationships and structure of the communities are represented in what is known as a “bibliometrics map” or “science map” (Buter, Noyons, Van Mackelenbergh, & Laine, 2006; van Eck, Waltman, Dekker, & Van den Berg, 2010). Science maps are symbolic representations of scientific fields or organizations in which the elements of the map are associated with topics or themes. Elements are positioned in the map so that other elements with related or similar characteristics are located in their vicinity, whereas those elements that are dissimilar are positioned at distant locations (Noyons, 2001). The elements in the map can be words and terms, phrases, authors, publications, institutions, scientific topics, or instruments. The purpose of the representation is to enable the user to explore relationships among the elements.

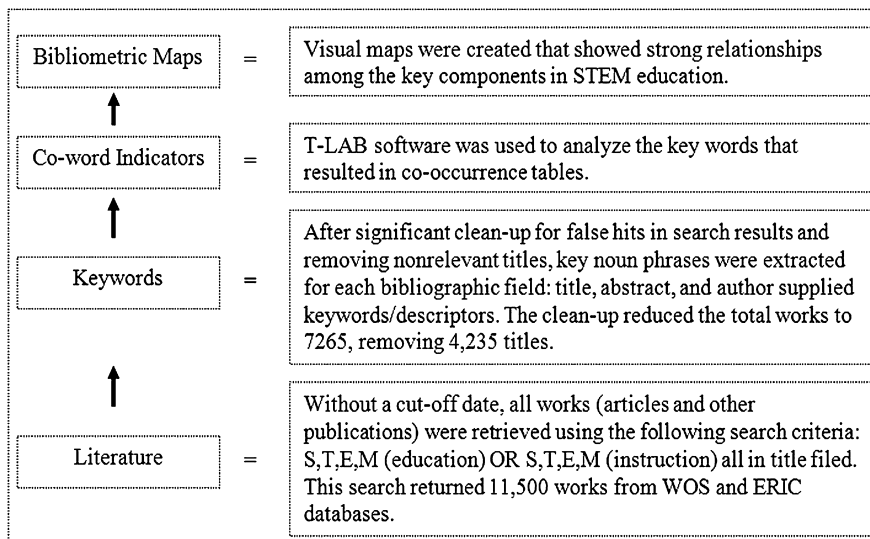


FIG. 1. Summary of method.

TABLE 1. Types of documents retrieved from the databases.

Subject area	Time period	Total documents	Note
Science Education/ Instruction	1925–2010	3270	The first paper on Science education in the title field from the databases considered appeared in 1925
Technology Education/ Instruction	1962–2010	1508	The first paper on technology education started in 1962
Engineering Education/ Instruction	1901–2010	1042	Engineering education dominated the retrieved documents through mid-1950s
Mathematics/Math Education/ Instruction	1917–2010	1425	After 1917, the next paper on Math education in the title field appeared in 1953
STEM Education	2006–2010	20	The first paper on STEM education in the title field from the databases considered appeared in 2006

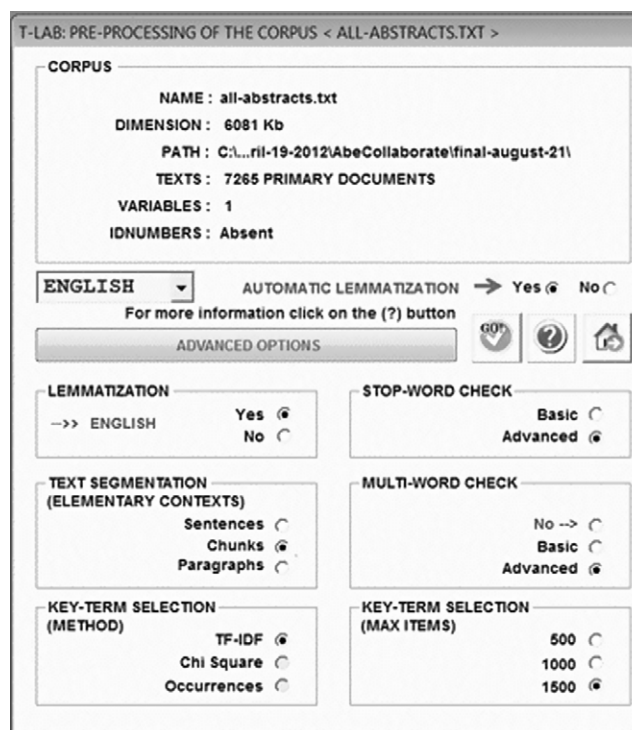


FIG. 2. Corpus preprocessing option.

Mapping and clustering are complementary to each other. Although mapping can be used to represent a fairly detailed picture of the structure of a bibliometric network, relations in more than two dimensions will usually not be visible. Clustering, on the other hand, is not constrained by dimensional restrictions except that it works with binary rather than continuous dimensions (Waltman et al., 2010). Another widely used technique to create visual maps of different bibliographic elements in a low-dimensional scale is multi-dimensional scaling (MDS) (van Eck et al., 2010).

In addition to representing relationships and creating a structure of a domain, bibliometric maps or science maps are used for a wide range of purposes, including to determine innovation activity (Agarwal & Searls, 2009); to support science policy (Buter et al. 2006; Leydesdorff & Wagner, 2009; Noyons, 2001; Noyons, 2005; Noyons, Buter, & van Raan, 2002); as a trend analysis to determine evolution or development of scientific disciplines (Callon et al., 1986;

Marshakova-Shaikevich, 2005; Zitt & Bassecoulard, 1994); and to measure academic productivity (An et al., 2002). In the social studies of science, constructing a map and analyzing science and the activities of scientists through their works have been recognized as methods to understand a field (Callon et al., 1986). Other studies have developed “knowledge domain visualization” (KDV) to understand the development of a given knowledge domain (Chen, 2004).

Previous Research

Apart from using bibliometric techniques to identify and characterize topics in specific science and technology subjects, no study in the extant literature used similar methods to understand the fields of STEM and STEM education in general.

Analysis of documents based on co-occurrences of words is often regarded as a way of grasping the conceptual framework of a field of science (Marshakova-Shaikevich, 2005). Over the years, co-word analysis and co-citation analysis have been used extensively to map the structure of science in

a given field of study (Courtial & Law, 1989), for example, in library and information science and women’s studies (Marshakova-Shaikevich, 2005). Co-word analysis was used to determine intellectual coherence in artificial intelligence (Courtial & Law, 1989), diversity and coherence as indicators of interdisciplinarity in bionanoscience (Rafols & Meyer, 2010), and identification of technological topics in molecular biology (Roche, Besagni, François, Hörlesberger, & Schiebel, 2010).

Based on the co-occurrence of pairs of words, co-word analysis seeks to extract the themes of science and detect the linkages among these themes directly from the subject content of texts. The method of building such thematic or lexical maps usually proceeds in two steps, i.e., identify the list of lexical units used in the language of the given field of science, and identify systematic relationships between these units (Marshakova-Shaikevich, 2005). In some studies the map is created using a list of frequently appearing words or phrases taken from the titles or abstracts of the papers. Because of that, caution is exercised to understand the research theme objectively and detect the dynamics of

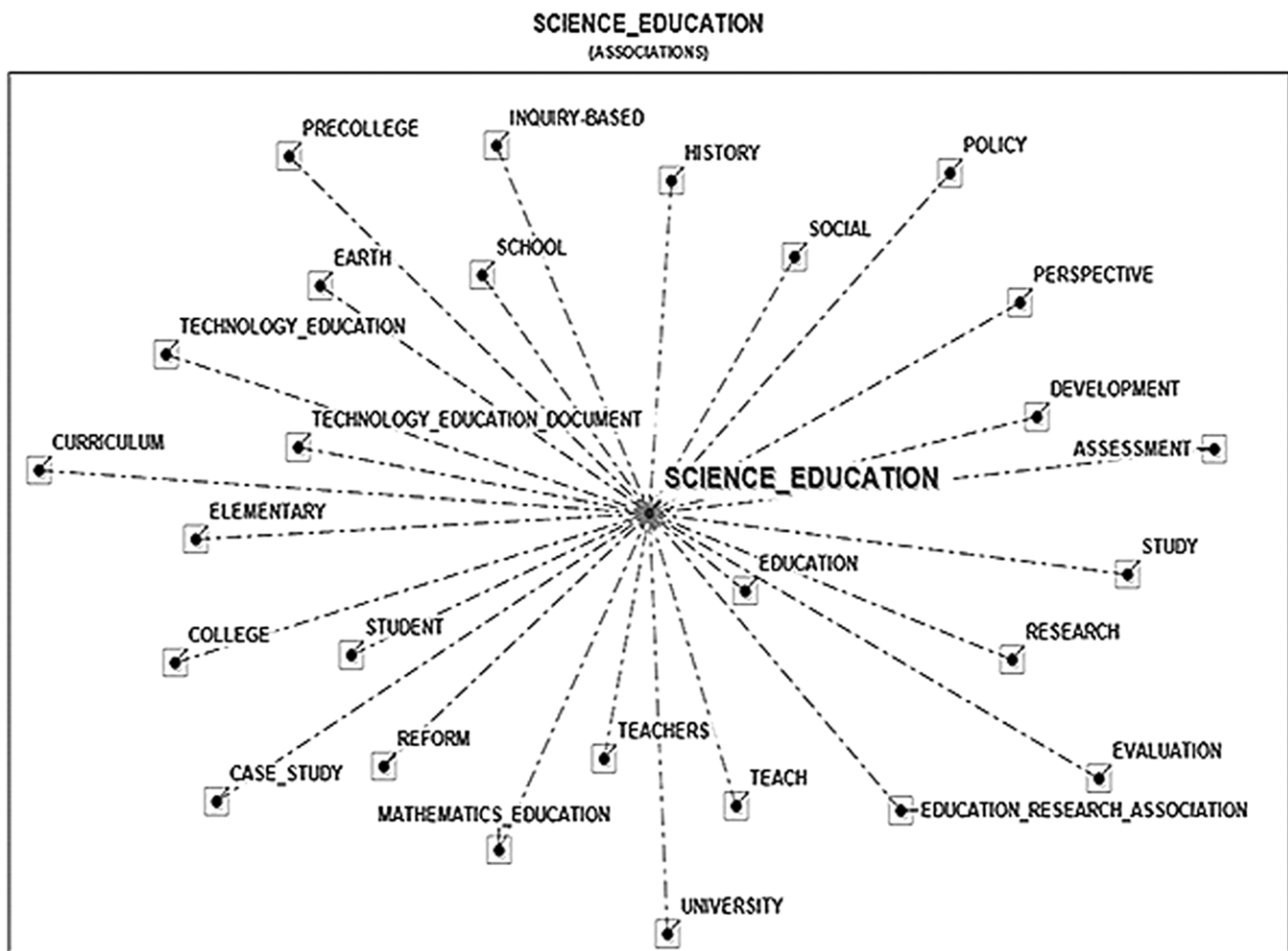


FIG. 3A. Word associations between “science education” and other lemmas in the titles corpus subset.

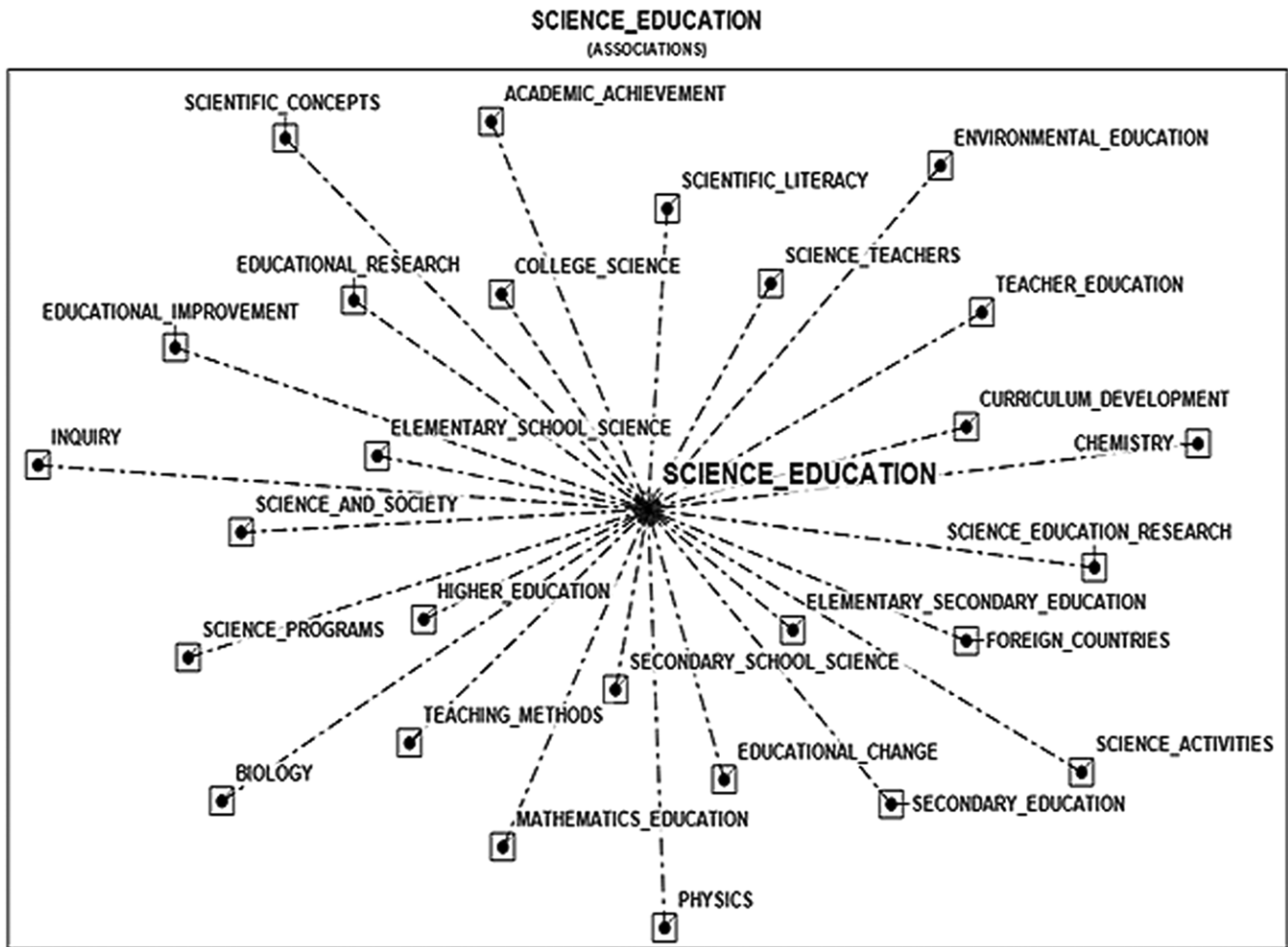


FIG. 3B. Word associations between “science education” and other lemmas in the author-supplied keywords/descriptors corpus subset.

science without reducing them to the extremes of either internalism or externalism (Callon et al., 1986).

