



The emergence and evolution of ambiguous ideas: an innovative application of social network analysis to support systematic literature reviews

Thomas Cowhitt¹ · Joshua Travis Brown² · Anthony Lising Antonio³

Received: 25 April 2024 / Accepted: 20 August 2024
© The Author(s) 2024

Abstract

Systematic literature reviews are attempts to understand conversations between researchers working to develop solutions to common problems. These conversations often stretch back decades and can involve the participation of dozens of authors. Traditional approaches to systematic reviews are ill-equipped to make sense of the sheer volume of relevant literature when exploring the emergence and evolution of ambiguous ideas across large knowledge communities. This article presents three innovative applications of Social Network Analysis (SNA) methods to explore the emergence and evolution of accountability in higher education across a collection of 450 peer-reviewed articles published from 1974–2017 and their corresponding 12,270 references. First, qualitative data from articles and references were integrated into new interactive joint displays called Narrated Network Diagrams, creating opportunities to more accurately assess themes and meanings in literature by connecting structures in co-citation networks with relevant relational stories. Second, time was elevated in the analysis procedure to capture the dynamism of knowledge formation. Third, underutilized descriptive network statistics were applied to the co-citation network analysis to generate new insights such as different mechanisms for authors gaining influence in a knowledge community. Ultimately, this article presents an innovative longitudinal Mixed Methods Social Network Analysis (MMSNA) approach to systematic literature reviews, significantly advancing previous SNA methods integration in this critical research practice.

Keywords Systematic literature review · Social network analysis · Co-citation network · Accountability · Higher education · Education policy

Mathematics Subject Classification 97B40 Educational policy for higher education · 91D30 Social networks; opinion dynamics · 91C20 Clustering in the social and behavioral sciences

JEL Classification D83 Search Learning · Information and Knowledge · Communication · Belief · Unawareness · D85 Network Formation and Analysis: Theory · Z18 Public Policy

Extended author information available on the last page of the article

Introduction

An ambiguous idea rarely remains confined to a specific domain. Some, like free speech, values, and diversity are applied to broad swaths of society, becoming relevant to a wide array of citizens in many different contexts (Altomonte, 2020). The wide application of ambiguous ideas creates different emergent and evolutionary pathways for their development and use, resulting in a proliferation of literature and competing or even contradictory terminology. One such idea in American higher education is accountability, which since the 1970s has emphasized assessing value-added quality on the one hand and performance-based funding on the other (Brown, 2017, 2018). Developing insight into such ideas that diverge or coalesce within a community-like accountability in higher education—requires tools capable of untangling their lineage across a wide array of academic disciplines over several decades. Traditional sensemaking approaches are ill-equipped to systematically engage the sheer volume of literature emanating from researchers and practitioners working in these spaces. New methodological approaches are needed to better understand the emergence and evolution of such ideas, whose continued use impacts many different areas of society.

This article presents innovative applications of Social Network Analysis (SNA) methods for understanding the ascent of ambiguous ideas across knowledge communities. The application of SNA tools in systematic literature review protocols is still relatively new but part of a growing trend for researchers utilizing bibliometrics to analyze publication proliferation (Cowhitt et al., 2019). Tracing the rise of accountability as an influential concept within higher education required us to engage with 450 journal articles published from 1974 to 2017, along with the corresponding 12,270 references. This sizeable collection of literature represents a decades-long conversation between a myriad of stakeholders working in different contexts. It also represents diverse types of publication materials such as peer-reviewed articles, government reports, books, news media, and grey literature.

We present three innovative applications of SNA tools for helping researchers make sense of such large collections of literature. These innovations culminated in a new approach for systematic reviews, which will be referred to as a longitudinal Mixed Methods Social Network Analysis (MMSNA) approach. First, qualitative excerpts from articles and references were embedded into network diagrams to generate new interactive joint displays called Narrated Network Diagrams (NNDs) (Cowhitt et al., 2023). NNDs allow readers to call on text by clicking on or scrolling over a tooltip, which is a visual marker for readers to identify latent text. The integration of qualitative excerpts from literature into network visualizations creates opportunities to account for the relational dimensions of research and writing—an analysis approach we refer to as *tying structure to story*. The use of tooltips allows for the integration of qualitative excerpts from each piece of literature from a review into a social network visualization by embedding text into nodes, which represent articles and references, or into edges, which connect the articles and references based on a predefined and pertinent type of relationship. This data visualization method allows researchers to engage with the content of published literature more authentically, pushing beyond article titles and abstracts when trying to understand a collection of literature. Significantly, the use of NNDs and the ability to engage with excerpts surrounding the use of a reference in co-citation networks provides unique insights into the evolutionary journey of ideas and relevant terminology across knowledge communities of dozens or even hundreds of authors.

Second, the new literature analysis protocol makes use of publication dates to elevate time as a key variable. Time is an underutilized variable to support the cleaving of large networks and is especially relevant when attempting to identify the inception and track the evolution of ideas. We use an informed partitioning protocol, which we refer to as *policy pivot points*, to separate the co-citation network into discrete communities at five-year intervals based on events in education policy legislation in the United States. The resulting segments of network data, hereafter referred to as policy panels, take on further significance in the analysis and discussion of the data.

Lastly, we apply descriptive network statistics in new ways to advance the analysis of a large collection of literature. Specifically, we illustrate how a triad census, which accounts for the different variations of three-entity configurations across a network, can be used to identify influential authors and the genesis of innovative ideas within knowledge communities. The application of additional descriptive network statistics to aid with literature analysis is an intentional effort to consider a wider array of local structural configurations when developing narratives of publication outputs by a research community. Using triads to tie structure to story advances efforts to further deploy SNA methods to aid our understanding of literature development and proliferation.

Applications of social network analysis tools in systematic literature reviews

As fields of study mature, literature accumulates, providing a rich, though often murky landscape for researchers seeking to understand different knowledge communities. Alternative motivations and special interests can cause a proliferation of competing narratives that create challenges for those attempting to understand the evolution of influential ideas. This has led to researchers using bibliometrics, or the statistical analysis of publication metadata, to explore rapidly expanding collections of literature through systematic literature reviews (Godin, 2006; Rousseau, 2014).

A review of SNA reviews

To understand previous applications of SNA tools in bibliometrics, we conducted an interdisciplinary search of published literature reviews in Web of Science that utilized key word co-occurrence, co-citation, or co-authorship analyses. This Boolean search yielded 5024 publications. These results were then filtered for literature reviews and sorted by publication date. The resulting 585 literature reviews represents a review of SNA reviews (Pollock et al., 2023) from 127 different fields of research, the most prominent fields being Computer Science ($n=166$), Business Economics ($n=154$), and Environmental Sciences Ecology ($n=136$). Full-text copies of the reviews were compiled into a PDF file. A series of keyword searches were then conducted on the full-text collection to identify the application of mixed methods and longitudinal network analysis tools and techniques, the application of specific descriptive network analyses (i.e., community detection, density, triad, and census) and prominent network and qualitative analysis software (Rstudio, Bibliometrix, VOSviewer, Citespace, Visnetwork, and NVivo). After analyzing the context of each keyword use, four themes emerged regarding the use of SNA tools in systematic literature reviews. First, there is a need for SNA tools to support simultaneous engagement with both network diagrams and the qualitative content of publications in large literature reviews. Literature reviews infrequently examine the content of literature beyond the title and abstract. Second, longitudinal network analysis was infrequent and justification for the creation of

panels was rarely given when this type of analysis method was pursued. Third, there is a general lack of transparency when partitioning networks (i.e., communities, cliques, and clusters). Finally, there is an overreliance on dyadic descriptive network statistics (i.e., in- and out-degree centrality). We discuss each of these themes below.

The need for SNA tools to support content analysis

The application of SNA tools in bibliometrics has increased efficiency in important literature analysis tasks such as identifying significant authors, seminal pieces of literature, emerging trends, and the prominence of terminology, by providing new tools for exploiting publication metadata. The most commonly used tools are Citespace (Chen, 2006), VOSviewer (Van Eck & Waltman, 2010), and Bibliometrix (Aria & Cuccurullo, 2017). Applying SNA in literature review protocols prioritizes relationships, allowing easier analysis of knowledge communities and the evolution of ambiguous ideas. The use of SNA tools in literature searching and analysis is an important development because it can help take into account social dynamics behind academic output.

However, SNA tools have yet to be deployed in systematic literature reviews to help researchers engage in content analysis beyond keyword co-occurrence. Examining keyword co-occurrence networks is a helpful technique for identifying macro topical trends and at the micro-level, particularly popular (high degree) topics or high-strength topic-pairs (Radhakrishnan et al., 2017). Keyword co-occurrence networks, however, are reliant on pre-determined keywords in document metadata that may a priori structure the relational data in a limited manner. A persistent challenge is engaging with the full content of publications beyond keywords, titles, and abstracts. Researchers need to engage with qualitative content beyond the title and abstract to gain more than a cursory understanding of the evolution of ideas through time in a particular discipline.

Additional qualitative analysis tools like NVivo could be added to literature review protocols to allow researchers to engage with more qualitative content beyond titles, abstracts, and keywords. We identified only three instances of researchers incorporating qualitative analysis software like NVivo. Two of these instances deployed NVivo to examine keywords and abstracts (B. Raza et al., 2021; Ribeiro et al., 2022) and another used NVivo to analyze word clouds of article titles (Lloyd et al., 2016). However, engaging with qualitative data more efficiently is only part of the solution. The application of SNA in bibliometrics creates an opportunity for using interactive displays and other tools to integrate large amounts of qualitative data into network visualizations, providing important context for readers about network ties.

To date, interactivity is underutilized in systematic literature reviews that deploy SNA tools. Interactivity, which refers to the two-way flow of information between a computer and user, allows for the embedding of large excerpts from articles through tooltips. Visual clutter is avoided as users can call on the qualitative content by engaging with a tooltip instead of permanently displaying large quantities of text. In this paper we propose the use of Narrated Network Diagrams (NNDs) to provide analytical support for simultaneously engaging with both network diagrams and the qualitative content of articles (Cowhitt et al., 2023).

