Received January 3, 2022, accepted March 10, 2022, date of publication March 14, 2022, date of current version March 28, 2022.

*Digital Object Identifier 10.1109/ACCESS.2022.3159674*

# Networks in Education: A Travelogue Through Five Decades

**MOHAMMED SAQR[1], OLEKSANDRA POQUET[2,3], AND SONSOLES LÓPEZ-PERNAS[4], (Graduate Student Member, IEEE)**

[1]School of Computing, University of Eastern Finland, Joensuu Campus, 80100 Joensuu, Finland
[2]Centre for Change and Complexity in Learning (C3L), University of South Australia, Adelaide, SA 5072, Australia
[3]Center for Research and Interdisciplinarity (CRI), University of Paris and INSERM1284, 75007 Paris, France
[4]Departamento de Sistemas Informáticos, ETSI Sistemas Informáticos, Universidad Politécnica de Madrid, 28031 Madrid, Spain

Corresponding author: Mohammed Saqr (mohammed.saqr@uef.fi)

**ABSTRACT** For over five decades, researchers have used network analysis to understand educational contexts, spanning diverse disciplines and thematic areas. The wealth of traditions and insights accumulated through these interdisciplinary efforts is a challenge to synthesize with a traditional systematic review. To overcome this difficulty in reviewing 1791 articles researching the intersection of networks and education, this study combined a scientometric approach with a more qualitative analysis of metadata, such as keywords and authors. Our analysis shows rapidly growing research that employs network analysis in educational contexts. This research output is produced by researchers in a small number of developed countries. The field has grown more recently, through the surge in the popularity of data-driven methods, the adoption of social media, and themes as teacher professional development and the now-declining MOOC research. Our analysis suggests that research combining networks and educational phenomena continues to lack an academic home, as well as remains dominated by descriptive network methods that depict phenomena such as interpersonal friendship or patterns of discourse-based collaboration. We discuss the gaps in existing research, the methodological shortcomings, the possible future directions and most importantly how network research could help advance our knowledge of learning, learners, and contribute to our knowledge and to learning theories.

**INDEX TERMS** Social network analysis, learning analytics, network science, bibliometrics, education.

## I. INTRODUCTION

Analyzing a phenomenon through the network lens invites a shift in thinking, from collections of independent observations, such as randomly sampled data points, to a relational view upon interdependent interactions between network elements. In the past two decades, such a relational view has gained prominence in scientific research, in part due to the advances in computational capacity to analyze a large number of relationships [1]. These sets of relationships describing a phenomena – networks – have helped discern patterns in texts [2], brain biology [3], human mobility [4], and spread of epidemiological diseases [5], just to name a few. A similar focus on relationships and networked structures between interacting elements of a studied phenomenon has been applied in educational settings. For instance, at the micro-level of mental maps, researchers examined students' conceptual networks [6]; at the meso-level of interpersonal relationships, studies inquire into student social networks [7], or at the macro-level view, educational systems are conceived as networks between staff, parents, policies, principals, and similar relevant stakeholders and objects [8]. These examples demonstrate how versatile the application of network analysis is in educational settings. This versatility is among reasons why the research on the intersection between networks and education is both widely used and interdisciplinary.

The interdisciplinary research on the intersection of networks and education has been established over a long history of scholarly work. More than five decades ago, researchers already noted the potential of using network analysis in education, reflecting on extensive methods in multiple areas of inquiry within this field [9]. Today, this multivocal research area attracts researchers with diverse applications in educational research. Such diversity of research foci,

The associate editor coordinating the review of this manuscript and approving it for publication was Chao-Yang Chen.

approaches, and methodologies is difficult to capture with a traditional synthesis method, such as a systematic review. Instead, we take advantage of scientometric methods to provide an overarching view of research applying network analysis in education. Scientometrics offers a quantitative tool to evaluate the state of the art of a scientific field [10]. Modern scientometrics reaches beyond the count of articles, authors and publication venues to include —inter alia— temporal trends, network analysis, statistical and machine learning modelling of research output [11], [12]. Using scientometrics, we map this broad research area, highlighting its theoretical foundations, major knowledge creators and publication outlets, describing the thematic evolution and interrelationships between research trends. We augment these quantitative methods with a nuanced qualitative analysis to describe the content of the foundational and popular papers, as well as thematic trends [13].

The present study is organized as follows. The next section provides basic understanding of the concept of network and approaches to network analysis, with a particular focus on the applications in educational research. The Related work section describes previous studies reviewing research on networks in education, concluding with the aim of this research review. The Method section details dataset selection and data analysis. Next, we present the review results, concluding with the Discussion where we synthesize observed findings, discuss implications, and note limitations of our work.

## II. BACKGROUND

### A. NETWORK AS A CONCEPT

Analysis of networks is central to several research strands, namely *social network analysis (SNA)*, *network analysis*, and *network science*. Despite some important differences explained below, these strands are not clearly delineated [14]. *SNA* is the oldest among the three, with a long-standing empirical tradition rooted in sociological research, contributing theories and evidence about human social networks [15], [16]. SNA has been mainstream in educational settings due to its focus on the social ties between actors and direct application to student and teacher relationships [17]. Drawing on its rich sociological tradition, SNA often presumes theoretical views about the nature and structure of social relationships using both quantitative and qualitative methods [18]. Some researchers frame their studies in educational research through these theoretical views [19], but more often than not, SNA is used in a much narrower sense – to indicate computational techniques that analyze networks. These techniques provide quantitative metrics that summarize patterns of network ties (known as network edges or links) at the level of the entire network, as well as at the level of parts of the network [20]. Although dyadic relations are the main unit of analysis in a network, SNA researchers have been also interested in the influence of nodes in the network due to their position, providing metrics to quantify various types of influence. In social network research, where network ties represent self-reported relationships between individuals, these metrics may indicate power, access to resources, and popularity [21], [22].

*Network science* has emerged as a research discipline during the last three decades with a focus on the study of complex networks, i.e., network structures that are complex in comparison to trivial structures, such as random graphs, spanning social, biological, and technological networks, just to name a few [23]. The field is interdisciplinary and relied mainly on computational approaches. Although it integrates some sociological concepts such as transitivity, it heavily draws on physics and mathematics (e.g., graph theory). Network scientific research embeds a conceptual assumption that network patterns can reflect universal laws that underpin the structure of various networks of, e.g., proteins, genes, or the web [1], [14], [24]. Many research domains apply selected network techniques, without necessarily subscribing to theoretical assumptions permeating either *SNA* or *network science*. In this instance, *network analysis* is another related area of application where networks of all types are examined, regardless of being "social" [25], and not necessarily based on the assumption that the described structures are universal. As researchers use the keyword *network analysis* to denote any of these three approaches, this review encompasses all network analysis application in educational contexts.