Methods

Co-occurrence Analysis

In the current study, to create the bibliometric maps, co-occurrence analysis (a word association method) was used. For each of the keywords selected under STEM education, the word association method determined a significant relationship between words that co-occur. In view of the goal of the study, the keywords selected, and the co-occurrence that was sought were “science education,” “technology education,” “engineering education,” and “mathematics education.” Figure 1 shows the materials and procedures followed in the analysis.

Data Analysis

The data used comprise a text corpus of titles, author-supplied keywords/descriptors/, and abstracts of 7265

documents retrieved from two databases (6101 from Web of Science/WoS and 1164 from ERIC). These two databases were selected for their appropriateness in indexing educational literature and topics in STEM education. The documents classified by the databases as “resource types” include books, conference proceedings, journal articles, dissertations and theses, and reports from 1901 to 2010. A summary of the profiles of the documents is presented in Table 1.

The text corpus was analyzed using T-LAB (<http://www.tlab.it/en/>), a content analysis, visualization, and mapping software package mainly used for word occurrence and co-word mapping. With the text corpus as input, T-LAB converts each word into a lemma (or lexical root) and makes distinctions between various grammatical categories such as verb, noun, and adjective. In other words, it groups a verb and all its variations (past tense, -ing form, etc.) under the root form of the verb. However, a noun and an adjective, even if they have the same root word, are grouped as separate terms. The graphs and maps produced by T-LAB are labeled using the lemmas (lexical units), which serve as

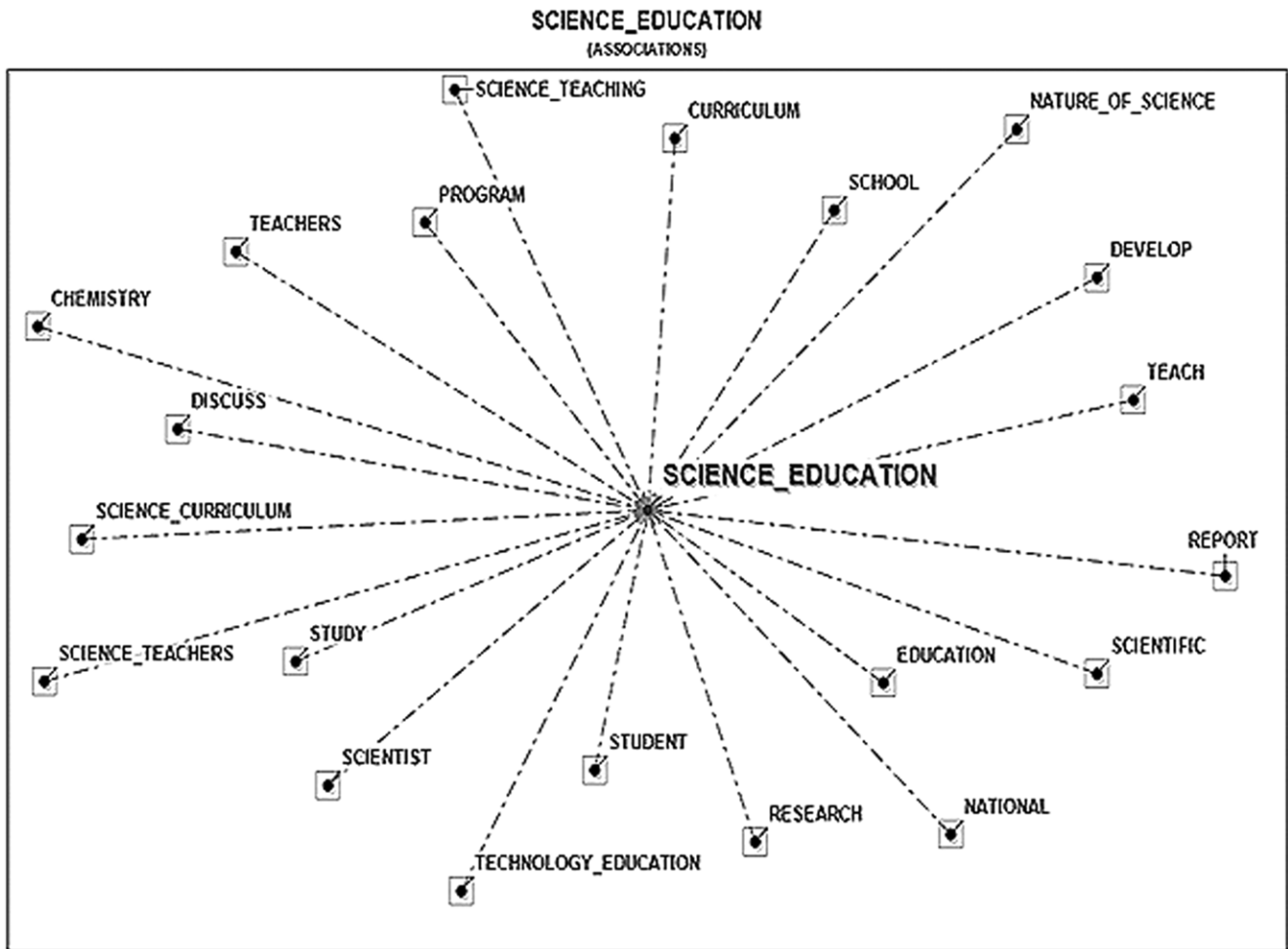


FIG. 3C. Word associations between “science education” and other lemmas in the abstract corpus subset.

units of analysis. T-LAB is an appropriate tool for studies such as this, which try to uncover the underlying structure of a complex field such as STEM education, because it allows analyses based on (1) the frequency of key terms, (2) the context and use of key terms in the sentence and paragraph in which they occur, and (3) the co-occurrence relationships between the key terms.

Apart from a content analysis of the entire set of full-text documents on STEM education, co-word analysis and mapping of the titles, keywords, and abstracts of the documents is the best approach to identify the core domains and their relationships. To determine key co-occurrence relationships surrounding STEM education and the components identified by the label “STEM,” the word association method was used. For each of the key terms selected under STEM, the word association method determined a significant relationship between words that shared co-occurrence. The TF-IDF (term frequency–inverse document frequency) measure, a feature available in the T-LAB software, was used to select key terms from the documents. The approach involved first preprocessing the corpus by specifying the

option, as indicated in Figure 2. The important options in the preprocessing step were the selection of multiword phrases and using chunks to segment the text within the elementary context in which it was found. After the preprocessing generated a multiword list, we used the list to transform the entire corpus, which resulted in selected key terms based on the occurrence threshold value we specified. T-LAB sets the minimum frequency threshold to 4 to guarantee the reliability of the statistical procedures. However, given the size of the three corpus subsets, we increased the threshold to 10 for title and keywords and descriptors corpora and to 70 for abstract corpus subsets.

Results

Research Question #1: What are the core knowledge areas and their relationships in STEM and STEM education?

The T-LAB software computed an association index (similarity coefficient) that essentially determined the number of times the selected key terms occurred across three

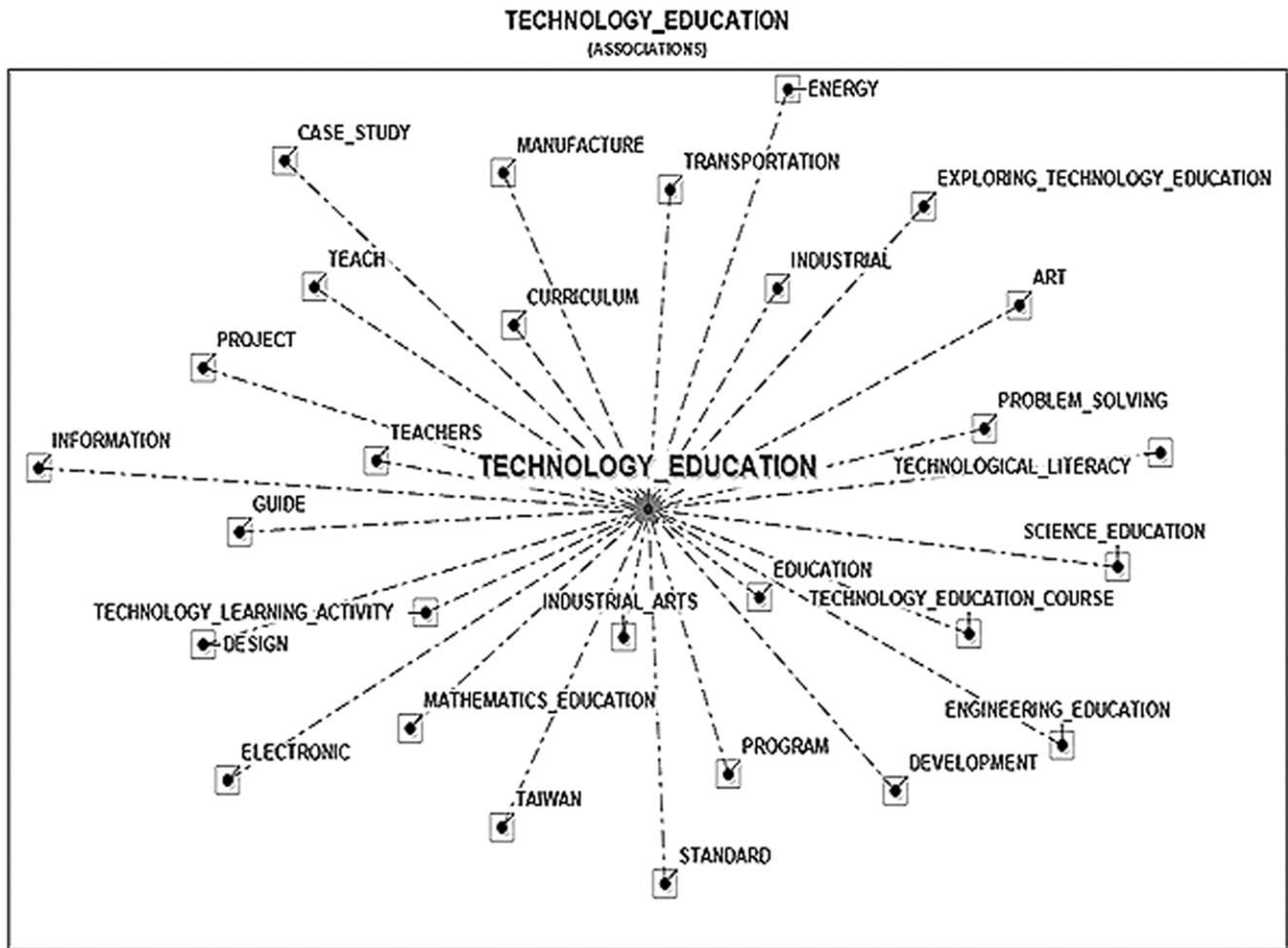


FIG. 4A. Word associations between “technology education” and other lemmas in the titles corpus subset.

corpus subsets, each having 7265 documents/records, and what other key terms co-occurred with them. For each of the selected key terms in the center, the other terms that share co-occurrence are distributed around it at a distance proportional to its degree of association resulting in a one-to-one relationship to the central key term. This means the key terms closer to the center co-occur most frequently with the term selected in the center. In addition, because TF-IDF is used as key term selection method, the algorithm evaluates the weight of a term within a document.

The total corpus was subdivided into three subsets, that is, title corpus, author-supplied keywords and descriptors corpus, and abstracts corpus. As shown in the figures below (Figures 3a–6c) for each of the four components in STEM, a visual map was generated to show the other key terms that co-occurred with the selected term in the center. This was repeated for each of the three corpus subsets. The results are presented in the order of the fields in the acronym STEM. The T-LAB software initially returns a map that has only 20 co-occurring key terms related to the term selected in the center. We used the add item feature in the software to

include terms we deemed appropriate to the map without unduly overcrowding the visualization map and, as a result, the number of key terms in Figures 3a through 8d vary.

Science Education. As shown in the word association map below, while Figure 3b (based on the author-supplied keywords and descriptors corpus) presents a better indicator of core knowledge areas and relationships (including those traditional areas commonly understood as belonging to the broader field of science, namely biology, chemistry, and physics), the title and abstract corpus subsets (Figures 3a and 3c, respectively) show very broad relationships. This is expected because not only are titles generally broad in nature, authors tend to use fewer specific terms and/or phrases in their titles than in keywords and/or abstracts. According to the extant literature retrieved from the two databases (ERIC and WoS) and based on the word association analyses, what can be safely said is that much of the discourse involves teaching science to students at elementary, precollege, college, and university levels. Also, evaluation, assessment, national policy, and reform topics dominate the map. In addition, distinct knowledge areas that

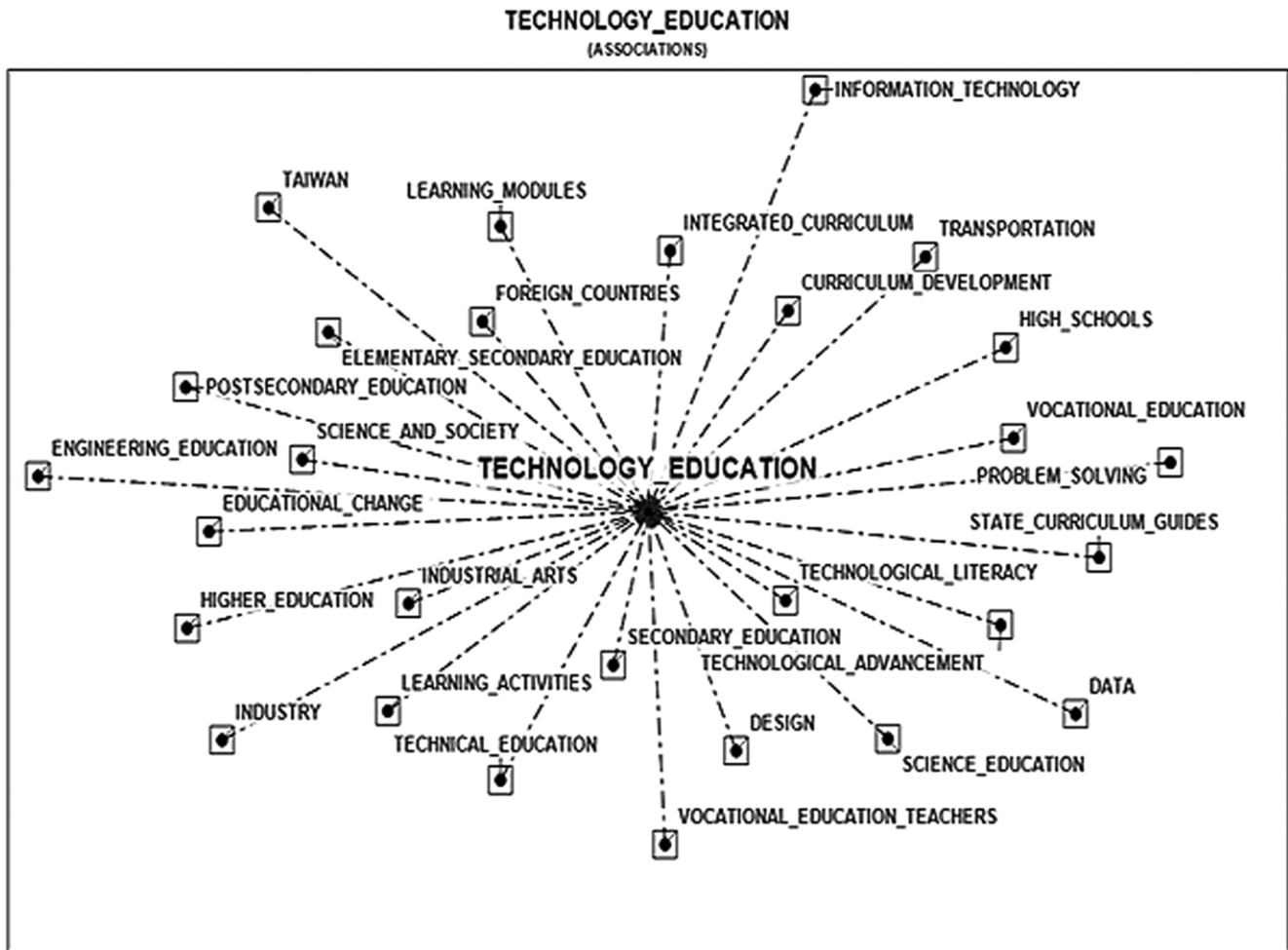


FIG. 4B. Word associations between “technology education” and other lemmas in the author-supplied keywords/descriptors corpus.