Longitudinal network analysis

Publication dates allow researchers to explore the dynamic nature of knowledge communities by examining literature through time. In the context of co-citation networks, longitudinal SNA can be used to understand how the influence of specific publications change. However, longitudinal network analysis requires the partitioning of relational data into discrete panels. Researchers can then track the evolution, dissolution, and strengthening or weakening of ties across panels to help explain outcomes of interest through the existence of specific structural configurations (Wasserman & Faust, 1994).

Our review of SNA reviews revealed limited use of panel data (Table 1). When panels were specified, justification for panel length was not provided. For example, one review conducted a quantitative analysis of keywords across five four-year periods to understand predominant and emerging research topics but provided no justification for the panel length (He et al., 2023). Another review conducted multiple bibliometric network analyses across two decade-long panels but did not justify the significance of decades as a unit of time when trying to understand the evolution of the field (Vaz & Araujo, 2022). The only explicit justifications for panel specification was a review that examined the outputs from a journal by decade at the journal’s 55th anniversary (Moreira Silva et al., 2019) and another that divided publications into 15 year increments to create equal variation between panels (Rossetto et al., 2018). Justifications for panel length should reflect significant contextual dimensions of the field of study under review. Our use of a policy pivot point to support the creation of panels is meant to encourage more authors to justify their use of longitudinal SNA methods when applying SNA methods to literature reviews.

Table 1 Prominence of keywords in the SNA literature review

| Category | Term | Frequency |
|------------------------|---------------------|-----------|
| Mixed Methods Research | Mixed methods | 32 |
| | MMSNA | 0 |
| | Joint display | 32 |
| | Interactive network | 2 |
| Time | Longitudinal | 175 |
| | Panel data | 21 |
| Analysis | Community detection | 26 |
| | Network density | 91 |
| | Triad | 117 |
| | Census | 0 |
| Software | Rstudio | 25 |
| | Bibliometrix | 177 |
| | VOSviewer | 202 |
| | Citespace | 2,661 |
| | Visnetwork | 0 |
| | Nvivo | 12 |

Collection was filtered for references, tables, and figures

Transparency in network partition

Many software options that allow researchers to apply generic SNA features to literature review protocols obscure critical decisions that can dramatically alter a network analysis. For example, researchers use SciMAT (Ghamari & Sharifi, 2021), Citespace (Sobral & Pestana, 2020), and VOSViewer (Vaz & Araujo, 2022) to help partition their keyword co-occurrence, co-citation, or co-authorship network into clusters or communities. Longitudinal analysis is then used to understand the influence of research clusters and the trajectory of articles within clusters (Phan Tan, 2021). However, there are many different methods for partitioning a network and few reviews included justification for their partitioning protocol or explained their community detection algorithm selection, which may indicate that these authors relied on the generic features of the bibliometric software rather than test alternative methods for best fit.

There are many different partitioning algorithms in network analysis and each protocol can yield different community or clustering formations from a given network. Reviews that justified a partitioning protocol tended to rely on the Louvain algorithm (Lin et al., 2022; Mazzu et al., 2022; Ophir et al., 2023; S. A. Raza & Govindaluri, 2021). However, the Walk-trap partitioning method (Ribeiro et al., 2022), Blondel algorithm (Aryadoust et al., 2019), and edge-betweenness (Luis Sanz-Cabanillas et al., 2017) were also used. We propose using dendrograms and community sampling as strategies to transparently validate the use of a particular partitioning protocol.

Overreliance on dyadic network statistics

Relatedly, for longitudinal network analysis to be productive, researchers need to identify relevant local structural configurations, track their development through time, and then connect their existence to an outcome or social phenomena of interest. There are many examples of descriptive network statistics in bibliometrics (i.e., centrality and density). However, descriptive statistical tools for identifying and tracking some of the most basic building blocks in networks, particularly triads and stars, are largely absent. Only one article references a triad clique (Yu et al., 2013) and there are no appearances of a dyad or triad census or k-star analysis in any of the articles in our review. Triads are foundational relational structures in the SNA literature and are at the center of many social phenomena (Heider, 1946). The methods protocol in this literature review proposes the use of 021U triads as a proxy for influence in a co-citation network.

Social network analysis prioritizes relationships when explaining phenomena like the emergence and evolution of ambiguous ideas. It has many practical uses in systematic literature reviews, both in the identification of relevant literature and in the analysis of publications, especially large collections of publications. But this process requires researchers to efficiently and effectively connect structure with story. It involves the use of network diagrams to show the stepwise process for relationships and structural configurations that change over time, as well as using qualitative data to make sense of the dynamism of networks (Stadtfeld & Block, 2017). This paper

offers several significant contributions to realizing innovative applications of SNA tools in bibliometrics.

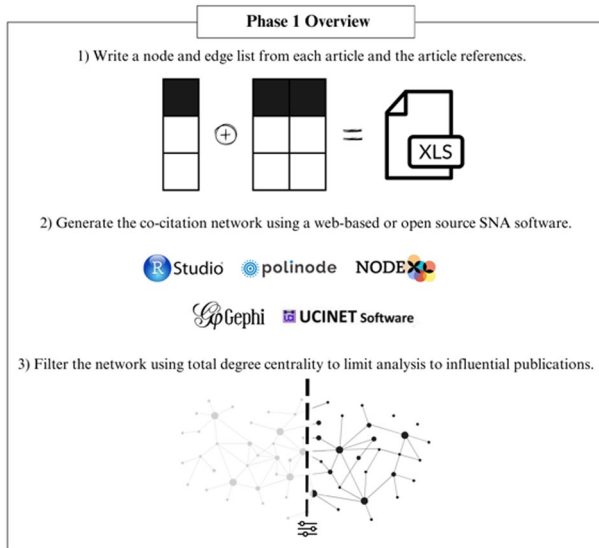
This review of SNA tools for analyzing literature highlights their widespread use by researchers working across dozens of fields. Their presence across myriad areas should not be surprising given the fundamentally relational nature of the development of knowledge communities. It also signals a broad interest in new applications for SNA methods. While the following sections demonstrate these new SNA methods using a literature collection on the growth of accountability in higher education over time, the methods are relevant to anyone assessing the emergence and evolution of big ideas across fields, knowledge communities, and society.

Materials and methods

The use of SNA tools to guide the analysis of a large collection of literature is presented in five phases. Phase 1 explains how to construct a co-citation network from a collection of articles and references. Citations serve as an indication of significance and visibility for literature (Aksnes & Rip, 2009). Examining co-citations across a collection of literature tracks the prominence of ideas and concepts. The resulting co-citation network included nodes and 14,281 edges. Phase 2 details a procedure for splitting the co-citation network into communities of articles and references to make the literature analysis manageable. Phase 3 outlines different attempts by the authors to account for time in the literature analysis protocol, ultimately settling on the use of prominent education policy to inform the partitioning of communities into panels. Phase 4 deploys a new data visualization technique for incorporating qualitative content into interactive joint displays, a critical tool for tying structure to story when developing narratives of knowledge communities through time. Finally, Phase 5 involves the sequencing of interactive joint displays into network schemas, resulting in a literature analysis protocol resembling longitudinal mixed methods social network analysis (MMSNA). The utility of this new approach is demonstrated in the findings, where the analysis protocol allowed the authors a more nuanced understanding of the evolution of ambiguous ideas through time.

Phase 1: Generating a co-citation network

A systematic search protocol was used to generate a collection of literature on accountability in higher education. Inclusion and exclusion criteria were then applied to the search results, narrowing the analysis to 450 journal articles published between 1974 through 2017. These original additions to the collection will be referred to as articles. An additional 12,270 pieces of literature were then added to the analysis for each reference cited by the articles. These secondary additions to the collection will be referred to as references. Readers interested in this systematic search and filtering protocol can consult the authors. The following methods protocol instead focuses on presenting a protocol for using SNA tools to make sense of the large collection of literature.



The protocol first requires the construction of a co-citation network. In a co-citation network, nodes represent each article and reference. Color is then used to differentiate between nodes that represent articles and nodes that represent references. For example, in Fig. 1, a blue node represents an article, and a purple node represents a reference. Finally, nodes are connected by discrete lines, often referred to as edges, representing a citation relationship between an article and each reference included in the article. Co-citation networks can

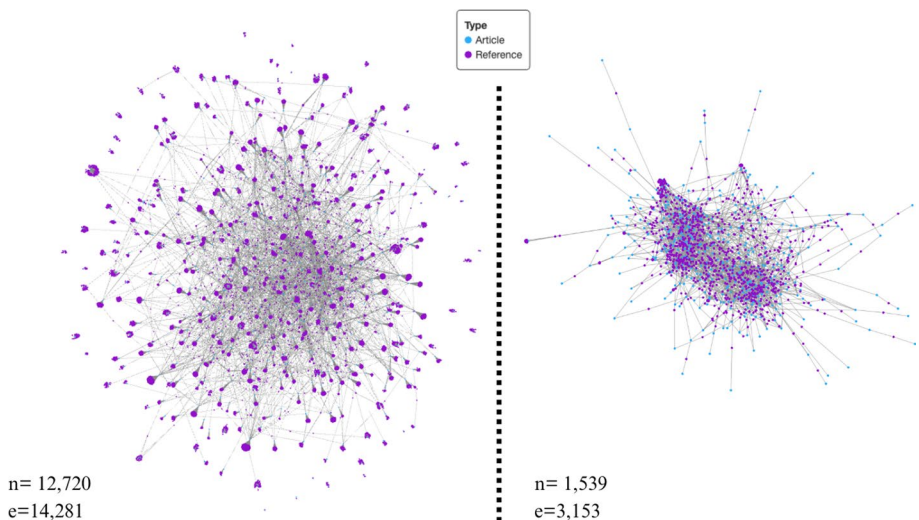


Fig. 1 Whole co-citation network before and after the total degree filter ≤ 1 is applied