### B. NETWORK APPLICATIONS IN EDUCATION

The flexibility of network methods and the wealth of insights offered by the applications thereof have given rise to the large-scale adoption and a diverse range of applications in education [26]. Quantitative analysis of networks in education covers a wide range of educational contexts. Studying networks of students —how they form, grow, increase, and build their friendship communities, and how such networks shape their behavior has been a common thread in educational research [27]. Similarly, networks of teachers, as well as leadership and organizational networks have been explored by educational researchers, for instance, to understand how they collaborate and innovate [8].

These analyses of interpersonal relationships, have been complemented by a surge of work focused on learner collaboration, style of participation, and patterns of interactions, associated with the research strand focused on computer-supported collaborative learning (CSCL) [28] and computer-mediated communication (CMC). Areas of focus here included exploring how students build learning communities and what roles individuals take within collaborative groups [29]. For instance, network analysis has been applied to investigate knowledge building theory, where interpersonal verbal exchange and argumentation have been extensively studied with network methods to map evolving communities of knowledge, track students' engagement in discussions, and study the knowledge creation process [30]. Patterns of participation, interaction, and collaboration derived from network analysis have also been used to understand and inform different learning designs, especially in participatory and collaborative environments [31].

Driven by the increasing reliance on data-driven methods, researchers employ network analysis to identify metrics describing students' positions in communication, track their engagement and model their academic achievement [32], [33]. Modelling students' interpersonal discourse and dialogic interactions has also been used to visualize communication networks, capture similarity in student discussions, or capture conceptual views of the students [34]–[36].

In parallel to these more conventional applications, recent work has started to extend inquiry that is typical of social science and complex network research to educational settings, particularly those that are digitally mediated. Examples include research focused on network mechanisms (i.e., why digital networks form [37], (identification of network measurements that properly account for time in relational processes [38], and network approaches for the analysis of multivariate psychological survey data [39], [40]. Routinely collected digital data of student location, such as WiFi, have also been analyzed via network approaches to understand student collocation in face-to-face settings in relation to performance [41]. In addition to analysis of relationships between actors and objects in learning settings, network analysis in education has been applied to a vast variety of other markedly different research problems: identification of the structure of different types of navigation with video resources [42], curriculum patterns and pathways [43], similarities in structures of in-course resources [44], and structural patterns in content analysis of various learning related texts through a method called Epistemic Network Analysis (ENA) [35].

## III. RELATED WORK

There have been several attempts to synthesize research involving networks in education. Over ten years ago, Biancani and McFarland [45] conducted a comprehensive review of SNA in higher education research. The authors systematically selected studies using keyword search, with the focus on social relationships between different actors in education. They identified two strands of literature around social relations in higher education: that of faculty collaborations and links between collaborative relationships and productivity, and that of student relations towards interracial ties and peer influences on student outcomes. Biancani and McFarland emphasized that research on social relationships in education has grown but, despite its distinct foci, it did not belong to a recognizable research community. In their review, they further highlighted that research work that treats university as a complex system, with diverse network actors has been largely underdeveloped.

With the increasing of adoption of technology, several literature reviews focused on summarizing studies in technology-mediated environments. Much of the focus of these reviews is on the methods and types of applications, rather than thematic overview of the field. For instance, Sie *et al.* [46] presented a "primer" analysis on the applications of SNA in technology-enhanced learning settings. Using a dataset selected via a flexible search strategy with keywords that included —inter alia— "social science", "psychology", "computer science", and "information science", the authors summarized data collection methods (e.g., survey methods) and offered an overview of the used network metrics (e.g., centrality measures and graph-level measures). They also summarized the types of SNA applications observed in technology-mediated learning, namely, network visualization, analysis of learners' networks, simulations of networks of interactions, and interventions to optimize learning. The review by Cela *et al.* [17] on SNA and education included 37 studies which were reviewed in relation to applications and methodological details of SNA, such as node and tie classification, as well as used metrics and software. Dado *et al.* [26] focused on the methodological approaches for SNA research in computer-supported CSCL. The authors reviewed 89 articles and concluded that SNA offers a worthwhile and appropriate research method for understanding CSCL. Other focused reviews were the ones by Forehlich *et al.* [47], evaluating network research that used mixed methods (i.e., studies combining quantitative and qualitative analysis), and the one by Jan *et al.* [48], covering a small selection of ten network studies focused on communities of practice [49] and communities of inquiry [50].

As much as these extant reviews provide an in-depth examination of specific niche areas of network research in education, a comprehensive view of this growing area of inquiry, its prominent strands, authors, and venues has not been conducted. Such a review is timely to help understand strength and limitations of this research domain. What is more, a coherent overview of network research in education across diverse applications and contexts, as a whole, can help identify an underlying structure of this research area, as well as gaps and future research directions that require more attention, and are otherwise overlooked.

## IV. AIM OF THIS STUDY

Against this backdrop, the current study summarizes network-related research in education. In doing so, we strive to present an overarching quantitative mapping using state of the art scientometrics enriched by qualitative insights observed by the authors through content analysis of selected metadata from the studies [13]. To provide such a mapping this paper describes and analyzes:

1) knowledge production in the field, by examining authors, countries, and institutions, as well as their collaboration, main research themes and contributions.
2) venues through which knowledge were disseminated and their citation impact through the analysis of top journals and conferences as well as their citations, and trends.

3) seminal manuscripts and their contributions, offered through a qualitative review of the historical papers and the papers with highest number of citations.
4) theoretical foundations upon which this research area was established, achieved through the co-citation networks of referenced papers.
5) research themes, trends, and clusters of research, conducted via review of important keywords, keyword clusters as well as their temporal trends.

Prior to describing findings of this comprehensive mapping of research on the intersection of networks and educational research, we detail methods specific to each of the analysis steps.