co-occurred with “science education” include biology, chemistry, physics, technology education, environmental education, and mathematics education. For instance, results of the analysis of the author-supplied keywords and descriptors corpus (Figure 3b) shows better identifiable science subject areas including biology, chemistry, physics, mathematics, environment, and other scientific activities and science concepts and concept formation. Other visible and prominent co-occurrences indicate the emphasis given to teacher education and research activities in science education. The science curriculum topic spans elementary science to secondary and college-level science (Figure 3b). It is also possible to gather from these maps (Figures 3a and 3c) that “science education” is related to inquiry and inquiry-based activities and this is probably an indication that based on the discourse in the literature, there exists a closer association between science education and inquiry based activities.

Technology education. Following a similar approach to “science education,” an analysis of the word association for “technology education” produced the visualization maps

shown in Figures 4a, 4b, and 4c. Unlike the maps for “science education,” in the case of “technology education” we observe the presence of a consistent set of knowledge and skill areas across the three corpus subsets. The core knowledge and skill areas co-occurring and related to “technology education” include industrial arts, design, electronic, Internet, information technology, vocational education, energy, manufacture, and transportation. Other related themes across the three maps include engineering education, technological literacy, mathematics education, and standards. A facet of the maps for “technology education” that is similar to those for “science education” is that it spans all educational levels (from K–12 to postgraduate), albeit with both “science education” and “technology education” having a closer association with elementary education and secondary education than others. There also seems to be an emphasis on literacy (e.g., as evidenced by the concept technological literacy).

One other key term that frequently co-occurred with “technology education” across the three corpus subsets (titles, keywords and descriptors, and abstracts) is the

TECHNOLOGY_EDUCATION
(ASSOCIATIONS)

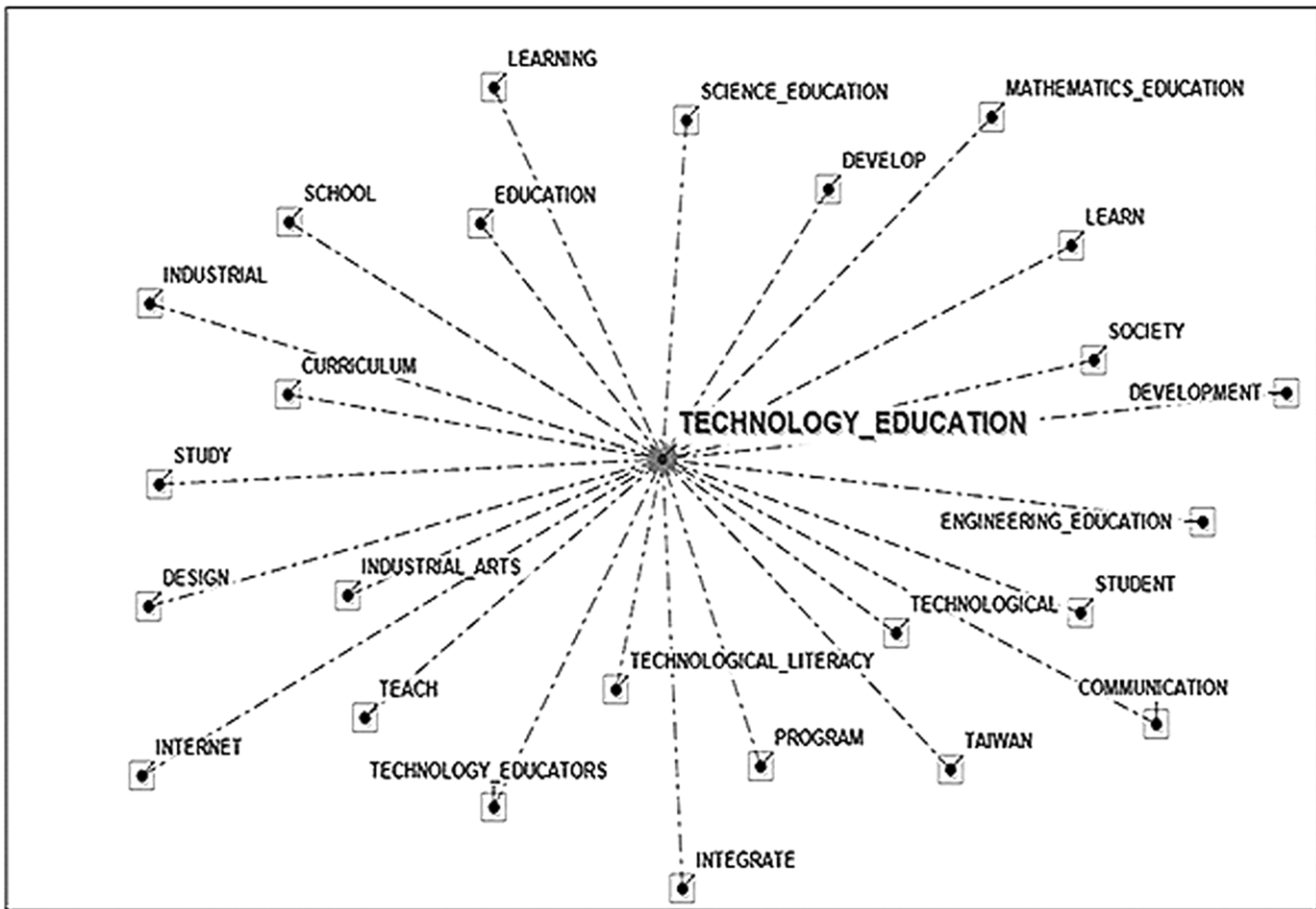


FIG. 4C. Word associations between “technology education” and other lemmas in the abstract corpus subset.

country name “Taiwan.” To identify the portions of text within which the word “Taiwan” occurs, a concordance analysis was conducted and the following contextual units were extracted (this is one instance out of the thousands of records in the corpus):

“In *Taiwan*, technology education used to be called “industrial arts,” but was changed to “living technology” in 1997. The goals of craftwork in elementary school are to provide an understanding in the areas of presentation, appreciation, and practical application of the arts. Curriculum content is divided into the three areas.”

The selected contextual unit above indicates why the phrase “industrial arts” is present as one core knowledge area in the word association maps for “technology education” (Figures 4a, 4b, and 4c). In addition, this text segment provides a plausible explanation in regard to the association between “industrial arts” or “technology education” and the country “Taiwan.”

Engineering education. The results of word association for “engineering education” are shown in Figures 5a,

5b, and 5c. Core knowledge areas that co-occur alongside “engineering education” include civil engineering, chemical engineering, electrical engineering, software, electronic, and technical education. One distinguishing characteristic of these word association maps for “engineering education,” compared to maps for the above two terms (“science education” and “technology education”), has to do with the core knowledge areas.

The core knowledge areas identified are all present in the context of undergraduate, graduate, college, and university levels. There is no mention of elementary or secondary schools as was the case with “science education,” “technology education,” and “mathematics education” (see Figures 6a, 6b, and 6c). In addition, technology education, science education, and professional education are the other core knowledge areas related to “engineering education.” Another distinguishing characteristic of the word association maps for “engineering education” is its closer association with continuing and professional education as well as two-year colleges.

ENGINEERING_EDUCATION
(ASSOCIATIONS)

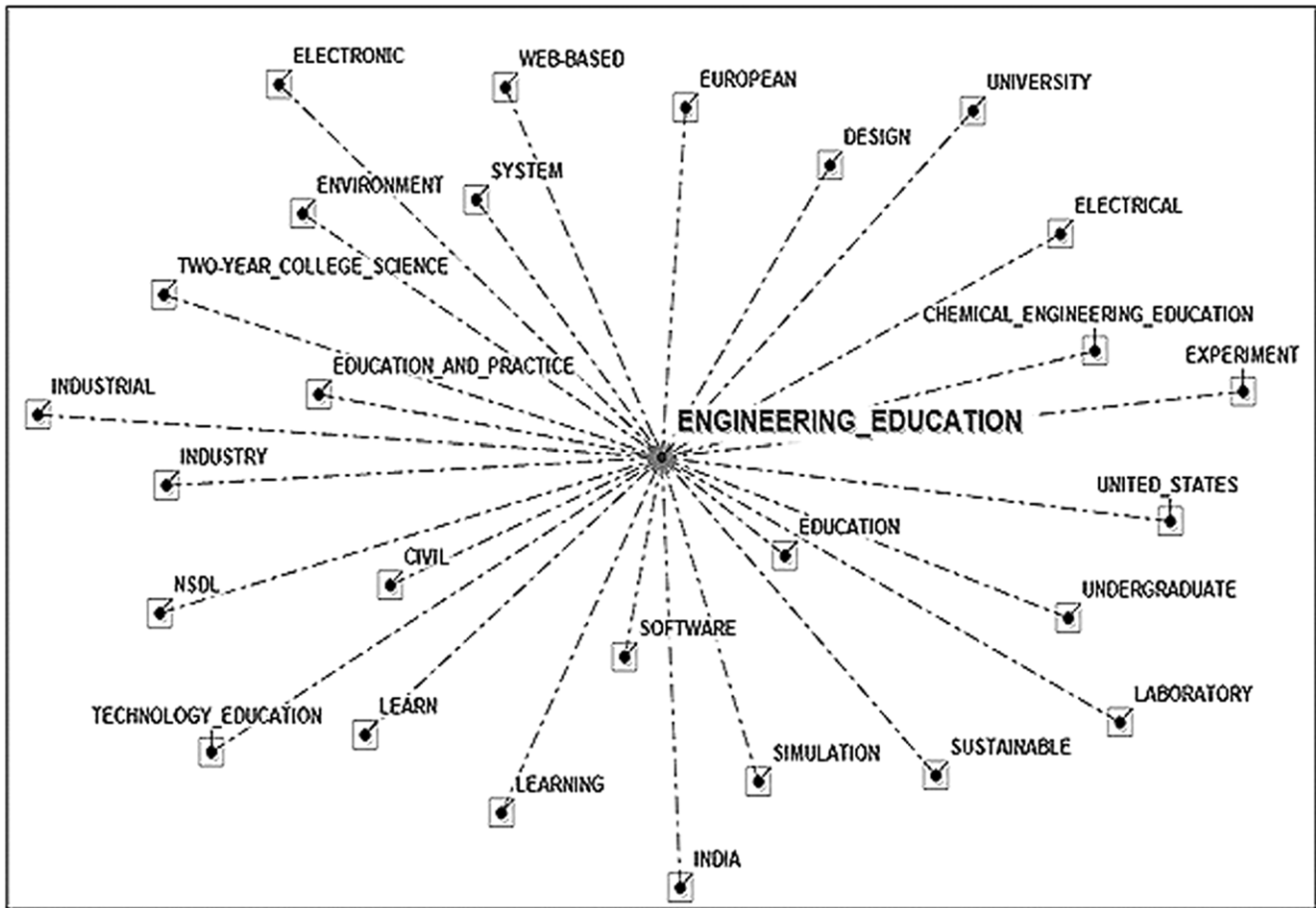


FIG. 5A. Word associations between “engineering education” and other lemmas in the titles corpus.

Mathematics education. Mathematics, the last of the STEM fields in the acronym, is not quite the megafield the other three are. As a result, its subfields are narrower than the respective subfields of the other three. This fact comes across in the word association maps. For instance, the core knowledge areas identified and related to “mathematics education” include algebra, geometry, calculus, number concepts, computer and computation, problem solving, mathematical concepts, and mathematical skills. The presence of these core knowledge areas manifests across the educational levels, starting from elementary to secondary and college-level mathematics, with the notable exception of early childhood education. Moreover, science education, technology education, and themes revolving around reform or educational change together with the National Council of Teachers of Mathematics (NCTM) co-occur with “mathematics education.”

To complement the above visual maps, a table containing the ten most frequently occurring lexical units and lemmas is presented (Table 2). The analyses of the three corpus sets resulted in a list sorted by the number of occurrences. For the titles and keywords and descriptors subsets, given that the corpus sizes were small, the threshold values were set to 10; for abstracts it was set to 70 because abstracts generated the largest number of occurrences. At the threshold value of 70, a total of 906 key terms were selected that together occurred 275,000 times. The top ten most frequently occurring lemmas (key terms) are shown in the table for the three corpus subsets. The percentages were calculated based on the total frequency of all lemmas in each corpus.

One important result that is evident from the above table and the word association maps is the finding that in the extant STEM literature, “science education” dominated the discourse and “engineering education” is the knowledge area least discussed.