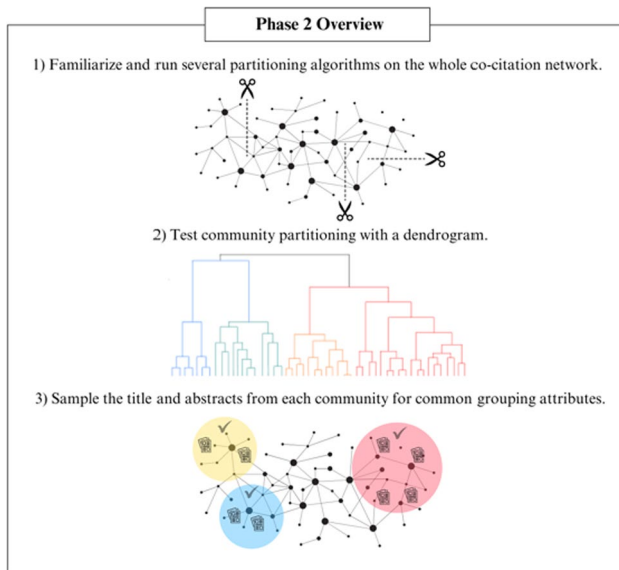
therefore be used to visualize and measure the intellectual lineage within fields of study by providing insight into the articles and references that researchers draw from as they contribute new work and develop a particular area of research.

The initial network diagram consists of 12,720 nodes and 14,281 edges (Fig. 1). Readers interested in developing their own relational dataset to generate a co-citation network can reference the underlying dataset for this co-citation network at the Harvard Dataverse using this link: <https://doi.org/https://doi.org/10.7910/DVN/VHB-WUD>. This dataset is a node and edge list, which in its simplest form is comprised of two tables. The node list is a single column table listing all article and reference titles. The edge list is a two-column table that lists each article in the left column and each reference in the right column. In this analysis, additional attribute columns were added to the node and edge list for publication year, journal source, authorship, organizational affiliation of authors, and relevant excerpts from the publication. This node and edge list was then uploaded to Polinode, a web application for network analysis. Other common software options for generating networks include the *igraph* package in RStudio, Gephi, UCINET, and NodeXL.

Basic descriptive network statistics were then used to de-clutter the full network. References that were cited by multiple articles were the focus of the analysis, as citation count is an indicator of potential influence on the development of ideas within a knowledge community. Total degree, the sum of all ingoing and outgoing edges, was calculated for each node, and nodes with a total degree less than or equal to 1 were filtered from the network. The resulting network diagram consisted of 1539 nodes and 3153 edges (Fig. 1). Those interested in reading about total degree centrality or other descriptive network statistics relevant to literature analysis are encouraged to consult *Social Network Analysis: Methods and Applications* (Wasserman & Faust, 1994), as a commonly referenced resource for scholars and practitioners working in SNA across disciplines.

Phase 2: Splitting the whole network into communities

In a document co-citation network, community affiliation determines collections of articles that use common references. These groupings represent discrete academic conversations within a larger field of study and can be used to help identify emerging ideas, experiments, or methods (Leydesdorff, 1998; Small, 1978). Co-citation networks vary in density, or the number of connections between authors and references. Researchers wanting to understand the evolution of ideas need to seek out highly connected areas in whole networks and examine central authors or publications. This will provide insights into what kinds of knowledge influences and defines a field of study as it matures.



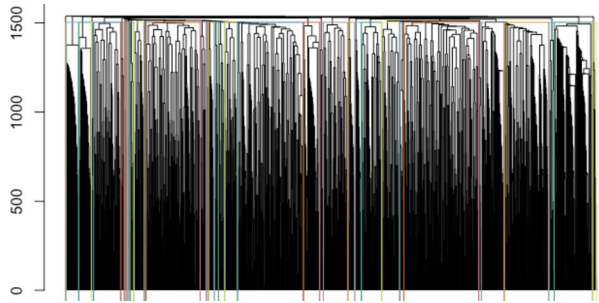
Community detection algorithms distinguish between subgroups within networks using either divisive, agglomerative, or optimization methods (Blondel et al., 2008; Javed et al., 2018). The quality of community detection is often measured by modularity, which is a scalar value comparing the density of links within and between communities (Newman & Girvan, 2004). To determine an appropriate community detection technique for the co-citation network, we tested five different algorithms (Table 2). The resulting modularity values and number of communities created by each algorithm were compared. Dendrograms were generated for each community detection algorithm to check that the height of the partition site across the hierarchy of stacked branches, or clades, generated substantially connected components of articles and references instead of smaller communities that would represent niche conversations within the field (Fig. 2).

The co-citation network was ultimately partitioned into communities using an advanced version of the Louvain algorithm, an optimization method that, “finds high modularity partitions of large networks in a short time and that unfolds a complete hierarchical community structure for the network, thereby giving access to different

Table 2 Comparison of partitioning strategies on the co-citation network

| Community Detection Algorithm | Modularity | Communities Generated |
|-------------------------------|------------|-----------------------|
| Leading eigen | .571 | 66 |
| Walk trap | .607 | 112 |
| Fast and greedy | .645 | 24 |
| Louvain | .672 | 23 |
| Edge-betweenness | .679 | 36 |

Fig. 2 Dendrogram of the Louvain algorithm partition protocol



resolutions of community detection” (Blondel et al., 2008, p. 3). Twenty-four communities were detected within the connected component of the whole co-citation network (Fig. 3).

As a final check on the partitioning technique, the content of articles and references from 25% of the resulting communities were examined for thematic coherence by two members of the research team. Four of these communities were selected for this sense-making exercise as they included articles across each of the five policy panels. Two additional communities were selected because they had distinctive structural configurations. Articles and

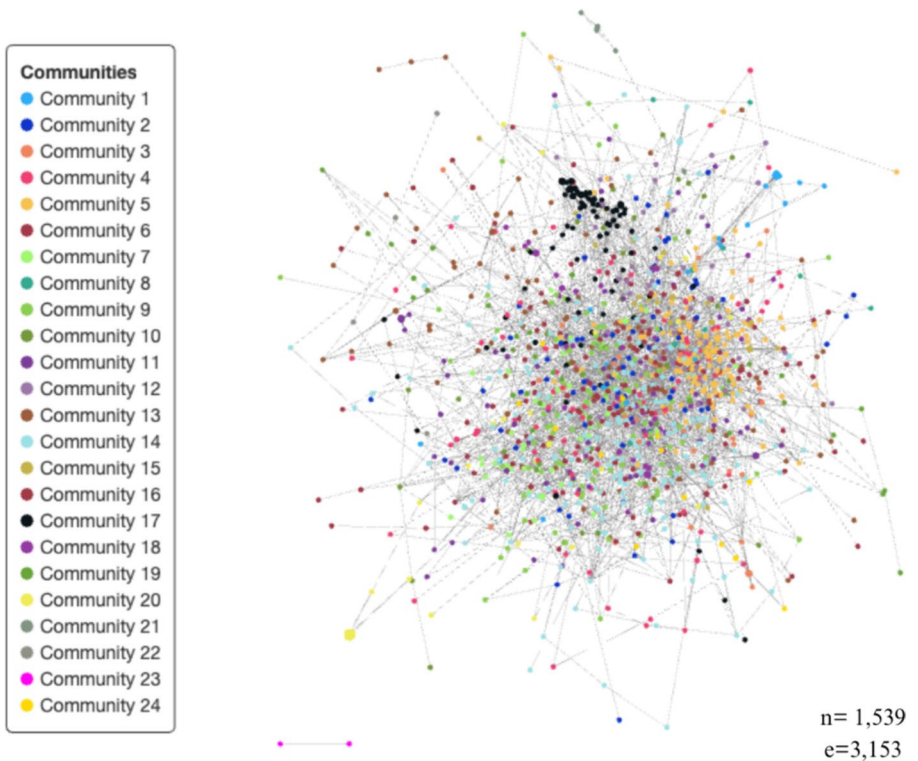
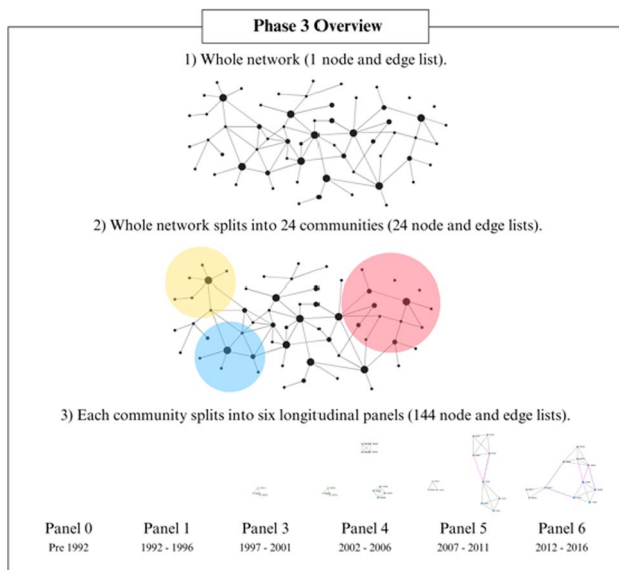


Fig. 3 Co-citation network partitioned into communities

references were examined for shared research focus, comparable methodological framing, the selection and use of similar methods, and common author attributes. This final check revealed general unifying characteristics within each of the six communities and provided confidence that analysis of the remaining communities would likely yield coherent lineages of research within the higher education accountability literature.