## V. METHODS

### A. SEARCH QUERY AND DATA SELECTION

The search process to identify the dataset of studies followed PRISMA-S (Preferred Reporting Items for Systematic reviews and Meta-Analyses literature search extension; [51]. The search was performed on the Scopus database, which besides including almost all journals in the Web of Science database, offers a broader coverage of social sciences journals and conferences relevant to our study [52]. Scopus offers a robust database with well-maintained metadata as well as a rigorous quality assurance procedure for the selection and curation of scientific journals or conferences [53], [54]. Scopus was chosen over the newer databases (e.g., lens.org and dimensions.ai), since the criteria for inclusion of articles and coverage in these databases are not clearly documented as Scopus. Several pilot iterations of the search query were performed to ensure the quality of the search. Based on these iterations, we selected two search queries that resulted in the most comprehensive collection of articles covering networks and education. The two search queries were:

1) *(“network analy** ***OR*** *“network* method*”* ***OR*** *“network* science”)* ***AND*** *(“educ*”* Some researchers frame their studies in educational research through these theoretical views *“learn*”* ***OR*** *“teach*”* ***OR*** *“student*”).* The first section included all possible variations of network methods, (e.g., network science, network analysis, network analytics, network analyses), whereas the second section accounted for the keywords related to education or learning contexts (e.g., learning, learners, learner, etc.).
2) *“network*”* ***AND*** *(“computer-supported”* ***OR*** *“CSCL”* ***OR*** *“computer-mediated”).* This query was conducted to obtain research on network analysis within these narrower academic communities. This choice was made as the previous query missed few prominent articles on networks in learning settings due to the predominant use of keywords that differed from those in the first query.

Our inclusion criteria included: 1) original articles in journals, conference proceedings or book chapters, and 2) articles published in English (to allow keyword comparison and aggregation). Our exclusion criteria included: 1) articles published during 2021 (which were excluded to retrieve complete years of manuscripts to allow trend comparisons), 2) non-empirical articles, e.g., reviews, editorials, or opinion articles as they offer no novel research findings.

The search using both queries resulted in 4,517 records retrieved with all available metadata. A total of 34 duplicate articles were removed, bringing the total down to 4,483. Two researchers reviewed a sample of 209 of these articles for inclusion in the study by reading the title, abstract, keywords and publication venues. The interrater agreement (Cohen’s Kappa) was 0.904. The raters reviewed the disagreements, resolved them, and one of them proceeded with the selection process. When the rater was uncertain, the articles were classified as “*maybe*”, and resolved by both raters, using the full text of the article. The final sample included 1,791 articles. The complete process of identification and screening of articles is summarized in Figure 1.

### B. DATA PRE-PROCESSING

A total of 1,791 articles selected for analysis were pre-processed and cleaned to improve the accuracy of the reported results. Author names were scrutinized manually for misspelling, duplicate names of the same author or name changes. Similar keywords were manually combined (e.g., SNA, social network analysis, social network analytics were combined). Furthermore, given the interdisciplinary and diverse nature of research on the intersection of networks and education, we had to further combine some of the keywords for the papers selected in the dataset. Three authors pre-processed the keywords making decisions on how some of the most granular level keywords were merged into larger categories. For compound keywords, such as *teacher social capital*, the keyword was split into two separate keywords, i.e., *teacher* and *social capital*. Figure 7 details examples of 20 keywords that were merged. A cleaning process was also applied to conferences and journal names where such consistency was required, e.g., “Lecture Notes in Computer Science” was merged with “LNCS”, or in cases where different editions of the same conference or workshop were named differently across different years.

### C. ANALYSIS SOFTWARE

The R statistical language with the Bibliometrix package [55], [56] were used for the analysis. Bibliometrix offers extensive tools for extracting, processing, and analyzing bibliometric metadata (e.g., authors, keywords, citations, and countries). Networks were plotted using Gephi, the open source network analysis software [57]. Frequencies, plots, and trends were computed and plotted using R statistical language [56].

### D. DATA ANALYSIS

#### 1) CO-AUTHORSHIP NETWORKS

Co-authorship networks can powerfully summarize and visualize collaboration and scientific production that has