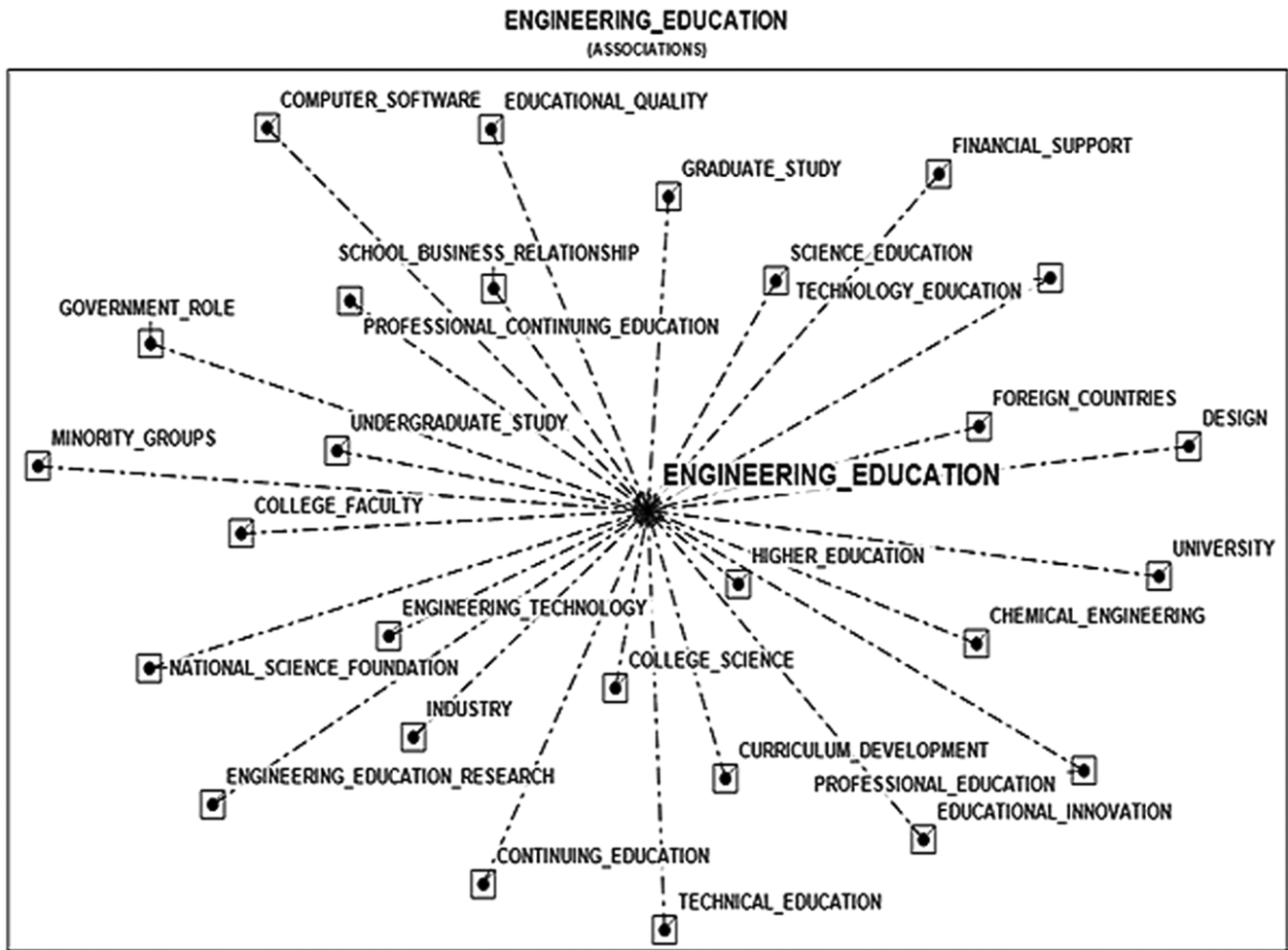


FIG. 5B. Word associations between “engineering education” and other lemmas in the author-supplied keywords/descriptors corpus.

Research Question #2: What can be learned from the visualizations and maps for the purpose of curriculum development?

This study primarily relied on a corpus of text retrieved from the Web of Science and ERIC databases on the topic of STEM education. Because it is difficult (if not impossible) to go through thousands of such works manually and find meaningful relationships, a statistical and linguistic content analysis software program (T-LAB) was used to gain knowledge and learn about the underlying structure of STEM and STEM education. Accordingly, to address this research question, we selected the lexical unit “curriculum development” to conduct word association analysis to see what other key terms co-occur with it. The same approach was followed as above to see how the key lexical unit “curriculum development” is related to other key terms across the three corpora: titles, keywords and descriptors, and abstracts. Figures 7a, 7b, and 7c are visual maps that show the relationships and co-occurrences of other

significant key terms associated with the chosen key term, “curriculum development.” Taking a closer look at these word association maps, the first observation we made was that “curriculum development” was closely related to all four components in STEM education, that is, science education, technology education, engineering education, and mathematics education, which means most of the scholarly works analyzed discuss STEM education within the context of curriculum development.

In addition to this obvious association, it can be safely argued that many of the key terms surrounding “curriculum development” in the visual maps across the three corpora reveal the need for a statewide and/or national map, guide, and/or framework for program development and curriculum evaluation. For elementary, secondary, and university levels, the other recurring themes include improving, changing, or transforming instruction, teaching methods, course materials, or course content. Again the author-supplied keywords and descriptors seem to offer a better visual map and richer associations. Also, teacher education, including

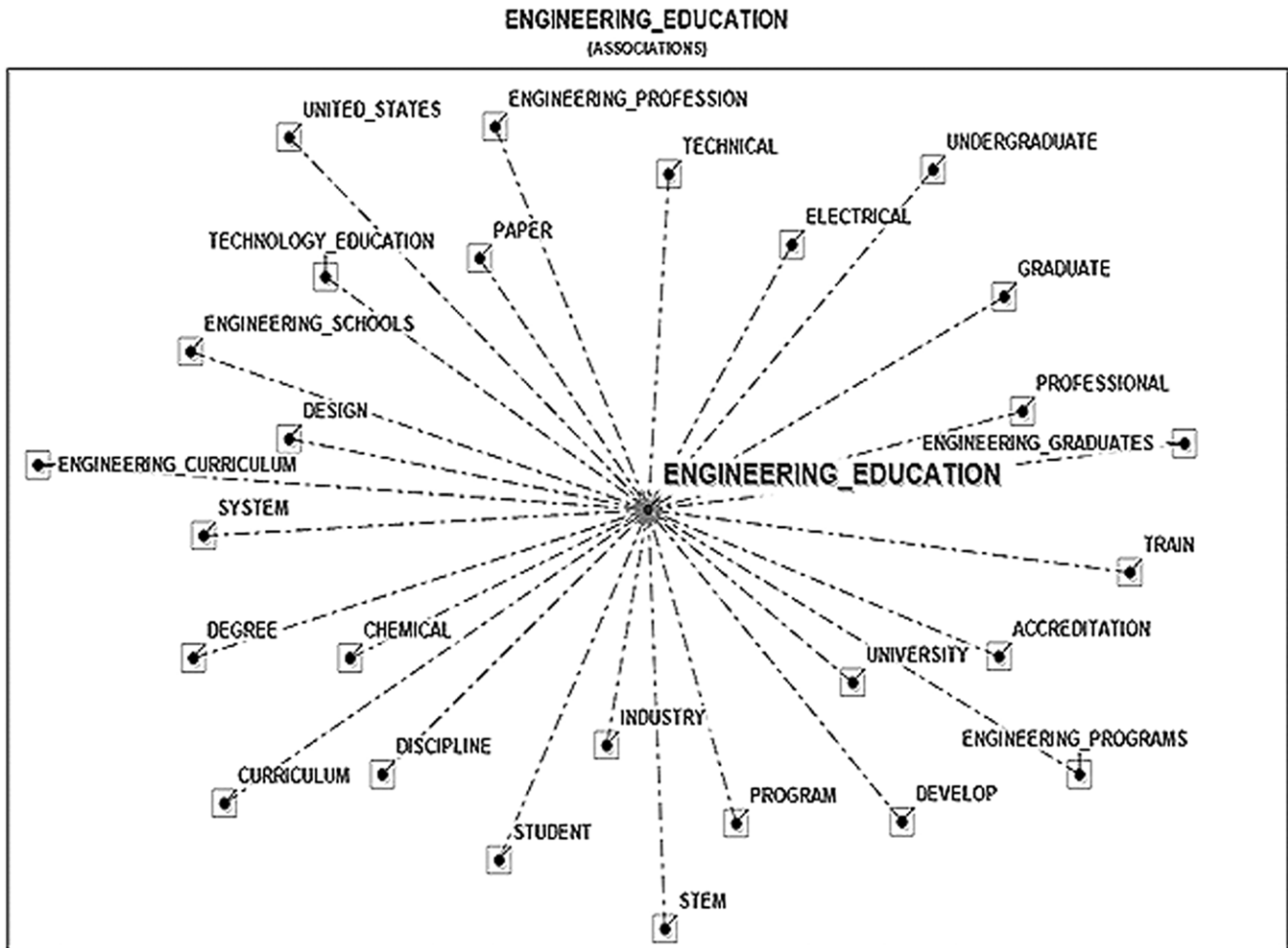


FIG. 5C. Word associations between “engineering education” and other lemmas in the abstract.

in-service education and concepts such as integration and integrated curriculum, are closely associated with “curriculum development.”

Research Question #3: What knowledge can be gained from the visualization maps for policymakers and other stakeholders in research as well as resource management and professional development activities?

As shown above, the author-supplied keywords and descriptors subset resulted in much better visualizations. To answer this research question, we used the keywords and descriptors corpus alone. We extracted key themes from the question statement itself to create the visual maps. Accordingly, from the key terms list generated by T-LAB, appropriate lexical units that helped answer the question were selected. The selected lexical units were “National Science Foundation” as a key institution and stakeholder fostering policy and supporting STEM activities (at least in the United States), “professional development” and/or “teacher education” (in-service and preservice), “policy formation,” and

“resource allocation.” These selected multiword key terms were generated from the analysis of the keywords and descriptors corpus by the software, and our role as researchers was to determine which of the key terms from the list offered a plausible explanation to the question by conducting further analyses (i.e., word association that results in visual maps). Figures 8a, 8b, 8c, and 8d present the result of the word association analyses.

Figure 8a shows the word association map for “National Science Foundation,” and from the co-occurring key terms in the map it can be easily observed that the extant STEM literature (body of knowledge) from which the corpus and the result were obtained depicts this U.S. federal agency as an institution providing financial support, grants, budgets, and federal aid. In addition, the map also shows the association between NSF and policy formation and federal legislation in science programs, natural sciences, and even social sciences across elementary, secondary, and higher education levels. It can also be safely concluded that the role of the government in building partnerships with schools and schools partnering with businesses is present in the

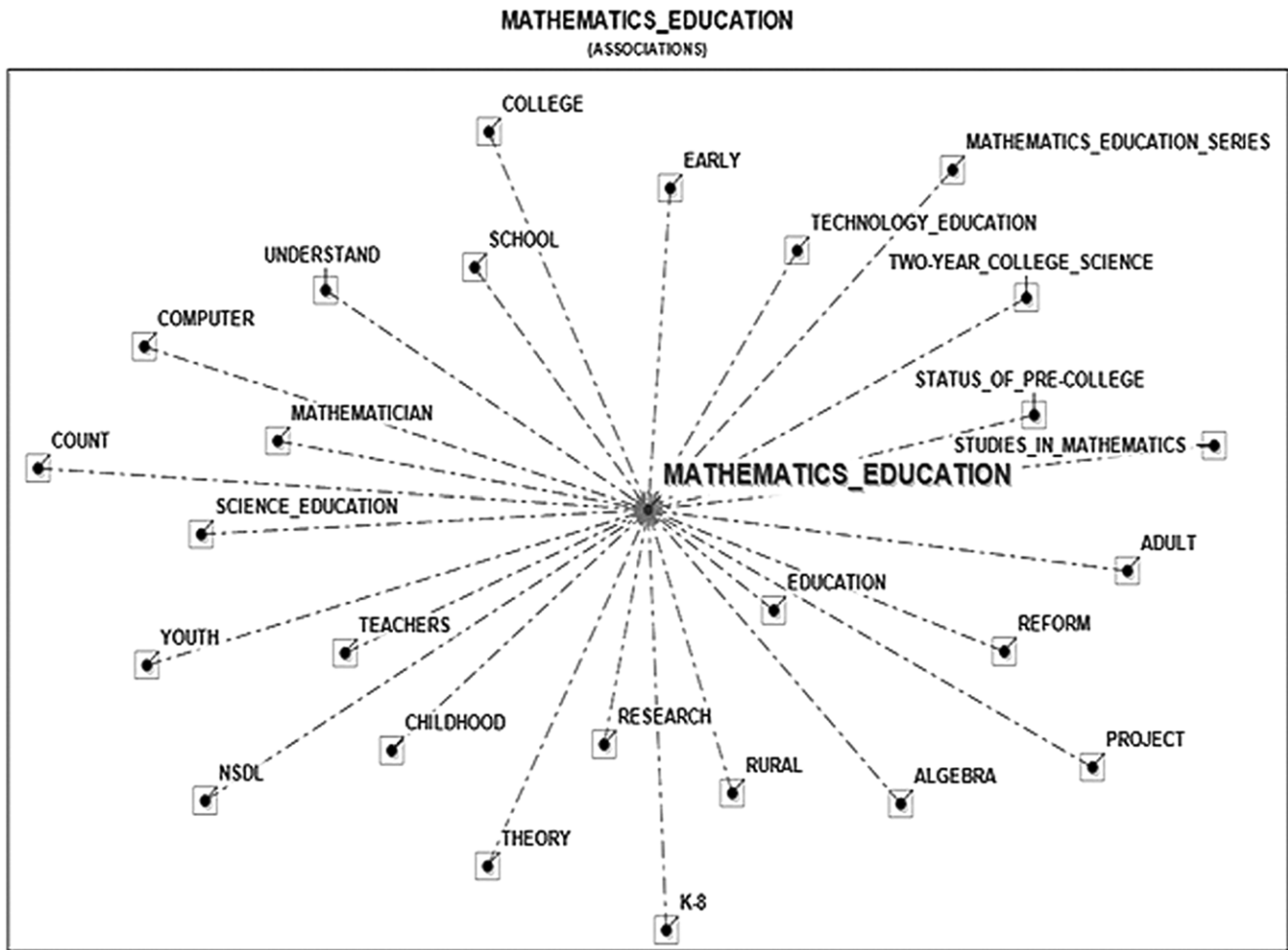


FIG. 6A. Word associations between “mathematics education” and other lemmas in the titles corpus.

discourse as evidenced by the multiword key terms generated in the vocabulary list GOVERNMENT_SCHOOL_RELATIONSHIP and SCHOOL_BUSINESS_RELATIONSHIP, respectively. In relation to the fields in STEM education, the map shows relationships between NSF and all but one component (technology education).