Phase 3: Accounting for time

Following the partition of the co-citation network into communities, each community was subdivided into longitudinal panels to capture the dynamism of knowledge creation. Partitioning co-citation or co-authorship networks into communities organizes large collections of literature into approachable quantities of relevant texts, which can help researchers in identifying different narratives or active areas of research within a field of study (Cowhitt et al., 2019). However, knowledge communities, like all social networks, are dynamic. They are constantly changing as membership churns. These knowledge communities are permeable to new ideas and reactive to new policies or events.



Several factors were considered when determining panels, including the size of the relational dataset and logical political and policy events relevant to the research context. First, decade panels were created. This generated unwieldy panels of several hundred nodes each and appeared to represent arbitrary cleaving of the network. Discarding that strategy, panels were then created to coincide with presidential administrations to reflect changing legislative priorities of the two prominent political parties in the United States. This strategy was also abandoned when single and two-term administrations created large variations

in the number of nodes across panels. These panels were also viewed as untenable representations of legislative context as Congress regularly switches party control within presidential administrations.

Finally, a panel protocol was developed around relevant education policy, notably the 2001 No Child Left Behind Act (NCLB), which codified accountability into education policy in the United States. Using 2001 as a policy pivot point, articles and references were sorted into panels of five year increments beginning in 2002. Methodologically, this involved splitting the node and edge lists of each of the 24 communities into panels based on the publication year of the articles and references. This procedure yielded five 5 year panels for each community: 1992–1996, 1997–2001, 2002–2006, 2007–2011, and 2012–2016. Prior to 1992, references and articles were sparse and an additional panel 0 was created for these pieces of literature (see Table 3).

Phase 4: Generating interactive joint displays

In addition to partitioning the co-citation network into communities and longitudinal panels, interactive displays were used to help researchers connect relational structures to relevant qualitative content in articles and references. This phase was instrumental in facilitating the qualitative examination of citation ties that can reveal insights into the evolution of a field—connecting structure to story. The text surrounding the use of a reference was extracted from each article and integrated into edge attributes to create narrated network diagrams (NNDs). Mechanically, this entailed adding a new column to the underlying edge list and populating it with the relevant qualitative content from each article. The qualitative excerpts are effectively hidden within edge labels using interactive tooltips to avoid visual clutter during analysis. NNDs allow readers to call on qualitative data by hovering their cursor over an edge with embedded text (Cowhitt et al., 2023) and are particularly effective resources for making sense of relational configurations in co-citation communities because they provide on-demand context to readers about each reference.

Table 3 Dates informing the formation of policy panels for longitudinal analysis

| Panel | Start Year | End Year |
|-------|------------|----------|
| 0 | Pre-1992 | 1991 |
| 1 | 1992 | 1996 |
| 2 | 1997 | 2001 |
| 3 | 2002 | 2006 |
| 4 | 2007 | 2011 |
| 5 | 2012 | 2016 |

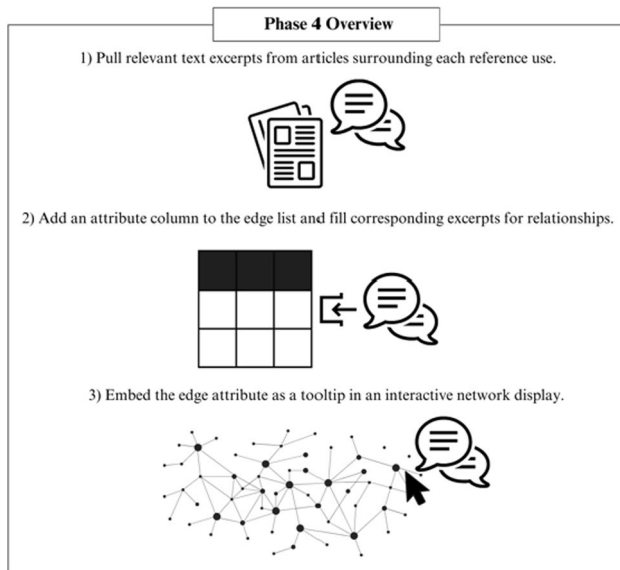


Figure 4 shows how NNDs connect structure to story using the tooltip feature in Community 5. In the second panel, the network is comprised of nine nodes— six articles labelled A1-A6 and three references labelled R1-R3 (A6 obscured by the text display). When the cursor is positioned over the edge that connects Article 4 (A4) with Reference 3 (R3), the qualitative text where R3 is cited in A4 is displayed in the tooltip callout feature. The interactivity of this display allows readers to access qualitative data on how each reference is used by authors.

Phase 5: Network schemas and community evolution

Social networks are inherently dynamic wherein new node relationships continuously form and existing edges strengthen, weaken, or collapse. This dynamism remains an essential feature of the academic writing produced by knowledge communities as new manuscripts are written, seminal sources become more central through continued citations, and others

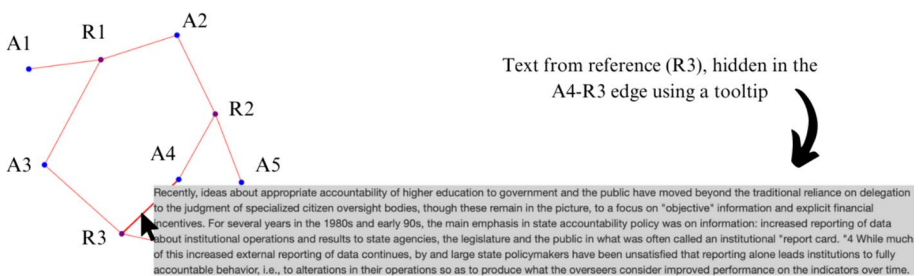
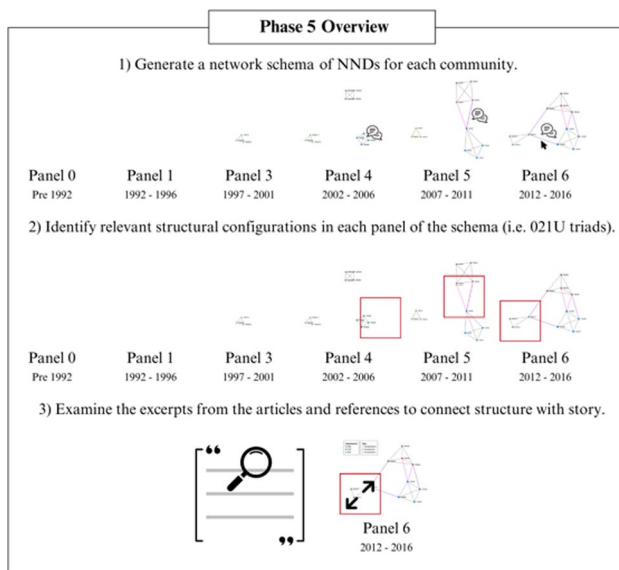


Fig. 4 Narrated Network Diagram from Community 5

inevitably fade as mentions from colleagues dwindle. Therefore, to understand the emergence and evolution of ideas within the higher education accountability literature, network schemas were created for each community.

A network schema is a sequence display, often used to highlight the evolution of structures within whole networks through time (Stadtfeld & Block, 2017). For this co-citation analysis, a network diagram was created for the articles and references within a community during each of the six policy panels. This revealed several different community growth patterns, with some communities gradually accumulating articles and coalescing around a coherent collection of influential authors. Other communities remained sparse for decades before rapidly amplifying within a particular panel. Some communities did not gain critical mass to establish themselves as influential components within the field.



Creating a visual sequence of co-citation communities through policy panels revealed the development, and in some cases, the collapse, of interesting network structures. In Findings and Results section, we present two examples of network schemas as a proof of concept for how these innovative SNA methods yield new insights when analyzing the emergence and evolution of big ideas across literatures, fields, knowledge communities, and society. We specifically examined the existence of triads and star configurations, as these structural phenomena in co-citation networks are translatable to significant narratives in knowledge creation, and in this specific case, structurally located significant developments in unfolding "story" of the accountability in higher education knowledge community.

This systematic review protocol deployed descriptive network statistics in new ways to understand dynamic network structures and their significance for literature analysis in understanding the emergence of big ideas. Namely, a triad census was deployed, with specific attention to identifying 021U triads which can be used as an indicator for influence. The resulting analysis and findings would not have been possible without exploring the development of 021U triads and the thematic content embedded within their edges in the narrated network diagram. The following section details two examples for how this

innovative approach to literature analysis can yield different understandings about the evolution of ambiguous ideas.

Findings and results

Two community analyses are provided here as proof of concept for the new deployment of SNA tools in literature review protocols. These two communities were chosen because their structural evolution resembled two different trajectories about accountability. In the first analysis, Community 5 represents the incremental evolution of an idea. The second analysis of Community 17 highlights an explosive expansion of a knowledge domain. Both communities were generated by applying the partitioning protocol as described in Phase 2 of the Materials and Methods section to the co-citation network. Each community was then split into panels using the No Child Left Behind Act as a policy pivot point for five-year increments. Finally, Phase 4 of the methods protocol was applied to generate interactive displays of the longitudinal network analysis and relevant descriptive network statistics were used to develop more nuanced understandings of knowledge outputs from each community regarding accountability in higher education.

Incremental evolution–community 5

Community 5 illustrates a traditional development of an education policy community. It originates with a single published article and over time a network of articles, linked by their references grows to form a large, interconnected community (Fig. 5). This developmental growth occurs gradually and is largely focused on a single research area—performance-based funding, an education policy regime in higher education that emerged in the early 1990’s (Miller & Morphew, 2017; Ortagus et al., 2020). The topic of performance-based funding is an area of higher education accountability that emphasizes the allocation of funding to colleges and universities by their respective state legislatures based on a set of performance metrics. Across the entire period of our study, this community consists of 35 articles and 121 attendant references (Table 4).