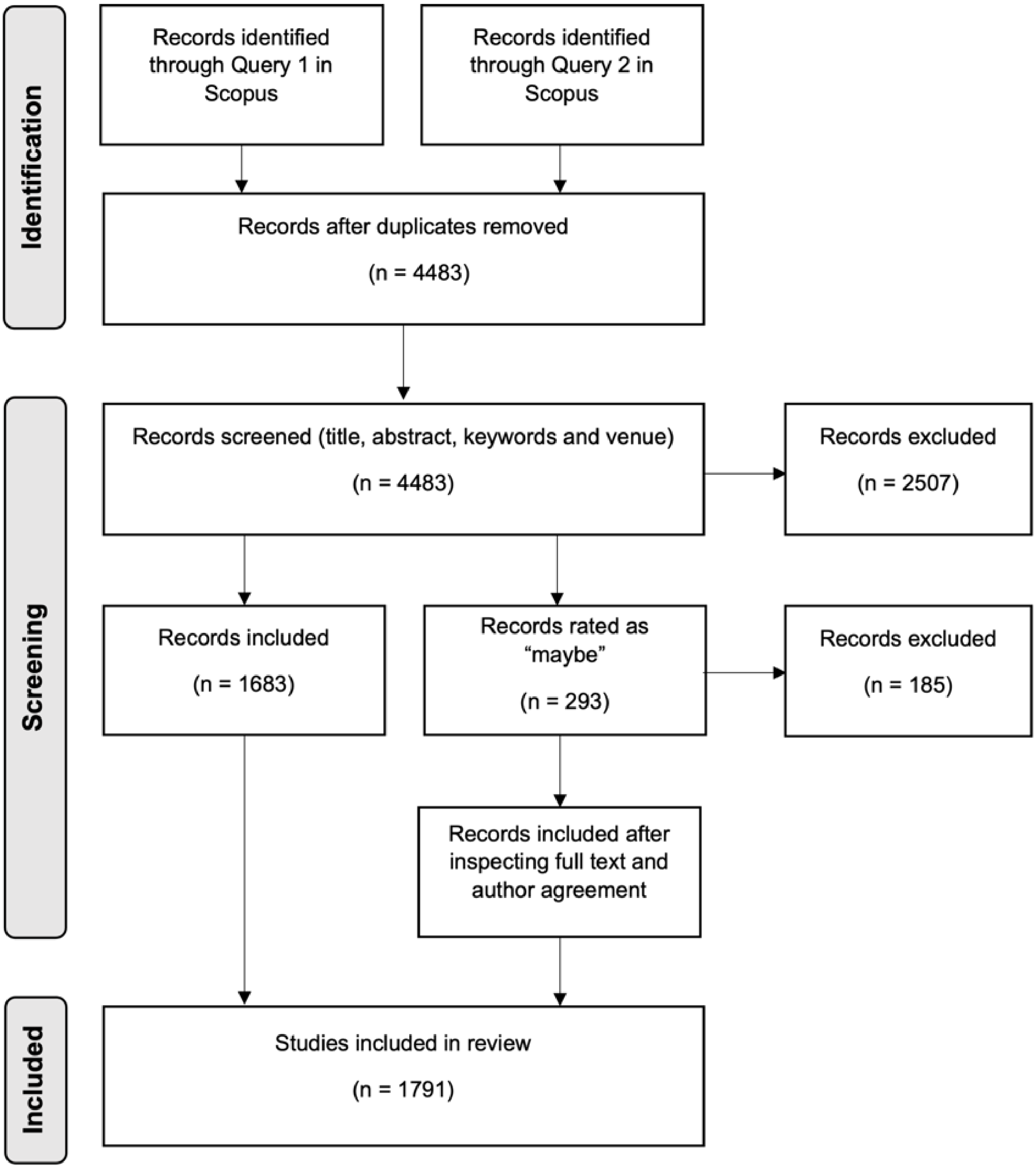



**FIGURE 1. Summary of the process of identification and screening of articles.**

shaped a scientific field. The fractional counting method was selected to construct a weighted co-authorship network, where co-authors of the same paper are considered connected and edge weights are inversely proportional to the number of authors. Such methods prevent assigning more weight to papers co-authored by many researchers [58]. For readability of the network graph, we limited the network to those authors who had minimum three co-authors with whom they collaborated and edge weights of three or more. In other words, a link between any two given authors was shown only

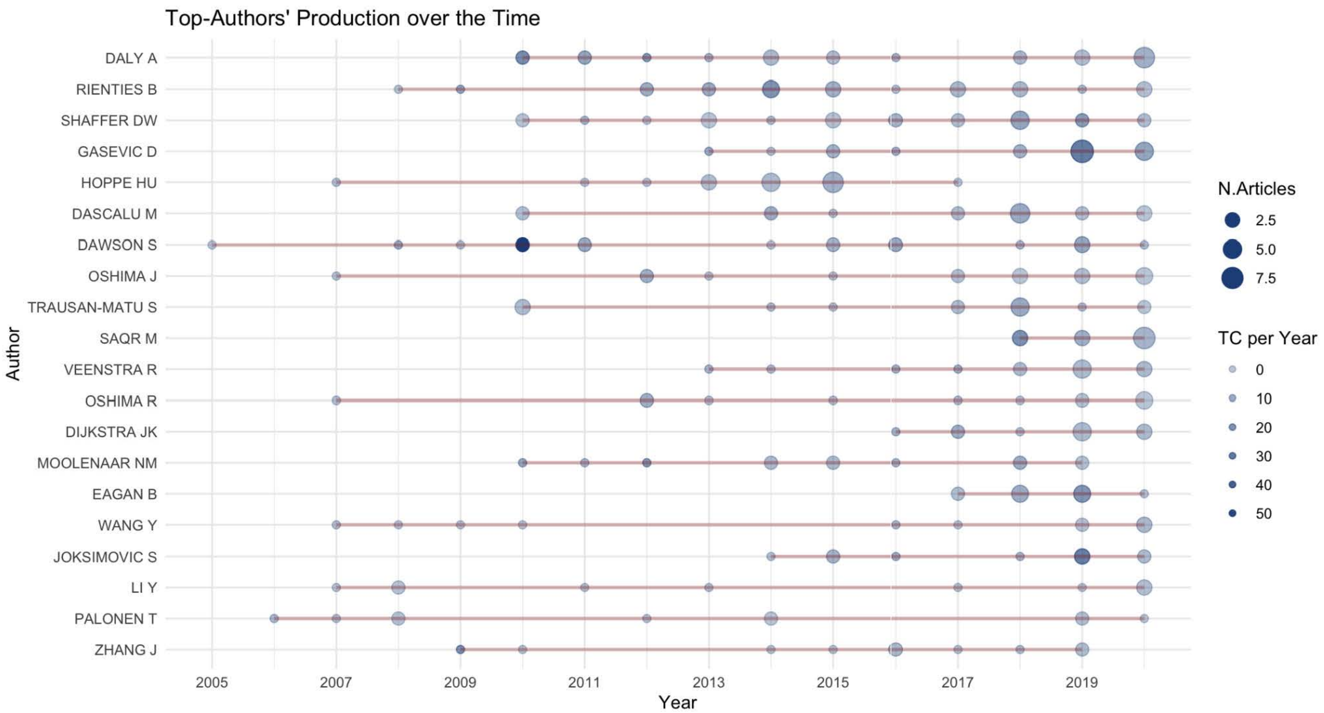



**FIGURE 2.** Top 20 authors with a timeline of their productivity and their number of citations.

if they collaborated at least three times. To map the patterns of co-authorship and the communities of authors who frequently collaborate, we applied Louvain modularity for community detection together [59]. The network was plotted with the Fruchterman Reingold [60] layout algorithm; different communities were colored differently in our visualizations.

### 2) GEOGRAPHY OF AUTHORSHIP AND COLLABORATIONS

As *country* is not a standard field in the article metadata, the country was retrieved according to the authors' affiliation at the time of the article's publication. For country productivity, the country of the corresponding author was considered as the article's country. Frequency of country productivity was calculated to show contribution to the field [55]. Multiple country articles were calculated to show the extent and structure of international collaboration. A country collaboration network was constructed using fractional counting with the affiliations of the contributing authors as the nodes and co-authorship as edges. Community detection was performed using Louvain modularity [59]. The network was plotted using the Fruchterman Reingold algorithm [57], [60].

### 3) CO-CITED REFERENCES

Studies that are frequently cited can be viewed as constituting the theoretical, methodological and research themes grounding the research domain. To identify these building blocks for the research intersecting networks and education, we analyzed groups of papers that were often co-occurring in reference lists. To explain, any two articles cited in the same paper were connected by a network tie [55] resulting in a network of citations co-occurring in the studies we analyzed. The more times two studies appeared together in the reference lists, the more weighted their link would be. Community detection with Louvain modularity was applied to identify sub-groups of papers that were most often co-occurring in reference lists together [59]. The network was plotted using the Fruchterman Reingold [60] algorithm to visualize these co-occurring foundational references.

## VI. RESULTS

Our study spanned over five decades from 1969 to 2020 with a total of 1791 documents (1024 journal articles, 698 conference papers and 69 book chapters). The average number of years since publication in the dataset was 6.2 years, indicating that most studies were published recently. The dataset included 3,821 unique authors. Most of the papers have been produced collaboratively with an average of 2.13 authors per document and only 13.8% of the dataset (248) were single author studies. Articles in our dataset were relatively highly cited with an average of 12.3 citations per article. During the last two decades, the average annual growth rate (average percentage change from year to year) of publications was 20.1%, indicating a significant yearly increase in the number of articles. These numbers suggest that networks in education represent a vibrant area of research, developed by a large number of authors and relatively impactful studies (if judged by the counts of Scopus citations). Table 1 presents a summary of the main results.