In Figure 8b, the word association for “professional development” shows the importance of mentoring (MENTOR) and teacher education, both in-service and pre-service education, for program and teacher effectiveness. Other visible themes in the map include cooperation and building knowledge bases for teaching (KNOWLEDGE_BASE_FOR_TEACHING) for science teachers and beginning teachers, again at elementary, secondary, and higher education levels. It also appears professional development activity is geared more toward science teachers and mathematics teachers.

When it comes to policymakers and key stakeholders, Figure 8c, the map for “policy formation,” shows closer co-occurrence relationships with “public policy” and “educational policy.” We also observe NSF closely co-occurring

with “policy formation” and other significant key terms that include advisory committees, program administration, curriculum evaluation, curriculum problems, educational quality, and educational improvement. Although we are exercising caution not to make generalized statements, the key terms that co-occur with the selected phrase, that is, “policy formation,” are indicative of the discourse in the extant published literature in STEM fields as far as the knowledge that can be gained for policymakers and stakeholders. This is one that highlights the need for policy matters and the role of the government through its agencies such as NSF to improve the quality of science education.

Another significant point in the question above (RQ#3) is “resource management.” Although we did not find the exact key term “resource management” in the list of multiword key terms, we chose the closest key term generated in the analysis, that is, “resource allocation.” Selecting this as a proxy to find more relationships and find explanations to the question of what can be learned from the extant literature on STEM education about resource management, Figure 8d offers yet another important insight. As always, our

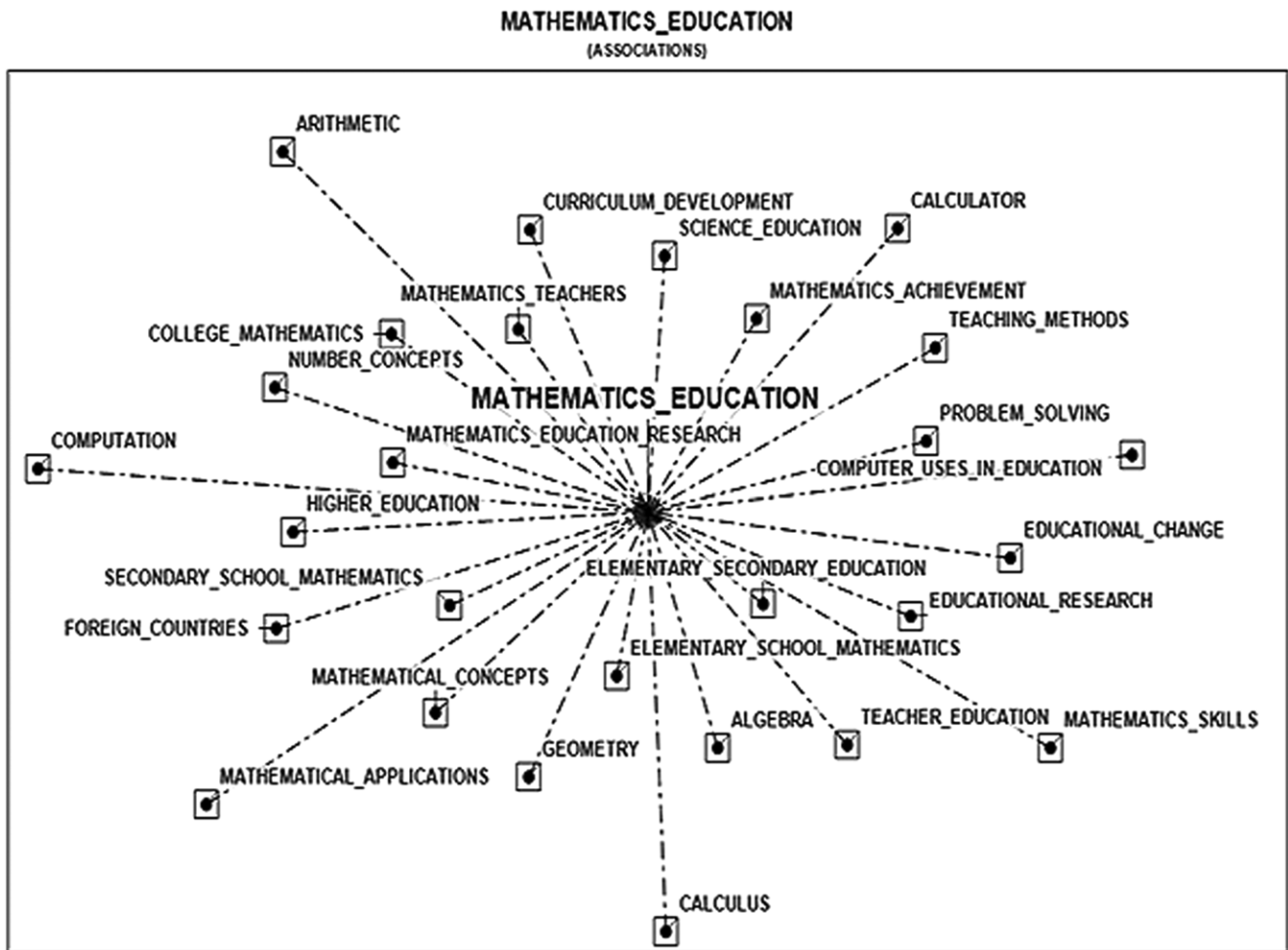


FIG. 6B. Word associations between “mathematics education” and other lemmas in the author-supplied keywords/descriptors corpus.

interpretation of the visual map starts from the key terms in the map that are closer to the selected lexical units in the center because the closer the line is to the center, the higher the frequency of association. Likewise, we observe “resource allocation” being closely related to literacy, educational facilities, financial policy, program administration, teacher salaries, community colleges, two-year colleges, state and federal aid, budget, statewide planning, and systemic change. The fact that the key term “rural areas” is closely related to “resource allocation” may be an indication of the need to expand the availability of resources to rural schools or rural areas, based on the STEM literature analyzed in this study.

Discussion

Scientometric and bibliometric methods are some of the ways, if not the only ways, through which we gain insight into the state and nature of fields or disciplines (Small, 1981). Our current work is unique in that although the history of the use of scientometric and bibliometric methods

and tools in the study of STEM fields dates back to the early days of scientometrics and bibliometrics itself (e.g., see Garfield, 1955, 1963a, 1963b), the uses of co-word analysis and word association maps are either new or few and far between. We have to note that some have used maps to study, among others, international collaboration by researchers in the sciences (e.g., Schubert & Braun, 1990) and global relationships and linkages between various disciplines (e.g., Small & Garfield, 1985). Small and Garfield’s (1985) and Noyons and Van Raan’s (1993) works are those that we could find that partially or most closely match our work.

Although Small and Garfield’s (1985) study was an attempt to map research in the sciences and social sciences with the help of citation data from the Science Citation Index/Social Sciences Citation Index for a single year, 1983, we set out to uncover the underlying structure of STEM education and the relationships between its core areas as well as draw implications with respect to curriculum and policy development using citation data that span several years. On the other hand, Noyons and Van Raan (1993)

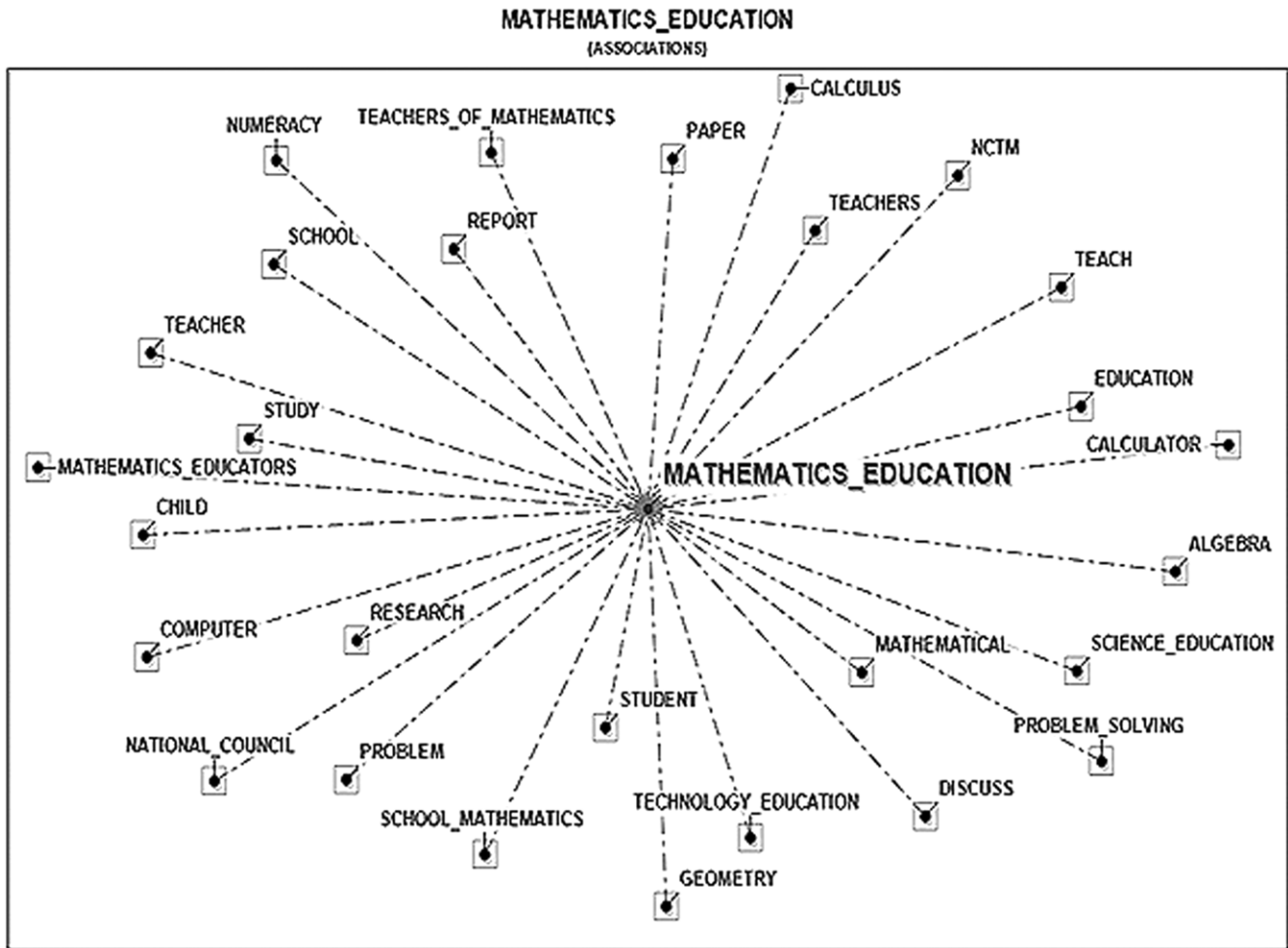


FIG. 6C. Word associations between “mathematics education” and other lemmas in the abstract corpus.

TABLE 2. The ten most frequently occurring lemmas (lexical roots) in titles, author-provided descriptors/keywords, and abstracts of STEM education publications.

Titles			Keywords			Abstracts		
Lemma	Freq.	%	Lemma	Freq.	%	Lemma	Freq.	%
Education	3,473	9.63	Science Education	4,431	7.03	Science Education	6,322	2.29
Science Education	2,091	5.80	Mathematics Education	2,603	4.13	Student	4,995	1.81
Technology Education	1,101	3.05	Elementary Secondary Education	1,921	3.04	Technology Education	4,536	1.65
Mathematics Education	887	2.46	Higher Education	1,593	2.52	Mathematics Education	4,118	1.49
Engineering Education	846	2.34	Technology Education	1,342	2.13	Education	4,104	1.49
Report	378	1.04	Curriculum Development	1,278	2.02	Program	2,884	1.04
Research	376	1.04	Foreign Countries	1,046	1.66	Study	2,670	0.97
School	266	0.73	Engineering Education	972	1.54	Engineering Education	2,413	0.87
Student	251	0.69	Teaching Methods	930	1.47	Research	2,271	0.82
Study	228	0.63	Educational Change	825	1.30	Curriculum Development	2,224	0.80
Total	9,897	27.41	Total	16,941	26.84	Total	36,537	13.23

produced maps for a narrower field in technology, optomechatronics, based on a set of technology patents, rather than on published documents such as journals and proceedings. Our work and similar works that use maps to study science

follow the tradition popularized, if not started, by DeSolla Price (1965).

More significantly, we were able to establish the structure as well as the themes of the scientific discourse in STEM

CURRICULUM_DEVELOPMENT
(ASSOCIATIONS)

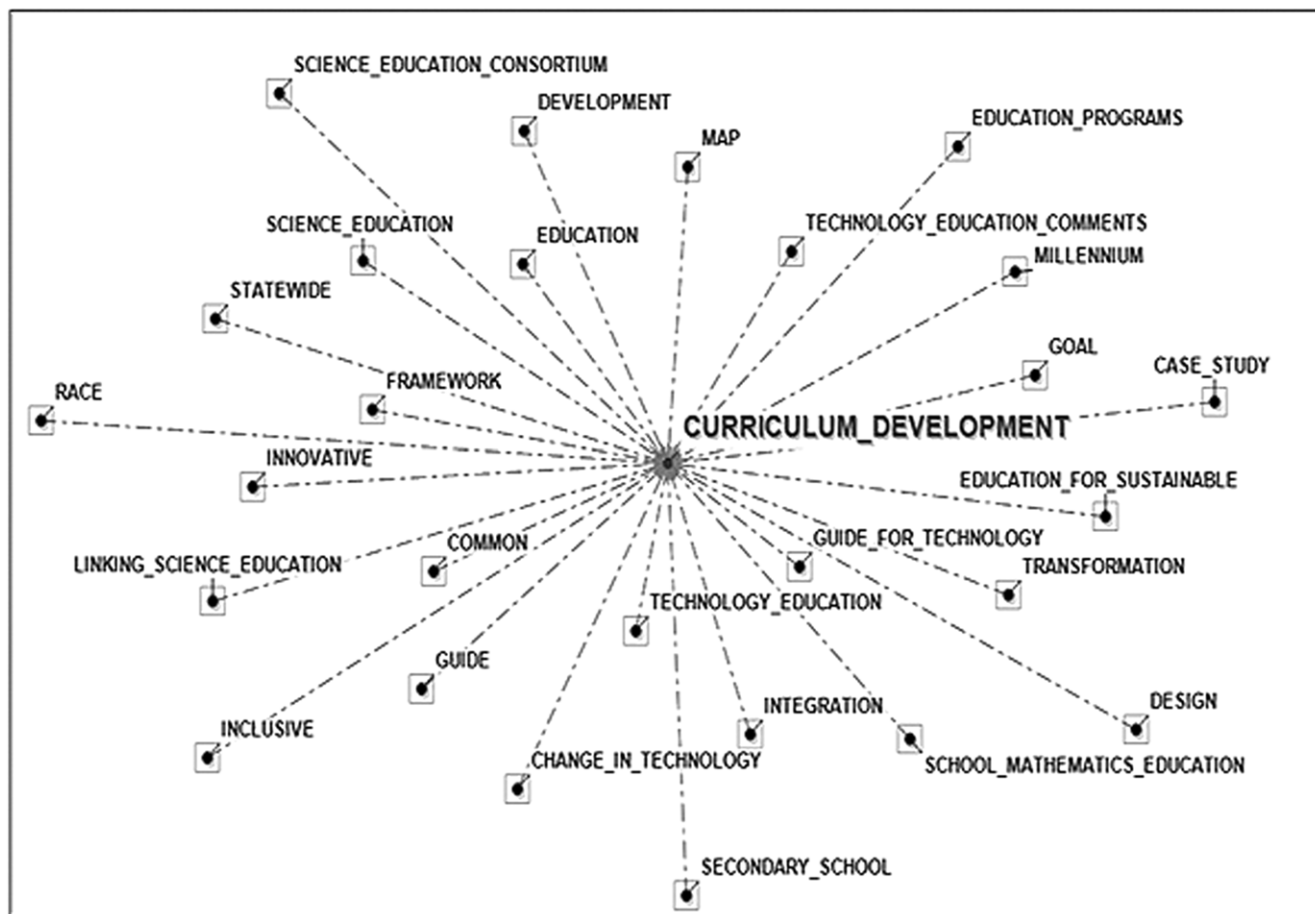


FIG. 7A. Word associations between “curriculum development” and other lemmas in the titles corpus.

fields. Our analyses as well as findings are consistent with the theoretical accounts that words and key terms, both single and multiword key terms, collectively represent the content of the text and that they can be analyzed statistically while maintaining semantic quality (Zhang, Yoshida, & Tang, 2011). The fact that we used TF-IDF as a key term selection method (hence the emphasis placed on the importance of a term in a document) and the advanced option is checked for multiword (hence the contextual information of individual words is captured) (see Figure 2), and later the multiword list is used to transform the entire corpus increases the reliability of co-occurrence analysis to reveal the structure of the scientific literature and, in turn, the fields under investigation.