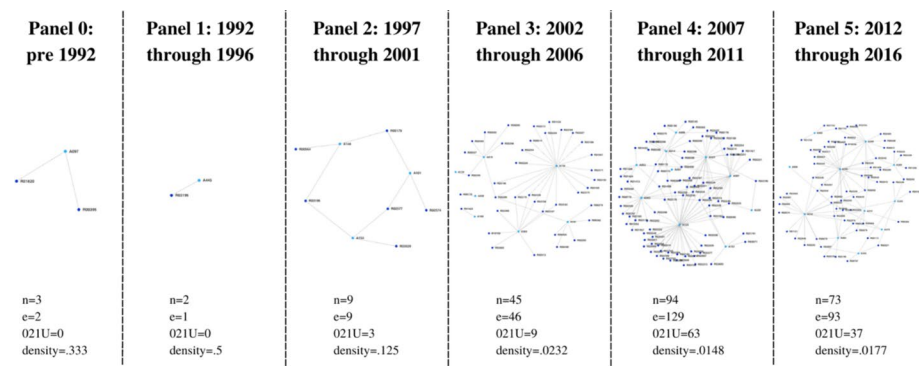


Fig. 5 A network schema of Community 5 from the co-citation network, including frequency of 021U triads and density across panels

Table 4 Descriptive data for communities 5 & 17

| | Community 5 | Community 17 |
|-----------------------|-------------|--------------|
| Article Total | 35 | 22 |
| References Total | 121 | 92 |
| Triads (021U)/Density | — | — |
| Panel 0 | 0 | 0 |
| Panel 1 | 0 | 0 |
| Panel 2 | 3 | 0 |
| Panel 3 | 9 | 0 |
| Panel 4 | 63 | 70 |
| Panel 5 | 37 | 131 |

Prior to 1996, we find only two articles in the performance-based funding community. It is not until Panel 2, in the 1997–2001 period, that we see an interconnected knowledge community linked by shared references. The three 021U triads form a ring, creating a community that substantively centers on state-level accountability in higher education. The articles focus on state accountability in higher education generally (e.g., A101, A153) or on the data system infrastructure demanded by a new state accountability regime (A148).

The references that connect the articles in Panel 2 include two reports and a popular press book, each touching upon the contemporary call for greater accountability of public institutions through performance-based funding policy. For example, Osborne and Gabler’s book (R03196) is used by Zumeta (A153) to identify the policy movement toward outcomes: “To summarize briefly, the rhetoric of the current accountability and quality improvement movement in business, government and higher education calls for a refocusing of attention, particularly in resource allocation, on results or outcomes of activities or programs, ideally in relation to explicit goals, as contrasted with the traditional focus on inputs” (Osborne & Gaebler, 1992). The two reports (R02577, R00178) serve to provide an empirical accounting of current state-level performance measures for scholars to illustrate the movement where Wells (A148) cites Christal (R00178) in this way: “One measure of the pervasiveness of this trend is the number of states that have adopted performance-based systems of funding over the past decade. According to a recent study conducted by Christal (1998) for SHEEO, thirty-seven states use performance indicators of colleges and universities to some extent.” This is notable because the reports, issued by Education Commission of the States (R02577) and State Higher Education Executive Officers Association (R00178), and the book (Osborne, a public policy consultant) can be characterized as emerging from the public policy community, not from peer academics (Osborne & Gaebler, 1992). The initiation of this community therefore coheres around the combination of data and policy work outside of the academy.

A small amount of growth occurs in Panel 3 (2002–2006). In this period there are six articles—four of which are interconnected—all focused on state-level accountability, performance-based funding, and implications for universities and states. A ring-structure is no longer present, and though there are three times the number of articles in the community, there are only four 021U triads. The community hinges, however, on a key reference at the center of three of the 021U triads, a report published by a university research institute, “Funding Public Colleges and Universities for Performance” (Burke and Associates, R00386). This report is a comprehensive accounting of then-current issues on performance-based funding for public institutions, including analyses of policies across several state cases. Two articles that share R00386, cite it for its case studies of state policy

examples in New York (A019) and Tennessee (A141). In A141, the context is clear: “This chapter examines the Tennessee accountability framework, offering an overview of the initiative, its measures, and its impact on institutions. The perceived value and effectiveness of Tennessee’s public accountability framework is suggested by its longevity and stability (Burke and Associates, 2002).” The central node of the one remaining triad is a report from the National Center for Public Policy and Higher Education (A321), a publication that grades each state in the U.S. on preparation for college, participation, completion, affordability, benefits, and learning. Coherence in the community remains dependent on reports rather than academic articles or books. Substantively, the community remains tightly focused on the emergence of performance-based funding policy and how states and institutions are adapting to the new higher education accountability regime.

In Panel 4 (2007–2011), there are ten publications, all of which are journal articles. The community is more robust in this period, with dozens of 021U triads present. Reports and trade books are present in the community as references and remain important. The National Center for Public Policy and Higher Education report (R00400) links together five articles—(A014, A052, A321, A367, A087)—each of which cite the report primarily to present analyses of the policy adoption of performance-based funding nationally. The majority of connecting references, however, are now journal articles. McClendon, Hearn, and Deaton’s journal article (A118) is a shared reference for five articles. It is a review and growth analysis of performance-based funding policies in state educational policymaking. Co-citing articles use the reference to primarily help establish the state of the research on performance-based funding. An example in A353 is clear: “Another important aspect of the state context considered is governance. Previous research in this area has focused mostly on examining the relationship between governance structures and other state-level policies, such as accountability (McClendon et al., 2006) and appropriations to research universities (Weerts & Ronca, 2006).”

In this period, the community begins to focus on specific outcomes legislatively specified in performance-based funding policies, particularly degree completion. McClendon, Hearn, and Deaton anchor two (A087, A353) of the three articles that exhibit this focus (R02206), (A087), and (A353). Other references that form linking triads around degree completion and performance-based funding include Archibald and Feldman (R03314), Ryan (R02245), Scott and Bailey (R02246), and Volkwein and Tandberg (R02253). Lastly, a new topic emerges in this period, made visible by references within 021U triads that discuss state politics and governance structures, including, for example, Archibald and Feldman (R03314), Lowry (R03182), Leslie and Berdahl (R02233), and Osborne and Gaebler (R03196).

The final panel analysis for Community 5 reveals that the number of articles (11) has stabilized from the previous period, possibly indicating maturation of the community. There are also fewer shared references among articles with forty percent fewer 021U triads present. Topically, the core of the community are five articles that focus on policy and the politics of performance-based funding, particularly at the state level (A068, A113, A388, A392, A404). The community maintains a cluster focused on outcomes, although there is a slight shift in emphasis toward retention (A350) and (A384) as the outcomes-based metric rather than graduation (A393). An emergent part of the community is internationally focused research, although the connection to performance-based funding and accountability is weak. The two articles are studies of institutional-level accountability metrics such as student evaluation (A008) and student cost-sharing policy in public institutions as a form of accountability (A270). They are relatively peripheral to the community but are critically linked to the high-degree article by Hearn et al. (A113) indirectly (A008) and directly

(A270) by Johnstone’s article (R02974) on worldwide trends in financing higher education, which includes discussion of cost sharing. An excerpt from A270 reads: “Among the many other policy alternatives that have been put in place is the introduction of cost recovery or cost sharing, deferred payment of loans, and pressure on universities to diversify resources to generate alternative sources of income just again to mention a few options (Johnstone, 2009).”

Explosive expansion–community 17

Community 17 is illustrative of the rapid development of an education-policy community researching low-stakes testing and assessment (Fig. 6). Although the duration of Community 17 is shorter than Community 5, the size of its overall network is somewhat similar, comprised of 22 articles and 92 attendant references (Table 4). The topic of low-stakes testing and assessment is an area of higher education accountability that emphasizes the application of psychometric methods to create tests and educational assessments within individual institutions to measure change in students. The application of psychometric principles notably diverges from the application of national education assessments that have “high stakes” for students during undergraduate and graduate admission (i.e., SAT, ACT, GRE, LSAT, etc.). We observe that the explosive nature of this community reflects the post-2001 testing emphasis that followed the NCLB federal education policies.

The network momentum of the low stakes testing and assessment community finds its origin in a seminal article in Panel 3 that examines the motivation of students in low stakes testing (A397). It addresses whether student scores are a valid measure of performance in low-stakes testing conditions where, “there are no personal consequences based on test performance for the individual test taker” (Wise et al., 2006, p. 65). The foundational nature of this piece by Wise et al. becomes clearer in Panel 4 when it is used by later authors to explain tensions in the use of low-stakes assessment in higher education accountability.

The transition from Panel 3 (2002–2006) to Panel 4 (2007–2011) highlights that the increased diffusion of higher education accountability created an ambiguous state of competing concepts within the low stakes testing and assessment knowledge community. These competing concepts are best captured in (A156). Here, the author questions whether conceptual advancements in the measurement of student learning and improvement are at risk

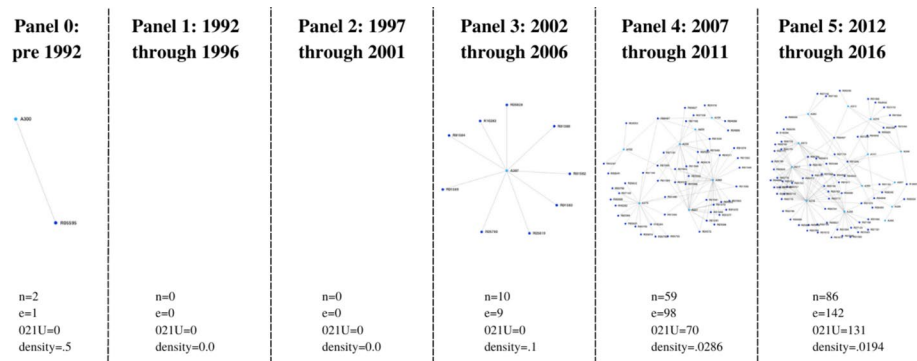


Fig. 6 A network schema of Community 17 from the co-citation network, including frequency of 021U triads and density across panels

of being overtaken by accountability-centered assessment promoted by policymakers. She asks, "...will an accountability tidal wave roll across the fields, crushing the fragile green sprouts of assessment for improvement that have begun to appear?" (Banta, 2007, p. 9). In the community's attempt to accommodate both perspectives—assessment for accountability and assessment for improvement—we observe the emergence of a new assessment discourse focused on value-added assessment. One author explained:

To measure instructional effectiveness in higher education, a term value-added was introduced in [the] voluntary system of accountability. Value-added is defined as the performance difference between first-year and fourth-year students on a standardized test (e.g., ETS Proficiency Profile, CAAP, CLA) after controlling for student admission scores (e.g., SAT, ACT).