**TABLE 1.** Summary of results extracted from the scopus metadata.

| MAIN INFORMATION ABOUT DATA | |
|---|---|
| Timespan | 1969: 2020 |
| Sources (Journals, Books, etc.) | 857 |
| Documents | 1791 |
| Mean years from publication | 6.16 |
| Mean citations per documents | 12.27 |
| Mean citations per year per doc | 1.505 |
| Total references | 62,233 |
| DOCUMENT TYPES | |
| Article | 1024 |
| Book chapter | 69 |
| Conference paper | 698 |
| AUTHORS | |
| Authors | 3821 |
| Author Appearances | 5493 |
| Authors of single-authored documents | 220 |
| Authors of multi-authored documents | 3601 |
| AUTHORS COLLABORATION | |
| Single-authored documents | 248 |
| Documents per Author | 0.469 |
| Authors per Document | 2.13 |

### A. KNOWLEDGE CREATORS

Authors are the primary creators of knowledge whose work can shape the development of a field, drive its advancement, and influence future generations. Understanding author networks help map the extent of collaboration, the breadth of influence, and the links between research groups. Therefore, to conduct our analysis of authors, we considered authors' productivity, co-authorship networks and their countries of affiliation.

#### 1) AUTHORS

The top authors in our sample of articles represent diverse groups from education technology, learning analytics, SNA tools, social sciences, and interdisciplinary researchers. Figure 2 shows the top authors' timeline of productivity and citations. They were predominantly from developed countries. The list of most authors with the highest number of publications in our dataset includes **Alan J. Daly**, Professor of Education, University of California, San Diego, who published SNA studies focused on policy, leadership and professional development of educational actors, such as teachers. **Bart Rienties**, Professor of Learning Analytics, affiliated with the Institute of Educational Technology, the Open University (UK), has worked on applications of SNA in education for over a decade, with contributions to self-reported student networks in educational settings. Within his work on networks and education, the most cited paper highlighted the significant correlation between internal motivation and students' contribution to the discourse [61]. **David Williamson Shaffer** is well known for his work on ENA, a technique foundational to the so-called quantitative ethnography used to represent thematically analyzed student text as networks. Within his work on networks and education, the highest citation paper combines SNA and ENA to study student ties to predict academic performance [62]. **Dragan Gašević**, Professor of Learning Analytics at the Faculty of Education, University of Monash, is a proliferative learning analytics researcher working with a diverse range of methods and topics in education technology and learning analytics. His earliest work and most cited paper in our dataset explored the association between students' social capital and their academic performance [63]. **Heinz Ulrich Hoppe,** Professor of Cooperative and Learning Support Systems at the University of Duisburg-Essen, Germany, has worked on a wide range of SNA applications with a focus of learning design [31] and on bipartite networks in computer-supported collaborative learning [64]. **Mihai Dascalu**, Professor at the Department of Computers, University Politehnica of Bucharest, is known for his work on ReaderBench, a tool that aims at supporting students' and teachers through cohesion-based analysis, identification of reading strategies, and identification and evaluation of textual complexity [65]. **Stefan Trausan-Matu,** Professor of Computer Science at University Politehnica of Bucharest, has published extensively about analysis of student discourse and collaborated with Dascalu on ReaderBench [66]. The two developed a technique for constructing networks where ties between students are based on cohesion metrics of the discourse they shared online. **Shane Dawson** is a Professor of learning analytics at the University of South Australia, a co-founder of learning analytics whose pioneering work analyzed relationships between positioning of learners in online networks, with their perceptions of belonging, their creativity, and academic performance. **Jun Oshima**, Professor at the University of Shizuoka University, is known for his work on Knowledge Building Discourse Explorer (KBDeX), an application that enables the analysis of students' discourse through SNA and is informed by knowledge building theory [30]. **Ritsuko Oshima**, Professor at the University of Shizuoka University, has published extensively in the areas of SNA in education and KBDeX. Professor Oshima is the only female author among the top cited scholars in the area of networks and education. **René Veenstra**, Professor of Sociology at the University of Groningen in the Netherlands, has published in different areas of psychology including social networks in educational settings [67]. **Mohammed Saqr** is a Senior Researcher at the University of Eastern Finland whose work focuses on temporal networks, learning analytics, as well as issues around centrality measures and improving methods [68].

#### 2) AUTHOR COLLABORATION NETWORKS

Figure 3 visualizes co-authorship networks of collaborative groups working on network analysis in education. Co-author clusters are largely based on the country affiliations and geographic proximity, with few international collaborations based on the shared thematic focus. For example, the Dutch research cluster (left) includes authors such as Rienties and Tempelaar who work on social networks, learning analytics, and higher education; Daly, whose work centers