In both the word association maps (Figures 3a through 8d) and the table showing the ten most frequently occurring key terms across the three corpus sets (Table 2), we were able to gain a better understanding of what authors have been writing about and discussing in the domain of STEM and STEM education. Because the goal of this study was to use the extant literature retrieved from two appropriate

databases, Web of Science and ERIC, as empirical data to understand the domain of STEM and STEM education, the key components of STEM education, that is, science education, technology education, engineering education, and mathematics education, were selected for analyses. In addition, key terms that help explain research questions 2 and 3 were selected; these were curriculum development, National Science Foundation (NSF), professional development, policy formation, and resource allocation. The discussion below sheds some light on major findings about the three research questions.

Of the three vocabulary sources for analysis, the author-supplied keywords and descriptors offered a better account of and served as better representations of the underlying structure of the literature in STEM education as shown in the visual maps in Figures 3b, 4b, 5b, 6b, 7b, 8a, 8b, 8c, and 8d. In an effort to map national research efforts world-wide, a study that analyzed more than 1.2 million works from Social Sciences Citation Index (part of WoS) had found the structure of the social and behavioral sciences together with developments in these disciplines using author-provided

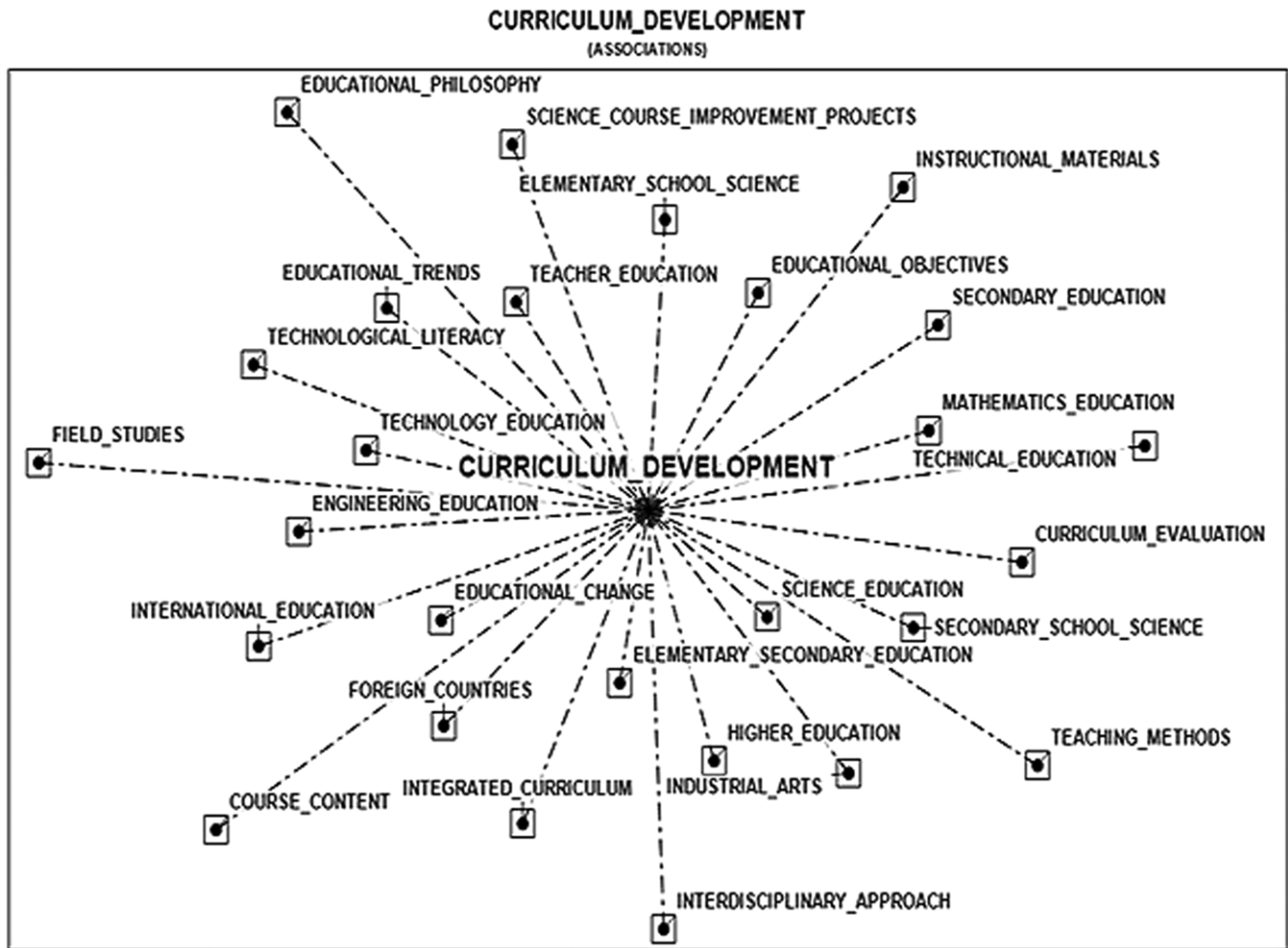


FIG. 7B. Word associations between “curriculum development” and other lemmas in the author-supplied keywords/descriptors corpus.

keywords (Nederhof & Van Wijk, 1997). It is also worth noting that the combination of author-supplied keywords and the subject descriptors from the databases performed better than the widely reported issues arising from the use of indexer assigned descriptors only, which is often referred to as the “indexer effect,” (Leydesdorff & Zaal, 1988).

Core Knowledge Areas

Given the fact that the acronym STEM combines four major disciplines that require investigations in their own right, the first research question and results of our analyses helped us understand what STEM education is and what knowledge areas are involved in the field. By analyzing thousands of scholarly sources in the various STEM fields, we were able to present bibliometric maps that show the underlying structure and core knowledge areas that emerged in the discourse. For each of the component areas in STEM education, we generated a number of bibliometric maps (word association maps) that showed the relationships between co-occurring key terms.

As broad a field as it is, in “science education” results of the analyses revealed topics and themes that revolve around teaching, assessment, and reform issues of science education at the elementary, high school, precollege, college, and university levels. The core knowledge areas that frequently co-occurred with “science education” were biology, chemistry, physics, mathematics, and inquiry-based activities.

In “technology education” we found a structure that is represented by knowledge areas such as industrial arts, design, electronic, Internet, information technology, vocational studies, technology literacy, engineering, and mathematics. In the United States, it is often the case that job training policies by educational institutions (such as community colleges), federal agencies, and administrations focus more on manufacturing and energy independence (or green energy technologies) and jobs in these sectors. Hence, the association of words such as “energy” and “manufacture” with “technology education” can be partly explained. With respect to maps for “technology education,” a surprising association was “problem solving” (at least for

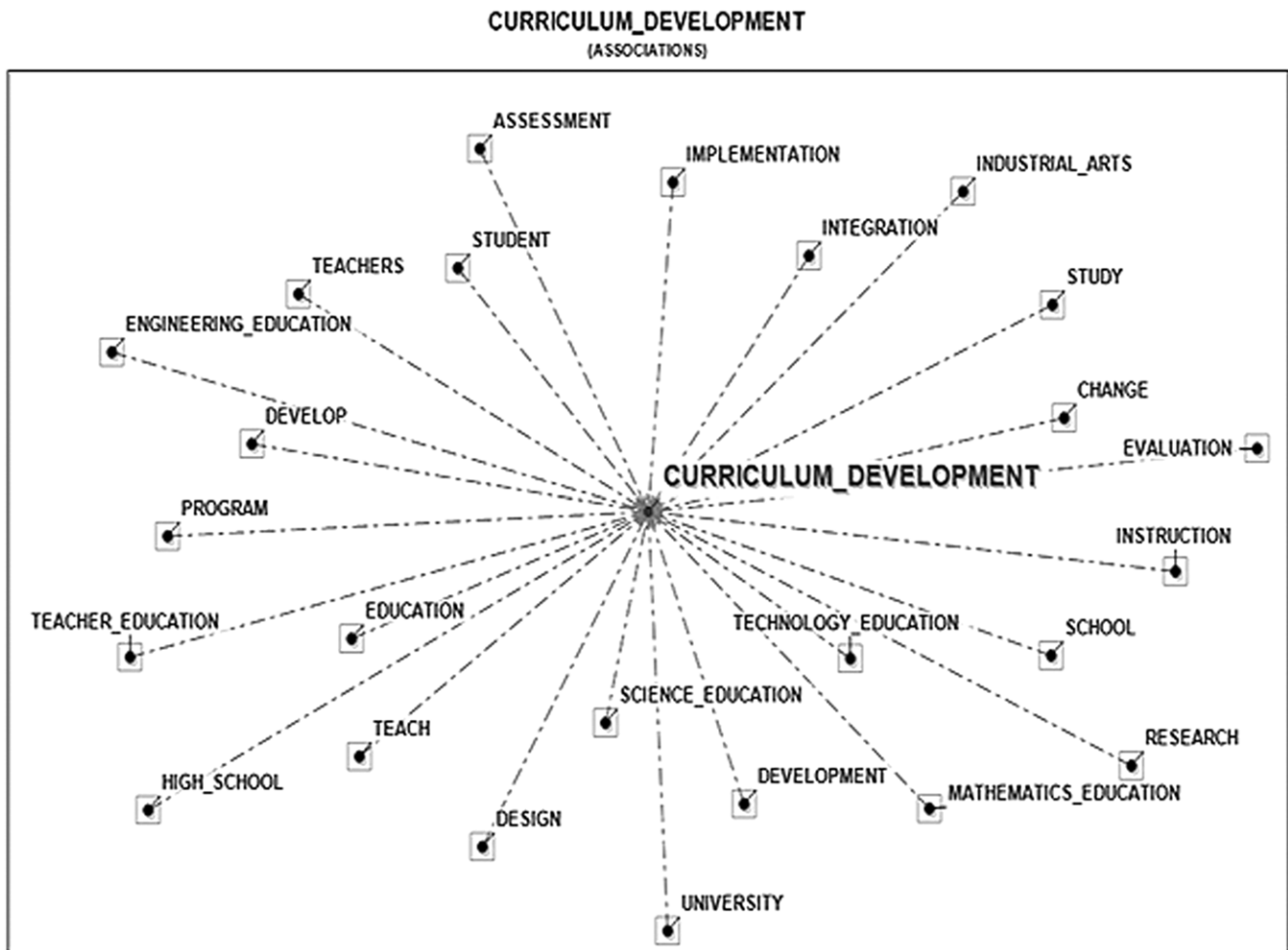


FIG. 7C. Word associations between “curriculum development” and other lemmas in the abstract corpus.

the title and keywords and descriptors corpus, Figures 4a and 4b) which, as should be expected, is a term that also had a closer association with “mathematics education.” One explanation is that technology is sometimes thought of as a tool for solving practical and work-related problems.

Regarding “engineering education,” the findings showed structures where all the core knowledge areas identified are closely associated with college and university-level education and/or instruction. From this finding, it can be safely concluded that “engineering education” is not well developed in elementary and precollege level curricula, although engineering is one of the four components of STEM education. The core knowledge areas that emerged in the word association map include civil engineering, chemical engineering, electrical engineering, software, electronic, and technical education.

The last component in STEM, “mathematics education,” is present across the educational levels. Unlike the other three components, mathematics education is detected from early childhood to the elementary, secondary, and college levels. The most significant themes underlying the

discourse as represented by the bibliometric maps include algebra, geometry, calculus, number concepts, computer and computation, problem-solving, mathematical concepts, and mathematical skills. In terms of interrelationships between the four STEM components, mathematics education appears to be a more visible or dominant knowledge area closely related to the other components in STEM education.

Based on the co-occurrence of pairs of significant words, the method used in this analysis, that is, co-word analysis, helped us extract key themes and knowledge areas and establish relationships between and among such themes. As discussed above, although the analysis of the corpus and text revealed a basic structure that represents the domain of STEM education, the findings equally show very general discussions related to policy, standard, reform, change, and improvement issues in STEM education from elementary through university levels. Most of the core knowledge areas identified were obvious, except the level of focus of “engineering education,” which was at the college and university levels.