Furthermore, in Panel 4, these two perspectives on assessment are noticeable in the configuration of the panel into two halves, one focusing on value-added and quality assurance (A060, A255, A256, A257, A122), while the other half focuses on the assessment of student learning and improvement (A156, A379).

The two halves of the network in Panel 4 are held together by 6 references that form O21U triads which bridge these two discussions (Fig. 7). When further examined, the contextual use of the reference in the triad embodies the tension between the two perspectives. For example, when an article (A379) from the half of the network focused on assessment of student learning and improvement cites Wise et al. (R01589), they state, "the strong emphasis on accountability, which requires universities to verify graduates meet performance goals, means that program assessment has high stakes for a department and its university, but not for its students, who are usually exempt from accountability for their performance" (Huffman et al., 2011, p. 90). Similarly, when an article (A060) from the half of the network focused on value-added

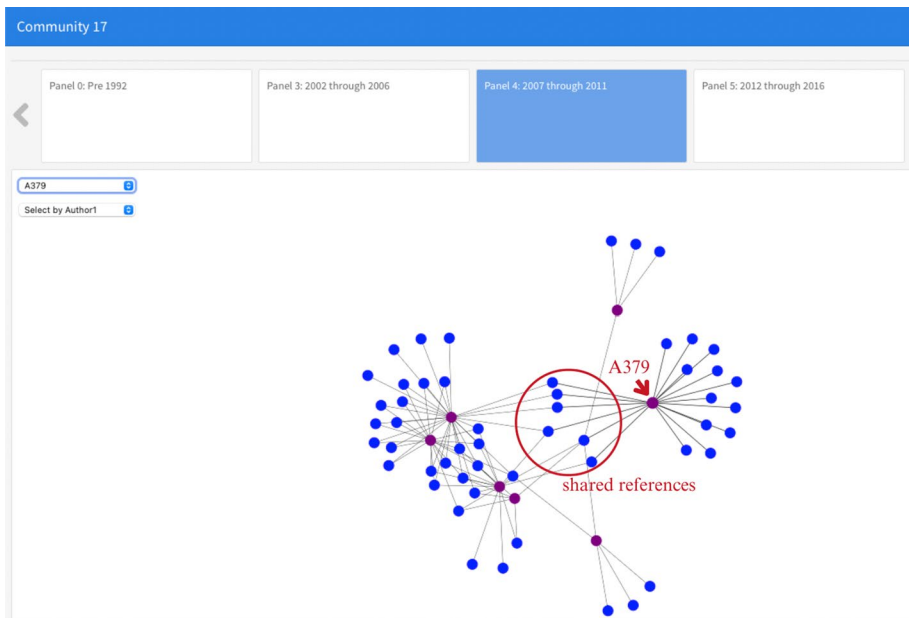


Fig. 7 Shared references connecting article A379 with five other articles in the Panel 4 network for Community 17

and quality assurance cites Wise et al. (R01589), they emphasize the, “lack of student motivation in test taking also leads to inaccurate estimates of student achievement and thus affects decisions about institutional efficiency” (Liu, 2011, p. 6). The actual text surrounding each of these bridging references highlights the important role 021U triads play not only within the network of actors, but also within the network of ideas.

Panel 4 illustrates the emergence of two distinct network halves with the adoption of low-stakes assessment instruments as an important form of measurement in higher education accountability. But it also highlights an explosive period of expansion in the low stakes and assessment community, which moves from 1 article and 9 references in Panel 3 to 7 articles and 52 references in Panel 4. In comparison, by this same 2007–2011 period, Community 17 had generated a similar amount of scholarship as Community 5, *but in half the time*.

The same “explosive expansion” trend is also seen in Panel 5 (2012–2016) when the low stakes testing and assessment network nearly doubles in size to 13 articles and 73 references. During this period, the community retains its shape as two network clusters held together by a small amalgamation of mostly references. One cluster continues its focus on motivation and value added (articles by Hawthorne, Finney, and Liu) while the other cluster continues to focus on student learning and improvement. The small group of “bridging” references situated between the two clusters takes on a new form in Panel 5 with most sourced from books (Assessing General Education Programs; Academically Adrift) or reports (Department of Education; OECD; Educational Testing Service) rather than peer reviewed articles. Situated between two sides of the community, these bridging references empirically focus on the sources of the discussion with an emphasis on policies (R00407) and testing instruments (R01565), or they empirically focus on measuring variation in learning gain (i.e., value-added) driven by different schools (R01236) or different levels of student motivation (R04973).

A further examination of the “bridging” references highlights an era (2012–2016) where researchers and practitioners continued to make sense of the shortcomings in using low stakes assessments for students as high stakes performance measures for colleges and universities: “Institutional accountability mandates prompt assessment of student learning. Although designed to accurately assess learning, many ‘accountability tests’ are low stakes for students...” (Finney et al., 2016). The mandates to which many refer are sourced to “A Test of Leadership” published the U.S. Department of Education, or more commonly referred to as the Spellings Report: “Student achievement, which is inextricably connected to institutional success, must be measured by institutions on a ‘value-added’ basis that takes into account students’ academic baseline when assessing their results” (Spellings, 2006, p. 4). In Panel 5 the community is awash in a discussion of myriad instruments to examine the value-added of individual institutions (e.g., CLA, MAPP, NSSE, etc.). And while the tension is not resolved in Panel 5, data emerges that seems to foreshadow future eras that must confront the negative biases in accountability measures seemingly rooted in test motivation and fatigue: “This study offers a direct response to these concerns and delivers evidence that decreasing examinee effort is negatively biasing our best estimates of student learning gains” (Finney et al., 2016).

Discussion

Sociocultural theory posits that knowledge is constructed, manifesting from our engagement with others, and is therefore inherently a social enterprise (Rieber, 1997; Vygotsky, 1980). SNA tools can strengthen systematic literature reviews by accounting for this relational dimension of knowledge creation and allowing for a more comprehensive

understanding of how new ideas evolve in different contexts. This article advances the application of SNA tools to systematic literature reviews in three ways. First, Narrated Network Diagrams, improves the ability of researchers to engage with qualitative content. Second, longitudinal network panels create new analytic opportunities for studying the development of knowledge communities when time is used to inform the partitioning of networks. Finally, the combination of these techniques in systematic literature reviews allows for new applications of descriptive network statistics and the detection of emergent findings. The application of these methods provides a protocol for researchers to engage with significant volumes of writing when making sense of a field of study. This longitudinal MMSNA approach can help researchers develop more nuanced understandings of the emergence and evolution of significant ideas.

Narrated network diagrams

Interactive displays that integrate different types of data can provide new analytic and interpretive opportunities (Guetterman et al., 2015). In this paper, we introduce the use of interactive displays called Narrated Network Diagrams (NNDs), which use tooltips to embed qualitative excerpts from articles and references into edge labels (Cowhitt et al., 2023). This connection of structure to story provides a convenient advantage of understanding themes and assessing meaning when developing interpretations of network structures.

For example, a synthesis of excerpts from articles and their references in a NND display allowed us to identify that Community 5 coalesced around the concept of performance-based funding, an education policy regime in higher education that has endured since the 1990s. NNDs proved advantageous for building nuanced interpretations of how and why ideas were linked within co-citation communities, particularly in their infancy. Similarly, articles and references in Community 17 coalesced around the idea of low-stakes testing. The influence and spread of this prominent idea is technically observable by manually scanning the hundreds of titles and abstracts across all the generated co-citation communities, but NNDs allowed us insight into the dynamics of this knowledge community by revealing dissonance in the shared references authors used. The tension between using assessment for institutional accountability versus the motivation of students was at the root of the emergence of value-added assessment, the key idea of Community 17.

These types of insights that focus on understanding themes and assessing meaning are much more difficult to realize through traditional categorizing and scanning of literature because they require time-intensive close analysis of full-text publications. Nuanced understandings of the academic discussions occurring in co-citation communities requires researchers to engage with the qualitative content of articles, even exploring how authors make use of shared references. However, the volume of literature identified during systematic literature reviews renders this type of close analysis unrealistic. Even the use of co-citation networks to categorize a large collection still makes engagement with qualitative content challenging as researchers must move between network diagrams and full-text publications. An MMSNA protocol that makes use of interactive displays provides integrated data for researchers to systematically engage with qualitative data during literature analysis. The result is a deeper understanding of the academic conversations underlying the evolution of seemingly ambiguous ideas.

Policy-informed longitudinal network panels

We demonstrate that partitioning co-citation communities into policy-informed longitudinal panels can both increase analytical acuity and contribute to the understanding of knowledge community evolution. In this case, NCLB was an appropriate policy pivot point as the legislation begins with the rationale, “to close the achievement gap with accountability...” and mentions the word accountability another 79 times (No Child Left Behind, 2001). More significantly, published literature in higher education accountability increased significantly beginning in 2002, immediately after the passage of NCLB (Fig. 8). A network schema of five-year panels on either side of 2001 offered insight with respect to the size of individual networks given the seminal time point of the passage of NCLB.