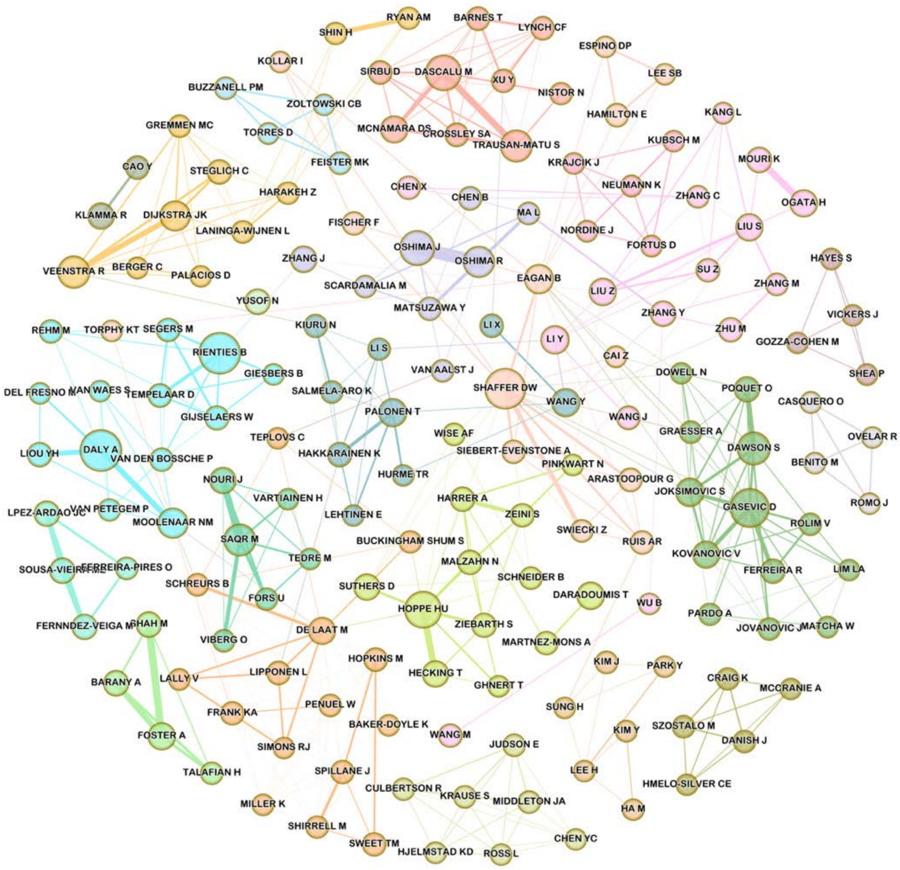



**FIGURE 3.** Constellations of co-authors who work on social networks and education.

on policy networks; and Moolenaar and Van Waes who research teacher networks. The Dutch cluster is linked to international researchers with shared interest in CSCL via Dutch researchers De Laat and Schreurs, towards an interlinked constellation of international authors (Suthers, Hoppe, and Buckhingham Shum). Hoppe brokers between this international community of CSCL researchers and his own research group and German collaborators (Hecking and Harrer). A similar dynamic where thematic national groups led by a prominent researcher are interlinked with international thematic research constellations is well exemplified through the first-degree connections by Shaffer, whose work centers on ENA. Shaffer and some of his national collaborators in ENA (Eagan and Swiecki) are interlinked with the international constellation of learning analytics researchers. This bridging is discernibly natural, given that combining SNA and ENA has been among proposed learning analytics methodologies. The international learning analytics constellation includes prominent scholars, such as Gašević and Dawson, and their national and international collaborators (e.g., Jovanovic, Dowell, and Rolim).

A similar pattern of how thematic and national affiliations overlap in co-authorship networks can be observed in other parts of the network. A knowledge-building research cluster presents researchers working on a distinct approach within CSCL and includes national collaborators (J. Oshima, R. Oshima, and Matsuzawa) who advanced analytical approaches to understand social and semantic networks of the knowledge forum. The cluster includes long-standing international contributors with links to Scardamalia (Canada) —a central scholar of knowledge building—, such as B. Chen, (US) and J. Zhang (Canada). Such patterns of country-based close collaborators linked to international collaborators within their broader research community (e.g., knowledge building nationally and CSCL internationally) can be traced across other research groups. These examples also demonstrate that researchers generally tend to use network approaches within otherwise less connected sub-domains of educational research (ENA, learning analytics, CSCL, professional learning, mixed methods in SNA, or statistical methods in network analysis).

### 3) COUNTRIES AND INSTITUTIONS

Our dataset contained only 61 countries, which is less than one third of all world countries. The most productive countries represented mostly developed countries from Europe and Asia, the US, and Australia. As such, no research on networks in education stems from the Global South., i.e., evidence produced within this research domain today is far from representative. Table 2 shows that the US was the most productive and cited country with 28.4% of all articles and 31.2% of all citations (21.1 citations per article). The

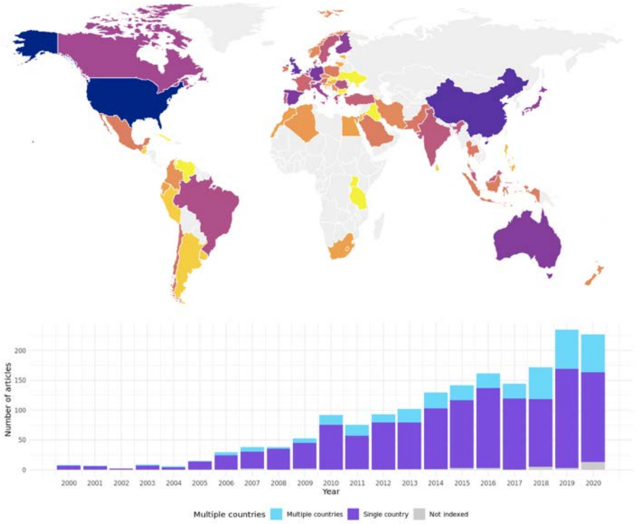



**FIGURE 4.** Frequency of publications per country (top) and evolution of multiple vs. single country publications (bottom).

**TABLE 2.** Top 10 most productive countries (Cit. = citations).

| Country | Articles | Frequency | SCP | MCP | MCP % | Total Cit. | Average Cit. | Ratio Cit. |
|---|---|---|---|---|---|---|---|---|
| UNITED STATES | 324 | 0.284 | 288 | 36 | 0.111 | 6844 | 21.12 | 0.312 |
| CHINA | 127 | 0.111 | 101 | 26 | 0.205 | 787 | 6.20 | 0.036 |
| UK | 86 | 0.075 | 61 | 25 | 0.291 | 1759 | 20.45 | 0.080 |
| GERMANY | 67 | 0.059 | 42 | 25 | 0.373 | 785 | 11.72 | 0.036 |
| SPAIN | 46 | 0.040 | 41 | 5 | 0.109 | 1293 | 28.11 | 0.059 |
| NETHERLANDS | 44 | 0.039 | 26 | 18 | 0.409 | 1144 | 26.00 | 0.052 |
| KOREA | 39 | 0.034 | 30 | 9 | 0.231 | 370 | 9.49 | 0.017 |
| AUSTRALIA | 36 | 0.032 | 28 | 8 | 0.222 | 675 | 18.75 | 0.031 |
| FINLAND | 36 | 0.032 | 26 | 10 | 0.278 | 570 | 15.83 | 0.026 |
| JAPAN | 32 | 0.028 | 26 | 6 | 0.188 | 335 | 10.47 | 0.015 |

second most productive country was China, with 11% of all articles and 3.5% of all citations (6.2 citations per article). European countries occupied the 3rd, 4th, 5th, 6th, and 9th place (United Kingdom, Germany, Spain, Netherlands, and Finland) with 24.4% of all articles and 25.2% of all citations. Korea occupied the 7th place with 3.5% of all articles, Australia with 3.1% and Japan occupied the 10th place with 2.8% of all articles. Around 32% of the included articles were produced with international collaborations (Multiple Country Publication, MCP). Figure 4 shows that there is a strong increase in articles in MCPs in recent years. Netherlands and Germany had the highest MCP with more than one third of their articles based on international collaboration.