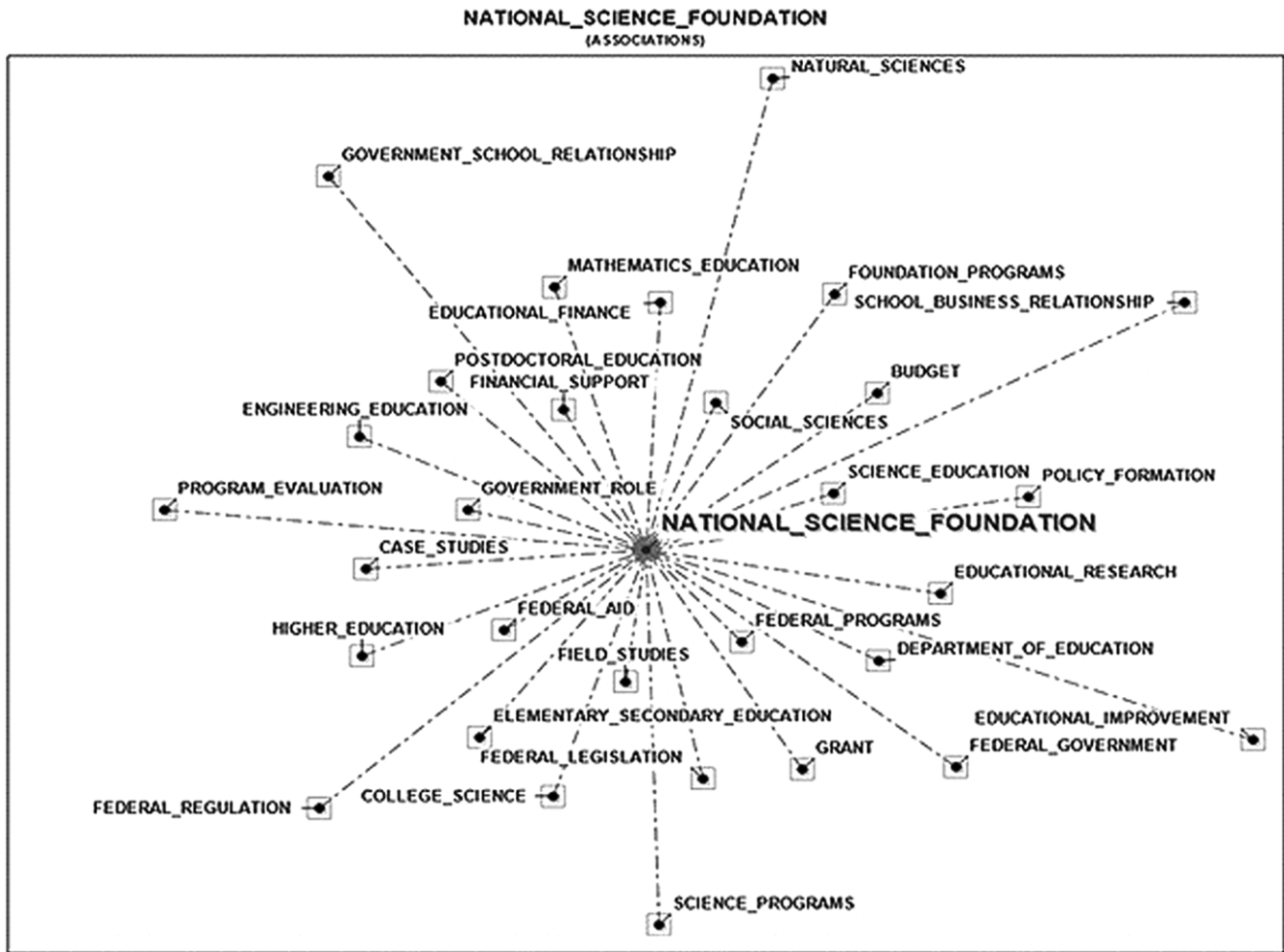


FIG. 8A. Word associations between “National Science Foundation” (NSF) and other lemmas in the author-supplied keywords/descriptors corpus.

Lessons for Curriculum Development

In addition to identifying core knowledge areas, we also investigated the body of literature to see if any lesson can be learned for curriculum development. As predicted, the text analysis identified “curriculum development” as one of the most frequently occurring multiwords. As shown in Table 2, “curriculum development” is the sixth most frequently occurring term in the author-supplied keywords and descriptors subset and the tenth in the abstracts subset. Although Table 2 presents only the ten most frequently occurring terms across the three subsets, “curriculum development” is the 12th most frequently occurring multiword in title corpus subset. These findings show the significance attached to “curriculum development” in the STEM education literature. Furthermore, by analyzing what other key terms co-occur with curriculum development, we were able to find a consistent result across the three subsets, one that signals the need for state and/or national maps, guides, and frameworks for curriculum development.

Knowledge Gained for Policymakers and Stakeholders

After examining the vocabulary list that was generated in the corpus preprocessing step, we identified key terms that potentially offer explanations to research question #3. These key terms include National Science Foundation (NSF), professional development, policy formation, and resource allocation. The NSF emerged as a key stakeholder in making financial resources available for research and professional and program development in the STEM area. Obvious results were also observed that suggest close associations between the NSF and grant, budget, and federal aid. Equally, as a federal agency that helps shape policy matters, interesting themes emerged that indicate the presence of relationships between “government and school” and “business and school.” Multiwords such as GOVERNMENT_SCHOOL_RELATIONSHIP and SCHOOL_BUSINESS_RELATIONSHIP were identified by the text analysis software as co-occurring with NSF.

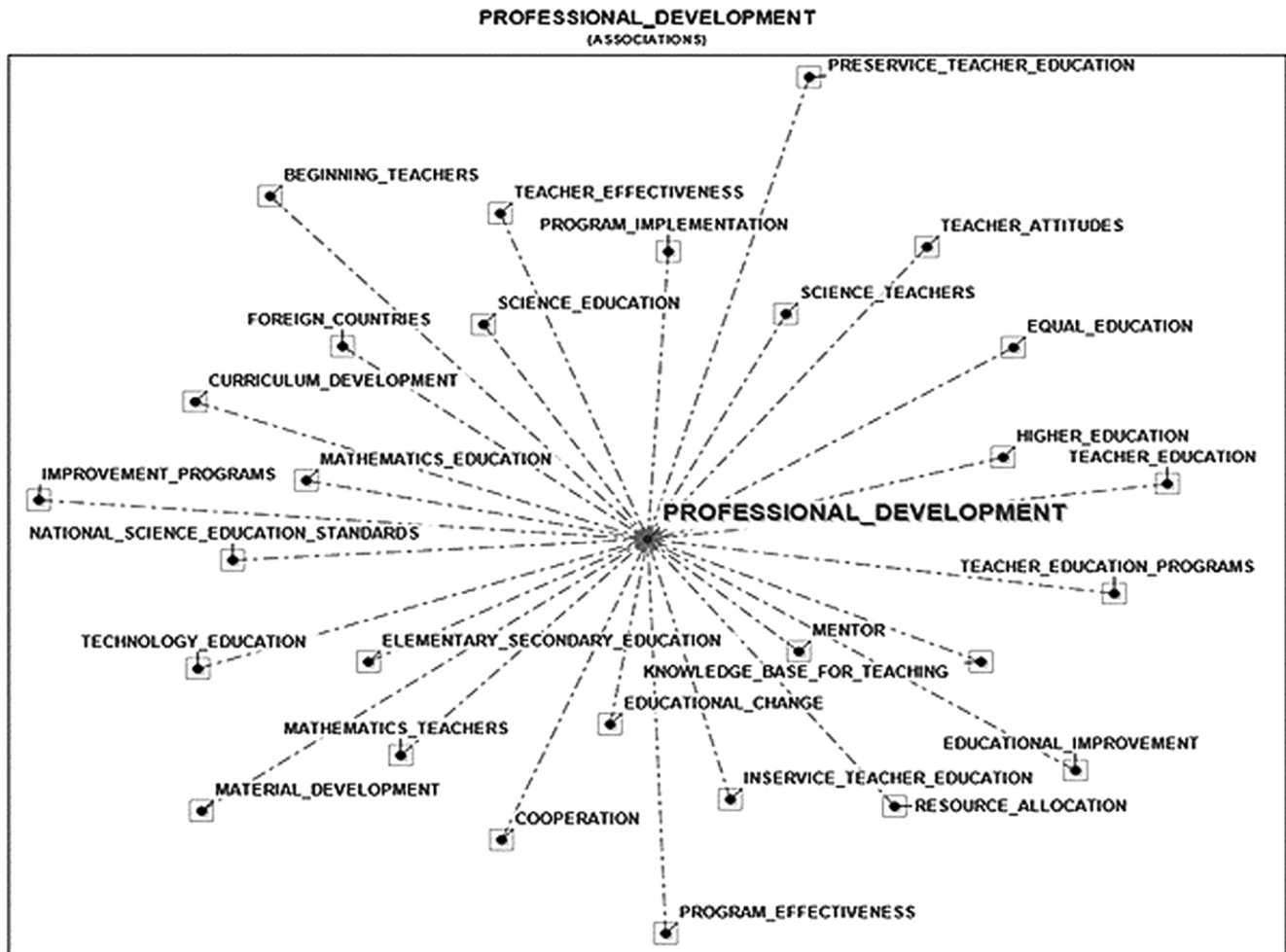


FIG. 8B. Word associations between “professional development” and other lemmas in the author-supplied keywords/descriptors corpus.

Regarding professional development, “mentor” was identified as the most frequently co-occurring key term. It could be argued that this is an obvious finding that does not require reporting. However, the important fact is that we used a body of scholarly output or literature in STEM education as empirical data and were able to corroborate some of the obvious associations reported in this study, “professional development” and “mentoring” being one example. Another multiword that was identified by the analysis is “knowledge base for teaching.” It is apparent that this suggests the presence of a discussion in the STEM education literature that associates professional development and the knowledge base for teaching. Regarding “resource management,” close associations between “resource allocation” and educational facilities, financial policy, program administration, teacher salaries, community colleges, and rural areas were observed.

In regards to knowledge gained for policy makers, the closest key term selected was “policy formation,” and as shown in Figure 8c, the most frequently co-occurring key terms include public policy, educational policy, advisory committees, evaluation needs, program administration,

educational quality, government role, moral issues, organizational objectives, federal legislation, institutional role, and federal programs. From this map, it can be concluded that the presence of public and educational policy in conjunction with institutional and federal roles to evaluate and improve the quality of science education is required.

Conclusion

Bibliometric mapping of scholarly outputs in a given domain is often used to chart the fields and knowledge areas and help establish the underlying structure (Chen, 2004). Following a similar approach, this study investigated the domain of STEM education to understand the core knowledge areas and what can be inferred from the emerging structure for important activities, including curriculum development, policy formation, professional development, and resource management. In view of the complexity of the domain of STEM education, the significance of this study lies in the core knowledge areas that it identified. Although most of the knowledge areas identified are predictable, an important implication is the absence of

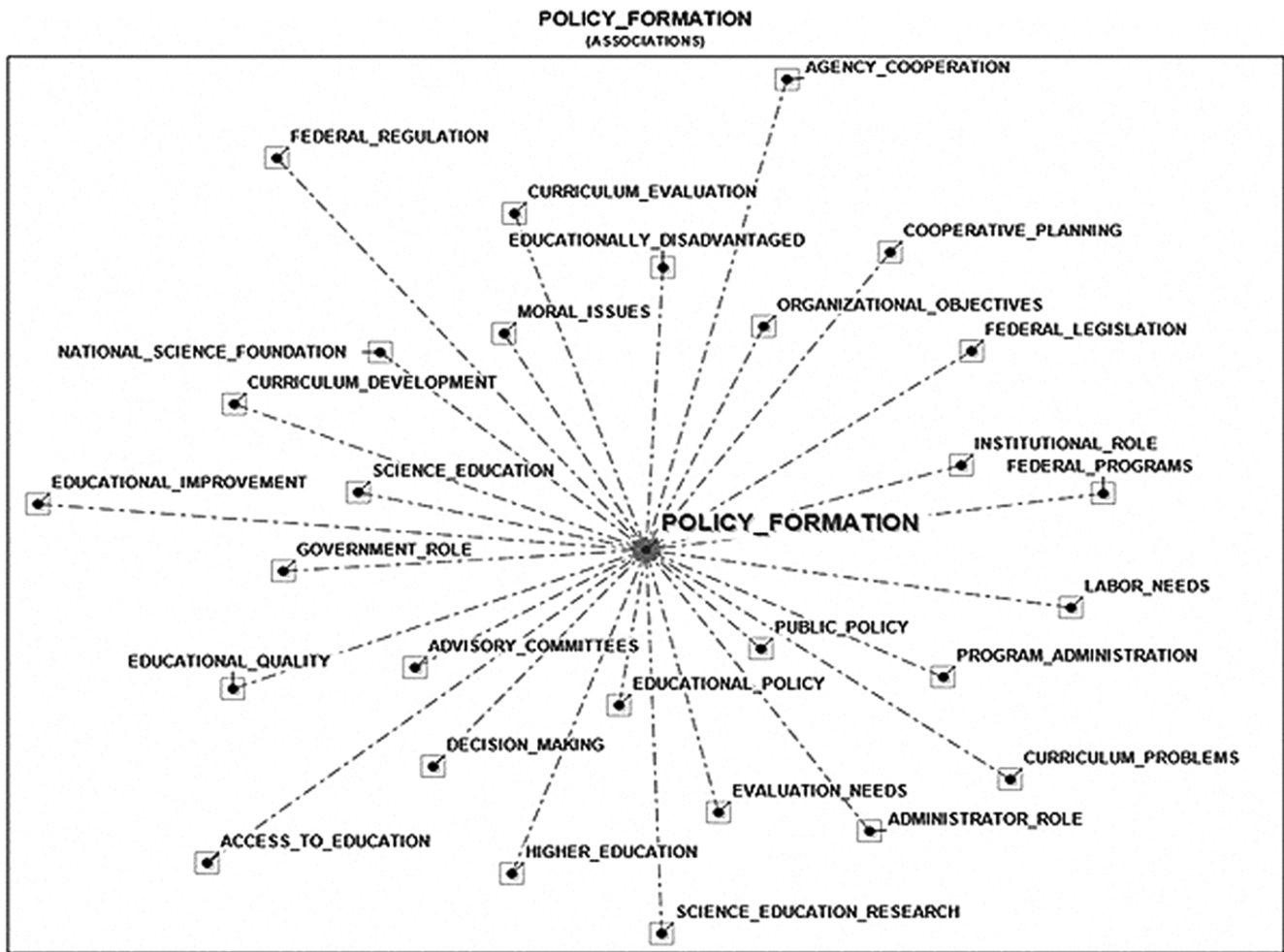


FIG. 8C. Word associations between “policy formation” and other lemmas in the author-supplied keywords/descriptors corpus.

discourse on “engineering education” at the elementary and secondary education levels. The lack of programs and content on engineering education in K-12 studies has been documented in the literature (Rockland, Bloom, Carpinelli, Burr-Alexander, Hirsch, & Kimmel, 2010). It is still important to develop a clear typology to chart the domain so as to understand what actually constitutes STEM education. In regard to curriculum development, the most frequently co-occurring key terms, such as guides, maps, and/or frameworks, suggest the focus area that needs to be addressed to support STEM-related activities. In addition, the underlying structure of the STEM field as represented by the bibliometric maps show the presence of policy, financial resource, mentoring, and teacher education in both urban and rural settings.