Informed partition of co-citation communities using a combination of education policy and publication dates created new analysis opportunities. One area of analysis that became possible was insight into the evolution of each co-citation community. Community 5 is an example of incremental development, originating with a single text and consistently growing over several decades into a discretely identifiable knowledge community concerned with performance-based funding. Alternatively, Community 17 is an example of explosive expansion, with rapid development of the community focused on low-stakes testing and assessment occurring after 2006.

The rate of growth can provide unique insights into a field of study. For example, the explosive growth in Panel 4 of Community 17 was led by a single researcher who authored four of the seven articles from the Educational Testing Service. The succeeding 5th panel includes articles by twelve different authors, with the original researcher being prominently cited. This growth pattern within a knowledge community may indicate individuals’ capacity to shape novel research directions in their field. It also suggests organizations like the Educational Testing Service could potentially influence narratives surrounding these topics by fostering concentrated outputs before the community diversifies with authors from other institutions. Understanding the evolutionary trajectory of ideas and influence in co-citation or co-authorship communities could be an area for further development by researchers leading systematic literature reviews.

Accounting for time also allows researchers to further advance the interactive displays from the MMSNA protocol. After the longitudinal element was added to the MMSNA protocol, NNDs were generated for each panel and stacked into an interactive display. This

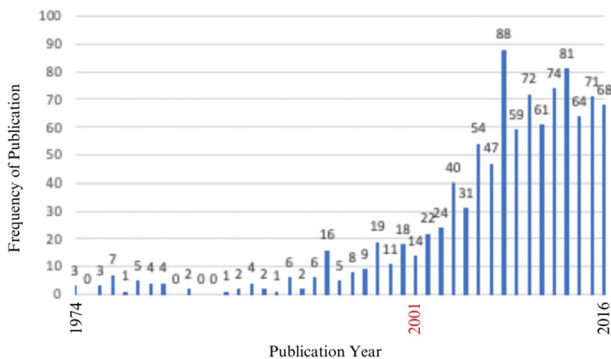


Fig. 8 Frequency of publication by year in the accountability in higher education literature

allowed for the formation of an interactive network schema, which consists of sequenced network diagrams used to depict evolutionary micro-steps of relational networks. These interactive displays that integrate network and qualitative data improved the analysis of Community 17 and allowed the review team to uncover a tension between two different perspectives on assessment- a degree of nuance that would have been far more difficult to discern without NNDs.

Emergent findings

New applications of descriptive network statistics, NNDs, and policy panels, led to emergent findings. The triad census helped identify influencers and innovators within discrete research communities. Their presence was uncovered by examining publication data over time and their impact was readily explained because researchers had access to qualitative content integrated with network diagrams in the NNDs.

The emergent findings from this analysis show that influencers exhibited one of two strategies to expand their role in scholarly discourse. We describe these two approaches as *persistence* and *saturation*. One approach employed by influencers can be seen in both Communities 5 and 17 which provide examples of seminal scholars who emerge and persist across multiple panels over time. *Persistence* involves an article transitioning to a reference in a later panel. Several articles make this transition and persist across the whole co-citation network, but fewer articles became references that functioned as vital connectors within their communities. This staying power by an article and their references represents potential influence, especially when references become shared by multiple authors.

In Community 17, article A397 emerges as the only article in Panel 3. The resulting structural configuration is known as a 9-star, resembling a wheel with the article in the center and nine references serving as spokes. By Panel 4, articles A397 and references R01588 and R015889 by the same author, are the highest cited publications in the community, with each being cited by six additional articles. Similarly, in Community 5, article A118 makes up the center of a 21-star. In panel 3, only six of these references (R00386, RR03156, R03159, R03176, R03182, R03187) were cited by additional papers. By Panel 4, twenty of the original twenty-one references are cited in other articles within the community. Structurally, when a reference is cited in an additional article, a 021U triad is formed. Because 021U triads represent a common citation, the frequency of 021U triads can be used to identify influential references and help uncover seminal authors through time. Among the 021U triads, references that are cited by three or more articles in a community occupy structural holes, which in a co-citation network means that these pieces of literature have a greater chance of containing good or innovative ideas (Burt, 2004). The influencers who authored these references can also be considered boundary spanners within their respective academic communities, serving to influence multiple works through the persistence of their own writing (Williams, 2002).

In contrast to persistence, we describe a second approach employed by influencers as *saturation*. In each of the communities an emergent author assumes a central role by focusing on producing a high frequency of publications that includes a combination of both articles and references. In Community 5, Burke achieves an essential position in the network with six publications in Panel 3 (A019, R00090, R00386, R00582, R01128, R02213). Similarly, in Community 17, Liu accomplishes influence within the network through seven publications in Panel 4 (A060, A255, A256, A257, R1571, R1572, R1574). In both instances, one of the references connects three articles, elevating both authors as potential

boundary spanners within the community. It is important to note that publications by Burke and Liu do not appear in their respective communities prior to the point of saturation. However, each author maintains a steady presence in the network following the period of saturation. This suggests that scholars can maintain an influential part of their respective network across time following their emergence through either persistence or saturation. Identifying these influencers and their specific role is only possible when deploying existing descriptive network statistics in new ways through longitudinal and MMSNA protocols.

Conclusion

Systematic literature reviews are attempts to understand a collection of related conversations between different groups of researchers working to develop solutions to common problems. These conversations occur through time, often stretching back decades. Even when research leads to false-starts or unproductive tangents, each published output becomes part of an intellectual mosaic documenting discrete contributions to the fickle momentum of learning about the world around us. Those incorporating SNA tools in their approaches to reviewing literature already see value in accounting for the social dimension of knowledge creation. This article advances these efforts by asserting that researchers must account for time and the qualitative exchange that occurs between authors to more accurately make sense of the social dimension of knowledge creation.

This article advances the use of SNA methods in systematic literature reviews by demonstrating how more advanced applications of SNA tools can be deployed to understand the evolution of ambiguous ideas that transverse knowledge communities. For example, mixed methods approaches provide opportunities for using interactive displays, which offer more integrated data for researchers when interpreting the outputs of research communities in co-citation networks. Elevating time to help inform the partitioning of networks also allows for opportunities to understand the dynamism of knowledge formation and learning through longitudinal analysis. Finally, when descriptive network statistics are applied to the longitudinal MMSNA protocol, emergent findings surface, such as clear mechanisms for writers gaining influence (i.e. persistence and saturation).

SNA applications to systematic literature reviews are still uncommon, and those that do only represent cursory attempts at applying the full power of SNA to this critical research practice. However, deploying SNA methods to aid in literature searching and literature analysis is not universally recommended. Research teams should conduct a cost-benefit analysis before applying new methods to literature review protocols to determine if the data are well suited to employing such techniques. That said, this article demonstrates how the sophisticated application of SNA methods in literature analyses equips researchers with tools to more accurately assess themes and meaning, while offering nuanced insights into knowledge formation and evolution across academic disciplines, fields, and society more broadly.

Acknowledgements For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) license to any Author Accepted Manuscript version arising from this submission. The authors would also like to acknowledge Dr. Sondra Barringer for her conceptual and methodological feedback, and Irene Toussaint for her assistance creating the accountability database.

Author contributions All authors contributed equally to this project and share equal authorship. Thomas Cowhitt, Corresponding author: Conceptualization, Methodology, Formal quantitative analysis, Writing—original draft, review & editing. Joshua Travis Brown: Conceptualization, Methodology, Formal qualitative

analysis, Writing—original draft, review & editing. anthony lising antonio: Conceptualization, Methodology, Formal qualitative analysis, Writing—original draft, review & editing.

Declarations

Conflict of interest The authors did not receive support from any organization for the submitted work. All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no competing interests to declare that are relevant to the content of this article.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Aksnes, D. W., & Rip, A. (2009). Researchers' perceptions of citations. *Research Policy*, *38*(6), 895–905. <https://doi.org/10.1016/j.respol.2009.02.001>
- Altomonte, G. (2020). Exploiting ambiguity: A moral polysemy Approach to variation in economic practices. *American Sociological Review*, *85*(1), 76–105. <https://doi.org/10.1177/0003122419895986>
- Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An R-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. <https://doi.org/10.1016/j.joi.2017.08.007>
- Aryadoust, V., Tan, H. A. H., & Ng, L. Y. (2019). A scientometric review of Rasch measurement: The rise and progress of a specialty. *Frontiers in Psychology*, *10*, 2197.
- Banta, T. W. (2007). Can assessment for accountability complement assessment for improvement. *Peer Review*, *9*(2), 9–12.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, *2008*(10), P10008. <https://doi.org/10.1088/1742-5468/2008/10/P10008>
- Brown, J. T. (2017). The seven silos of accountability in higher education: Systematizing multiple logics and fields. *Research & Practice in Assessment*, *11*, 41–58.
- Brown, J. T. (2018). Leading colleges & universities in a new policy era: How to understand the complex landscape of higher education accountability. *Change: The Magazine of Higher Learning*, *50*, 30–39. <https://doi.org/10.1080/00091383.2018.1483175>
- Burt, R. S. (2004). Structural holes and good ideas. *American Journal of Sociology*, *110*(2), 349–399.
- Chen, C. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, *57*(3), 359–377. <https://doi.org/10.1002/asi.20317>
- Cowhitt, T., Butler, T., & Wilson, E. (2019). Using social network analysis to complete literature reviews: A new systematic approach for independent researchers to detect and interpret prominent research programs within large collections of relevant literature. *International Journal of Social Research Methodology*, *0*, 1–14. <https://doi.org/10.1080/13645579.2019.1704356>
- Cowhitt, T., Greany, T., & Downey, C. (2023). Storytelling with networks: Realizing the explanatory potential of network diagrams through the integration of qualitative data. *International Journal of Qualitative Methods*, *22*, 16094069231189368. <https://doi.org/10.1177/16094069231189369>
- Finney, S. J., Mathers, C. E., & Myers, A. J. (2016). Investigating the dimensionality of examinee motivation across instruction conditions in low-stakes testing contexts. *Research & Practice in Assessment*, *11*, 5–17.
- Ghamari, H., & Sharifi, A. (2021). Mapping the evolutions and trends of literature on wayfinding in indoor environments. *European Journal of Investigation in Health Psychology and Education*, *11*(2), 585–606. <https://doi.org/10.3390/ejihpe11020042>
- Godin, B. (2006). On the origins of bibliometrics. *Scientometrics*, *68*(1), 109–133. <https://doi.org/10.1007/s11192-006-0086-0>

- Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *The Annals of Family Medicine*, 13(6), 554–561. <https://doi.org/10.1370/afm.1865>
- He, J.-H., Xie, X.-Y., Ling, C.-D., & Feng, W. (2023). Look back, move forward: A bibliometric review of team literature. *Group Dynamics-Theory Research and Practice*, 27(1), 1–27. <https://doi.org/10.1037/gdn0000192>
- Heider, F. (1946). Attitudes and cognitive organization. *The Journal of Psychology*, 21(1), 107–112. <https://doi.org/10.1080/00223980.1946.9917275>
- Huffman, L., Adamopoulos, A., Murdock, G., Cole, A., & McDermid, R. (2011). Strategies to motivate students for program assessment. *Educational Assessment*, 16(2), 90–103. <https://doi.org/10.1080/10627197.2011.582771>
- Javed, M. A., Younis, M. S., Latif, S., Qadir, J., & Baig, A. (2018). Community detection in networks: A multidisciplinary review. *Journal of Network and Computer Applications*, 108, 87–111. <https://doi.org/10.1016/j.jnca.2018.02.011>
- Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, 43, 5–25.
- Lin, C.-L., Zhu, Y.-H., Cai, W.-H., & Su, Y.-S. (2022). Recent synergies of machine learning and neuro-robotics: A bibliometric and visualized analysis. *Symmetry*, 14(11), 2264.
- Liu, O. L. (2011). Outcomes assessment in higher education: Challenges and future research in the context of voluntary system of accountability. *Educational Measurement: Issues and Practice*, 30(3), 2–9. <https://doi.org/10.1111/j.1745-3992.2011.00206.x>
- Lloyd, J., Patterson, T., & Muers, J. (2016). The positive aspects of caregiving in dementia: A critical review of the qualitative literature. *Dementia*, 15(6), 1534–1561. <https://doi.org/10.1177/1471301214564792>
- Luis Sanz-Cabanillas, J., Ruano, J., Gomez-Garcia, F., Alcalde-Mellado, P., Gay-Mimbrera, J., Aguilar-Luque, M., Maestre-Lopez, B., Gonzalez-Padilla, M., Carmona-Fernandez, P. J., Velez Garcia-Nieto, A., & Isla-Tejera, B. (2017). Author-paper affiliation network architecture influences the methodological quality of systematic reviews and meta-analyses of psoriasis. *PLoS ONE*, 12(4), e0175419. <https://doi.org/10.1371/journal.pone.0175419>
- Mazzu, M. F., Baccelloni, A., & Finistauri, P. (2022). Uncovering the effect of european policy-making initiatives in addressing nutrition-related issues: A systematic literature review and bibliometric analysis on front-of-pack labels. *Nutrients*, 14(16), 3423. <https://doi.org/10.3390/nu14163423>
- McLendon, M. K., Hearn, J. C., & Deaton, R. (2006). Called to account: Analyzing the origins and spread of state performance-accountability policies for higher education. *Educational Evaluation and Policy Analysis*, 28(1), 1–24. <https://doi.org/10.3102/01623737028001001>
- Miller, G. N. S., & Morphew, C. C. (2017). Merchants of optimism: Agenda-setting organizations and the framing of performance-based funding for higher education. *The Journal of Higher Education*, 88(5), 754–784. <https://doi.org/10.1080/00221546.2017.1313084>
- Moreira Silva, J. T., Ablanedo-Rosas, J. H., & Rossetto, D. E. (2019). A longitudinal literature network review of contributions made to the academy over the past 55 years of the IJPR. *International Journal of Production Research*, 57(15–16), 4627–4653. <https://doi.org/10.1080/00207543.2018.1484953>
- Newman, M. E. J., & Girvan, M. (2004). Finding and evaluating community structure in networks. *Physical Review E*. <https://doi.org/10.1103/PhysRevE.69.026113>
- Ophir, Y., Walter, N., Walter, D., Velho, R. M., Lokmanoglu, A. D., Pruden, M. L., & Andrews, E. A. (2023). Vaccine hesitancy under the magnifying glass: A systematic review of the uses and misuses of an increasingly popular construct. *Health Communication*, 38(10), 2106–2120. <https://doi.org/10.1080/10410236.2022.2054102>
- Ortagus, J. C., Kelchen, R., Rosinger, K., & Voorhees, N. (2020). Performance-based funding in american higher education: A systematic synthesis of the intended and unintended consequences. *Educational Evaluation and Policy Analysis*, 42(4), 520–550. <https://doi.org/10.3102/0162373720953128>
- Osborne, D., & Gaebler, T. A. (1992). *Reinventing government: How the entrepreneurial spirit is transforming the public sector*. Basic Books.
- Phan Tan, L. (2021). Mapping the social entrepreneurship research: Bibliographic coupling, co-citation and co-word analyses. *Cogent Business & Management*, 8(1), 1896885. <https://doi.org/10.1080/23311975.2021.1896885>
- Pollock, M., Fernandes, R., Becker, L., Pieper, D., & Hartling, L. (2023). Chapter V: Overviews of reviews. In *cochrane handbook for systematic reviews of interventions* (version 6.4). Cochrane. www.training.cochrane.org/handbook
- Radhakrishnan, S., Erbis, S., Isaacs, J. A., & Kamarthi, S. (2017). Novel keyword co-occurrence network-based methods to foster systematic reviews of scientific literature. *PLoS ONE*, 12(3), e0172778.

- Raza, B., Ahmad, R., Nasir, M. H. N. B. M., & Fauzi, S. S. M. (2021). Socio-technical congruence as an emerging concept in software development: A scientometric analysis and critical literature review. *IEEE Access*, *9*, 129051–129077.
- Raza, S. A., & Govindaluri, S. M. (2021). Omni-channel retailing in supply chains: A systematic literature review. *Benchmarking: An International Journal*, *28*, 2605–2635.
- Ribeiro, H., Barbosa, B., Moreira, A. C., & Rodrigues, R. (2022). *Churn in services—A bibliometric review*.
- Rieber, R. W. (1997). *The Collected Works of L.S. Vygotsky: The history of the development of higher mental functions: Vol. Four*. Plenum Press. <https://books.google.com/books?hl=en&lr=&id=Wvy8rt0kzf0C&oi=fnd&pg=PA1&dq=the+genesis+of+higher+mental+function&ots=anCc2lh2nE&sig=t8HaekRqX2mmMsDmV5hegz9hw60>
- Rossetto, D. E., Bernardes, R. C., Borini, F. M., & Gattaz, C. C. (2018). Structure and evolution of innovation research in the last 60 years: Review and future trends in the field of business through the citations and co-citations analysis. *Scientometrics*, *115*(3), 1329–1363. <https://doi.org/10.1007/s11192-018-2709-7>
- Rousseau, R. (2014). Forgotten founder of bibliometrics. *Nature*, *510*(7504), 218–218.
- Small, H. G. (1978). Cited documents as concept symbols. *Social Studies of Science*, *8*(3), 327–340. <https://doi.org/10.1177/030631277800800305>
- Sobral, M., & Pestana, M. H. (2020). Virtual reality and dementia: A bibliometric analysis. *European Journal of Psychiatry*, *34*(3), 120–131. <https://doi.org/10.1016/j.ejpsy.2020.04.004>
- Spellings, M. (2006). *A test of leadership: Charting the future of U.S. higher education*. U.S. Department of Education.
- Stadtfeld, C., & Block, P. (2017). Interactions, actors, and time: Dynamic network actor models for relational events. *Sociological Science*, *4*, 318–352. <https://doi.org/10.15195/v4.a14>
- Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, *84*(2), 523–538.
- Vaz, N., & Araujo, C. (2022). Failure factors in healthcare quality improvement programmes: Reviewing two decades of the scientific field. *International Journal of Quality and Service Sciences*, *14*(2), 291–310.
- Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications*. Cambridge University Press.
- Weerts, D. J., & Ronca, J. M. (2006). Examining differences in state support for higher education: A comparative study of state appropriations for research in universities. *The Journal of Higher Education*, *77*(6), 935–967. <https://doi.org/10.1080/00221546.2006.11778952>
- Williams, P. (2002). The competent boundary spanner. *Public Administration*, *80*(1), 103–124. <https://doi.org/10.1111/1467-9299.00296>
- Wise, V., Wise, S., & Bhola, D. (2006). The generalizability of motivation filtering in improving test score validity. *Educational Assessment*, *11*(1), 65–83. https://doi.org/10.1207/s15326977ea1101_3
- Yu, Q., Shao, H., He, P., & Duan, Z. (2013). World scientific collaboration in coronary heart disease research. *International Journal of Cardiology*, *167*(3), 631–639. <https://doi.org/10.1016/j.ijcard.2012.09.134>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Thomas Cowhitt¹  · Joshua Travis Brown²  · Anthony Lising Antonio³ 

✉ Thomas Cowhitt
Thomas.Cowhitt@glasgow.ac.uk

Joshua Travis Brown
joshua.brown@jhu.edu

Anthony Lising Antonio
aantonio@stanford.edu

- ¹ School of Education, University of Glasgow, 11 Eldon Street, Glasgow G3 6NH, UK
- ² School of Education, Johns Hopkins University, 2800 N. Charles Street, Baltimore, MD 21218, USA
- ³ Graduate School of Education, Stanford University, 520 Galvez Mall, Stanford, CA 94305, USA