The network of countries (Figure 5 - top) and the identified communities reflect the pattern of the authorship, geographical relations, and spoken languages and mirror national authorship patterns observed in the author networks. Whereas the US had a relatively low MCP, it occupied the most central position in the collaboration network with strong links to most highly productive countries. The main communities that can be noticed are a Nordic community represented by Finland, Denmark and Sweden; a large community of

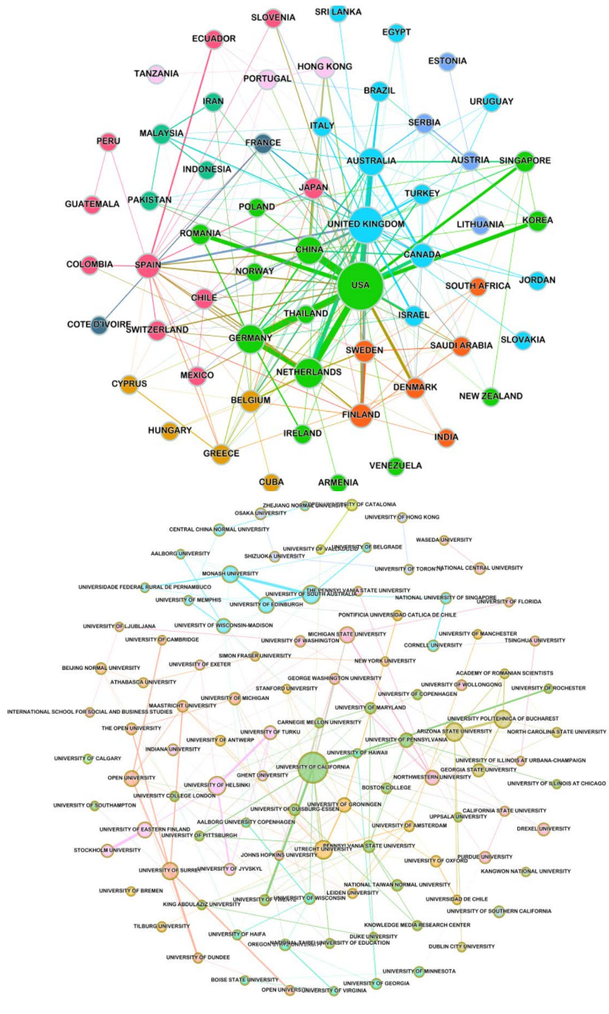



**FIGURE 5.** Collaboration network between countries (top) and institutions (bottom).

Spain and Spanish speaking countries; and a community of UK, Australia and Canada with other countries. The network of institutions (Figure 5 - bottom) was a reflection of the most proliferative authors' affiliations which was dominated (according to number of connections) by University of California, Michigan State University and Arizona State University from USA, University of Eastern Finland and University of Helsinki and from Finland; Utrecht University form the Netherlands, University of South Australia and Monash University from Australia; University Politehnica of Bucharest from Romania and Open University, University of Surrey, and University of Edinburgh from the UK.

### B. KNOWLEDGE DISSEMINATION

The manuscripts in our dataset were distributed over a wide variety of disciplines including education, computer science, social sciences, and interdisciplinary venues making 857 unique venues in total (counting different conference editions as a single venue). The majority of publication venues were not specialized in networked approaches, i.e., some 624 (73%) publication outlets published a single article in this area, and some 116 (13.5%) published two

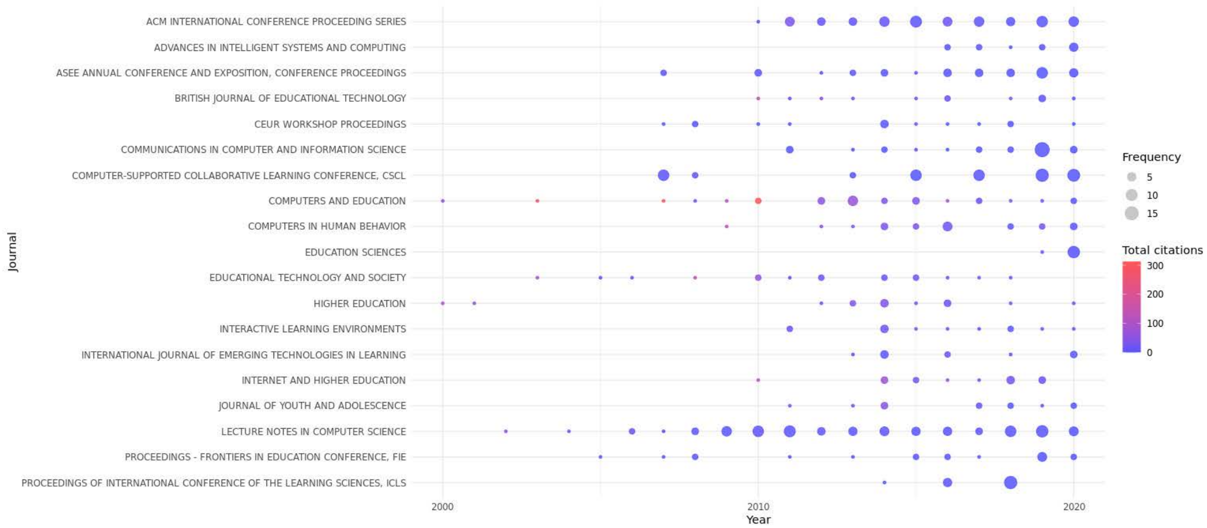



**FIGURE 6.** Top 20 venues of publication.

articles. In other words, these publication venues included an occasional study on networks in education. Most of the articles in our collection were published in journals (57%), followed by conferences (39%) and the rest in book chapters; which is in contrast to most Computer Science sub-domains —where most of the articles in our datasets were published — in which conferences represent the main outlet for research dissemination [69]. It was conference proceedings, however, that occupied the top positions of most frequent venues, as we explain in greater detail below (Figure 6).