Limitations of Current Work

This study is based on text analysis of the existing literature in STEM education fields. After preprocessing, cleaning, and corpus normalization, the analysis built a multiword

vocabulary list together with the frequency of occurrences for each of the key terms in the list. Further analysis was applied to the vocabulary list using a word association method and association index (similarity coefficients). This determined the co-occurrence of pairs of key terms across the more than seven thousand documents and records. Although the bibliometric maps make it easy to inspect the nature of structure and relationships among the key terms, it is difficult to make detailed inferences. Clearly, the key terms are surrogates for the titles, keywords, and abstracts from which they are extracted. However, it is often important to go back and confirm the contextual units within which the key terms occur, similar to what we did in the case of the country term “Taiwan” and technology or industrial arts. An attempt was made not to make sweeping generalizations and this is because of the limitations of the vocabulary in the text corpus as a complete representation of STEM education. Future studies can create corpus subsets by introducing a temporal variable to see trends and developments in curriculum development, policy, and other matters over time. T-LAB offers multiple correspondence analysis that

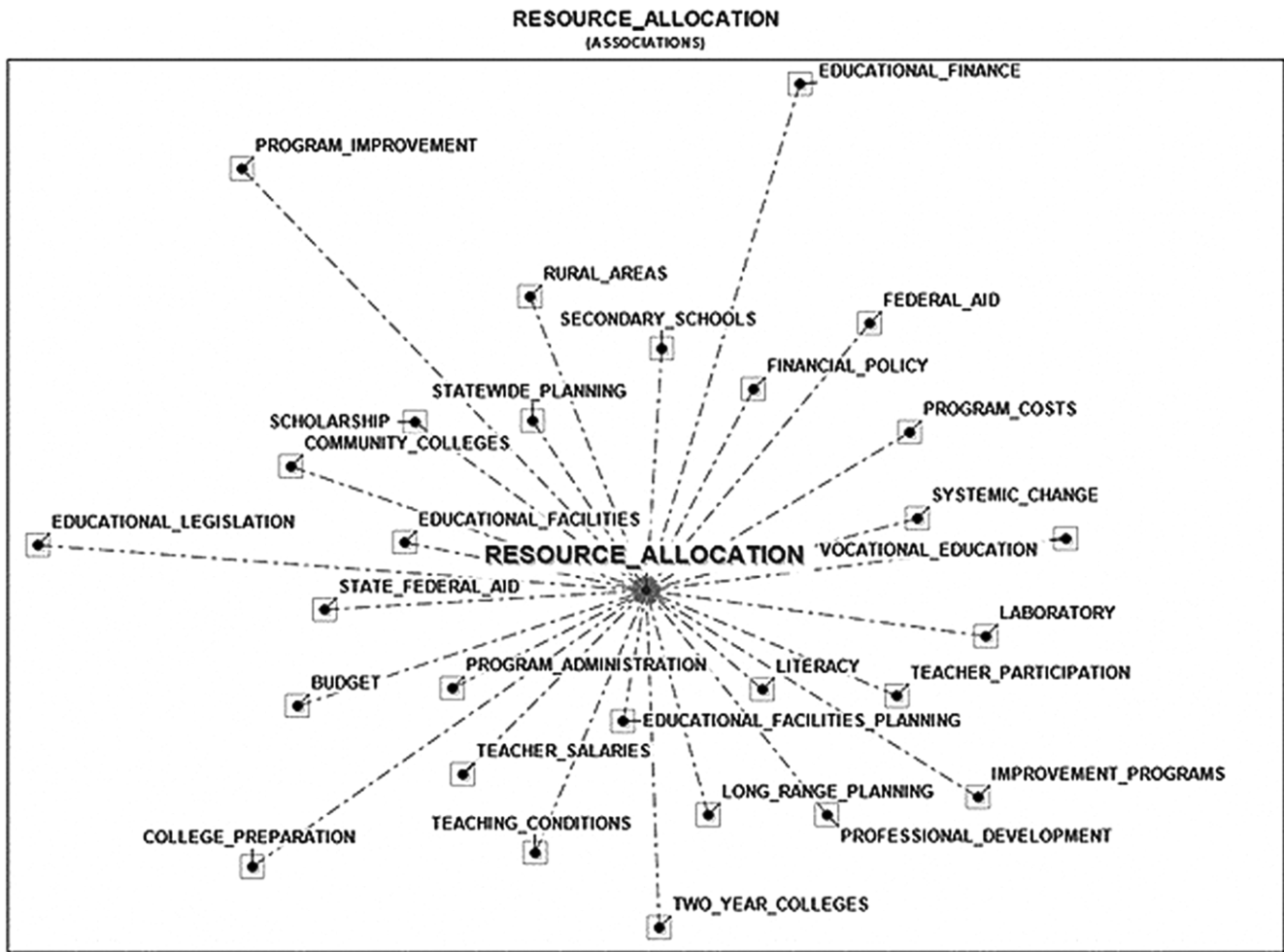


FIG. 8D. Word associations between “resource allocation” and other lemmas in the author-supplied keywords/descriptors corpus.

help gauge the relationships between two or more categorical variables.

Acknowledgments

We are grateful to Dean Gregory Anderson (Morgridge College of Education) for making funds available to purchase T-LAB software. Our sincere gratitude also goes to the anonymous reviewers for their comments.

References

- Agarwal, P., & Searls, D.B. (2009). Can literature analysis identify innovation drivers in drug discovery? *Nature Reviews Drug Discovery*, 8(11), 865–878. doi: 10.1038/nrd2973
- Althouse, B.M., West, J.D., Bergstrom, C.T., & Bergstrom, T. (2009). Differences in impact factor across fields and over time. *Journal of the American Society for Information Science and technology*, 60(1): 27–34. doi: 10.1002/asi.20936
- American Association for the Advancement of Science (1989). Project 2061: Science for all Americans. Washington, DC: AAAS.
- An, Y., Janssen, J.C.M., & Milios, E.E. (2002). Characterizing the citation graph as a self-organizing networked information space. In H. Unger, T. Bohme, & A.R. Mikler (Eds.), *Proceedings of Second International Workshop on Innovative Internet Computing Systems (IICS'02)* (pp. 97–107). Kùhlungsborn, Germany: Springer-Verlag.
- Borgman, C.L., & Furner, J. (2002). Scholarly communication and bibliometrics. *Annual Review of Information Science and Technology*, 36, 3–72.
- Bornmann, L., Mutz, R., & Daniel, H.-D. (2009). Do we need the h index and its variants in addition to standard bibliometric measures? *Journal of the American Society for Information Science and technology*, 60(6): 1286–1289.
- Breiner, J.M., Harkness, S.S., Johnson, C.C., & Koehler, C.M. (2012). What is STEM? A discussion about conceptions of STEM in education and partnerships. *School Science and Mathematics*, 112: 3–11. doi: 10.1111/j.1949-8594.2011.00109.x
- Buter, R.K., Noyons, E.C.M., Van Mackelenbergh, M., & Laine, T. (2006). Combining concept maps and bibliometric maps: First explorations. *Scientometrics*, 66(2), 377–387.
- Bybee, R.W. (2010). Advancing STEM education: A 2020 vision. *Technology & Engineering Teacher*, 70(1), 30–35.
- Callon, M., Law, J., & Rip, A. (1986). *Mapping the dynamics of science and technology: Sociology of science in the real world*. London: The Macmillan Press.
- Chen, C. (2004). Searching for intellectual turning points: Progressive knowledge domain visualization. *Proceedings of the National Academy of Sciences, USA*, 101 (Suppl. 1) (pp. 5303–5310).
- Courtial, J., & Law, J. (1989). A Co-word study of artificial intelligence. *Social Studies of Science*, 19(2), 301–311.

- Crouch, D., Irvine, J., & Martin, B.R. (1986). Bibliometric analysis for science policy: An evaluation of the United Kingdom's research performance in ocean currents and protein crystallography. *Scientometrics*, 9(5), 239–267.
- DeSolla Price, D.J. (1965). Networks of scientific papers. *Science*, 149(3683), 510–515.
- Franklin, J.J., & Johnston, R. (1988). Co-citation bibliometric modeling as a tool for S&T policy and R&D management: Issues, applications, and developments. In A. F. J. van Raan (Ed.), *Handbook of quantitative studies of science and technology* (pp. 325–389). Amsterdam: North-Holland.
- Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of ideas. *Science*, 122(3159), 108–111.
- Garfield, E. (1963a). Citations in popular and interpretive science writing. *Science*, 141(3579), 392.
- Garfield, E. (1963b). *Science Citation Index*. *Science Citation Index*, 1, v–xvi.
- Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). Highlights from TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth- and Eighth-Grade Students in an International Context. Washington, DC: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.
- Green, M. (2007). *Science and engineering degrees: 1966–2004 (NSF 07-307)*. Arlington, VA: National Science Foundation.
- He, Q. (1999). Knowledge discovery through co-word analysis. *Library Trends*, 48(1), 133–159.
- Healey, P., Rothman, H., & Hoch, P.K. (1986). An experiment in science mapping for research planning. *Research Policy*, 15, 233–251.
- King, J. (1987). A review of bibliometric and other science indicators and their role in research evaluation. *Journal of Information Science*, 13(5), 261–276.
- Leydesdorff, L., & Zaai, R. (1988). Co-words and citations relations between document sets and environments. In L. Egghe and R. Rousseau (Eds.), *Informetrics*, 87/88, (pp. 105–119). Belgium, Diepenbeek: Elsevier.
- Leydesdorff, L., & Wagner, C. (2009). Is the United States losing ground in science? A global perspective on the world science system. *Scientometrics*, 78(1), 23–36.
- Marshakova-Shaikovich, I. (2005). Bibliometric maps of field of science. *Information Processing and Management*, 41, 1534–1547.
- McCain, K.W. (1990). Mapping authors in intellectual space: A technical overview. *Journal of the American Society for Information Science*, 41(6): 433–443.
- National Academy of Sciences (2007). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, D.C.: The National Academies Press.
- National Commission on Excellence in Education (1983). *A nation at risk: The imperative for educational reform*. Washington, DC. Retrieved from http://datacenter.spps.org/sites/2259653e-ffb3-45ba-8fd6-04a024ecf7a4/uploads/SOTW_A_Nation_at_Risk_1983.pdf.
- National Science Technology Council. Committee on STEM Education (CoSTEM) (2011, December). *The Federal science, technology, engineering, and mathematics (STEM) education portfolio*. Retrieved from http://www.whitehouse.gov/sites/default/files/microsites/ostp/costem_federal_stem_education_portfolio_report.pdf.
- Nederhof, A.J., & Van Wijk, E. (1997). Mapping the social and behavioral sciences world-wide: Use of maps in portfolio analysis of national research efforts. *Scientometrics*, 40(2), 237–276.
- Noyons, E. (2001). Bibliometric mapping of science in a policy context. *Scientometrics*, 50(1), 83–98.
- Noyons, C. (2005). Science maps within a science policy context: Improving the utility of science and domain maps within a science policy and research management context. In H. Moed, W. Glanzel, & U. Schmoch (Eds.), *Handbook of quantitative science and technology research* (pp. 237–255). Dordrecht: Kluwer Academic Publishers.
- Noyons, E.C.M., & Van Raan, A.F.J. (1993). *Science base of technology: Bibliometric mapping as a tool for national science and technology policy: Part II: Optomechatronics*. CWTS Report 93-07, Leiden University.
- Noyons, E.C.M., Buter, R.K., & van Raan, A.F.J. (2002). Bibliometric mapping as a science policy tool. *Proceedings of the Sixth International Conference on Information Visualisation (IV'02)*, pp. 679–684. doi: 10.1109/IV.2002.1028848
- OECD (2011). *Lessons from PISA for the United States: Strong performers and successful reformers in education*. OECD Publishing. Retrieved from <http://dx.doi.org/10.1787/9789264096660-en>.
- Rafols, I., & Meyer, M. (2010). Diversity and network coherence as indicators of interdisciplinarity: Case studies in bionanoscience. *Scientometrics*, 82(2), 263–287.
- Roche, I., Besagni, D., François, C., Hörlesberger, M., & Schiebel, E. (2010). Identification and characterisation of technological topics in the field of Molecular Biology. *Scientometrics*, 82(3), 663–676.
- Rockland, R., Bloom, D.S., Carpinelli, J., Burr-Alexander, L., Hirsch, L.S., & Kimmel, H. (2010). Advancing the “E” in K-12 STEM education. *Journal of Technology Studies*, 36(1), 53–64.
- Sanders, M. (2009). *Integrative STEM education primer*. *The Technology Teacher*, 68(4), 20–26.
- Schubert, A., & Braun, T. (1990). International collaboration in the sciences, 1981–1985. *Scientometrics*, 19(1–2), 3–10.
- Small, H. (1981). The relationship of information science to the social sciences: A co-citation analysis. *Information Processing and Management*, 17(1), 39–50.
- Small, H., & Garfield, E. (1985). The geography of science: Disciplinary and national mappings. *Journal of Information Science*, 11(4), 147–159.
- Small, H. (1999). Visualizing science by citation mapping. *Journal of the American Society for information science*, 50(9), 799–813.
- Teaching Institute for Excellence in STEM (2010). *What is STEM Education?* Retrieved from <http://www.tiesteach.org/stem-education.aspx>
- The UTeach Institute (2012). *Replicating success*. Austin, TX: The UTeach Institute. Retrieved from <http://uteach-institute.org/replication>.
- van Eck, N.J., Waltman, L., Dekker, R., & Van den Berg, J. (2010). A comparison of two techniques for bibliometric mapping: Multidimensional scaling and VOS. *Journal of the American Society for Information Science and Technology*, 61(12), 2405–2416.
- Waltman, L., van Eck, N.J., & Noyons, E.C.M. (2010). A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics*, 4(4), 629–635.
- The White House (2009, November 23). President Obama launches “Educate to Innovate” campaign for excellence in Science, Technology, Engineering & Math (STEM) education. Washington, DC: Office of the Press Secretary.
- Whittaker, J.(1989). Creativity and conformity in science: Titles, keywords, and co-word analysis. *Social Studies of Science*, 19, 473–96.
- Zhang, W., Yoshida, T., & Tang, X. (2011). A comparative study of TF*IDF, LSI and multi-word for text classification. *Expert Systems with Applications*, 38(3), 2758–2765
- Zitt, M., & Bassecouard, E. (1994). Development of a method for detection and trend analysis of research fronts built by lexical or cocitation analysis. *Scientometrics*, 30(1), 333–351.

Copyright of Journal of the American Society for Information Science & Technology is the property of John Wiley & Sons, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.