The publication venue with this highest number of research papers on networks in educational contexts was *Lecture Notes in Computer Science (LNCS)* with 88 articles (4.9%). *LNCS* publishes proceedings of several conferences in computer science, including those that showcase applications of network science. *ACM Conference Proceedings* came second with 66 articles (3.6%) of which 36 articles were published in the *International Conference on Learning Analytics and Knowledge (LAK)*. *Computer-Supported Collaborative Learning Conference* came third with 56 articles (3.1%). The *American Society for Engineering Education Annual Conference and Exposition* came fourth with 38 articles (2.1%). *Communications in Computer and Information Science* —another conference proceeding series— came fifth with 33 articles (1.8%). Journals with the highest number of research on networks in education include those with considerable impact factors such as *Computers & Education* (29 articles), *Computers in Human Behavior* (21 articles), *Educational Technology and Society,* and *Internet and Higher Education* (15 articles). Overall, journal papers had a higher number of citations per article when compared to conference submissions. We observed that articles published in journals had an average of 18.2 citations, compared to the average of 4.4 citations per conference article. This difference was statistically significant with a moderate effect size: $t(1391.25) = 10.84$, $p < .001$, *Cohen's* $d = .46$. These results suggest that research on networks in education is vastly spread over multiple venues and disciplinary communities. Most of the conference papers were in fields related to computer science. Journals with the highest presence of networks and education studies were those publishing research on education and technology.

### C. MILESTONES AND SEMINAL MANUSCRIPTS

A full picture of the field of networks in education can only be obtained by considering the historical papers that may have shaped the field today. In this section, we briefly describe all the articles that were included in our dataset and published prior to 2000 as well as highly cited seminal papers that have shaped this research area.

#### 1) HISTORICAL PAPERS BEFORE 2000

The oldest article in our dataset dated back to the late sixties [9]. The authors described the process of network analysis at that time as ''may not be easy in face of the time and intellectual demands required''. The authors pointed out the ''possible economic, social or intellectual gains'' (p. 109), and summarized such gains as ''flexible planning, scheduling and control of programmes, operations, procedures and projects with a forecast of time required and a guide where time could be reduced at minimum cost.'' (p. 117). Other gains were summarized by the authors as: organizing courses, simulation of ''educational operations before being committed'' or making better use of resources. The second oldest article in our dataset used network analysis to analyze the syllabus of a course in librarianship and demonstrated how network analysis can help map the connections between learning objectives [70]. The third oldest article [71] used network analysis to help students conceptualize the behavior of organizations. Students were also given exercises to practice network analysis by constructing their own networks. Similarly, Raghavan and Glaser [72] used networks to help students engage in scientific reasoning by using scientific concepts as a network, offering one of the early examples of concept map representations [73]. By the late nineties, computers became more prevalent, and so did network analysis software giving rise to a wave of SNA research, e.g., studying how students' networks form, evolve and explain their differences [27], as well as their preferences for different contexts for learning science [74] or hypermedia systems [75].

#### 2) SEMINAL PAPERS

A review of papers that have the highest number of citations (Table S1) describes the themes of research that have influenced current research directions. The most cited manuscripts in the dataset covered a broad range of applications of network analysis that are common in educational settings, e.g., online communication networks and face-to-face social networks in primary and secondary school settings [76], [77]; teacher and principal networks [8], university student interaction networks in online settings [78], [79], adult learner networks in MOOCs [80], as well as networks with nodes that are non-human, such as brain regions [3]. Few highly cited papers focused on using SNA to derive metrics that can predict learner achievement. For instance, the most cited paper is Macfadyen and Dawson's [81], which analyzed digital learner traces collected from learner activity in online discussion forums to predict performance and at-risk students. A similar application in the top cited papers is that by Romero *et al.* [32] who selected features from discussion forum trace data to predict learner performance. Examples of trace data include indicators of forum activity, such as number of posts made and number of interacting peers, derived from network analysis of online discussions. In addition to using SNA-based discussion forum features to predict achievement [32], [80], [81], other applications of network analysis in top cited papers include:

- combining content analysis [33], [78], [82]–[84], survey information [85], or interviews [79] to better understand the content and nature of online communication ties between learners;
- exploring indicators of the entire network structure at different levels of analysis: communication networks of individual students to observe differences between higher and lower performers [85]; communication networks of groups of learners to understand the relationship between observed structures and group levels of cognitive engagement or quality of knowledge building within the network [76], [77], [82]; communication networks of the same group but in multiple media, to understand the quality of connections [86];
- collecting social self-reported educator networks to understand how network structure relates to the levels of policy implementation [8], [87], [88] or to the levels of student achievement supported by these educators [8];
- examining behavior change of individuals in the network, by implementing interventions targeting actors central to the network [89]; and applying network analysis to networks comprised of non-human nodes, such as analyzing co-occurrence of activated brain regions in response to different words [3].

The different ends of studying communication and interaction versus social networks should be noted. Communication networks in online and other technology-mediated settings were commonly analyzed in relation to the amount and structure of participation, quality of interaction, and their relationship to learning [90]. For instance, Aviv *et al.* [90] found that unattended online discussions, examined as networks, fostered low levels of knowledge construction, whereas scaffolded discussions helped develop cohesive cliques of learners and higher critical thinking manifested within the posts. Similarly, Lipponen *et al.* [76] found that student participation in the discussion was different from their position in communication networks, and that although most posts were related to learning discussions, they were short and of relatively low quality. Zhang *et al.* [77] present a rich study of knowledge building by primary school children. They use network analysis measures at the class-level to understand their relationship to students' level of inquiry and its quality, examined across three experimental group arrangements (fixed composition, fixed with scaffolds, and open emergent). Finally, studies such as that of Martinez *et al.* [79] demonstrate how descriptive network analysis of networks that represented various types of student relationships, including direct communication, as well as networks of similarity due to the use of a shared resource, can be combined with rich qualitative data sources, such as interviews and questionnaires.

In contrast to networks of students interacting online where ties are constructed using digital traces of behavior, in social networks, ties are constructed based on self-reported surveys disseminated to individuals. Highly cited studies of social networks focused on understanding how social structure helps shift adoption of a particular behavior [89]. Several highly cited papers focused on examining social relationships between the educators. For instance, Moolenaar *et al.* [91] examined the relationship between social teacher networks and student achievement and the mediating role of teachers' collective efficacy beliefs in 53 Dutch schools. They found that well-connected teacher networks were interrelated with teacher collective efficacy, and positively affected student achievement. In another study by Moolenaar *et al.* [8], network positioning of school principals was related to the school's innovative climate: the more connected principals were to teachers, the more likely teachers were to adopt new practices.

### D. RESEARCH THEMES

Our close review of keywords explains the main foci of research (social networks and communication networks), as well as main applications and educational contexts. Studies frequently include both keywords that refer to communication between peers in an educational setting (*interaction, collaborative learning*) as well as to relationships between them (*community of learners, social capital*). As previously noted, communication networks and social networks use different data sources for network construction. In many instances, studies of learner social networks and communication networks are informed by different disciplinary theories.

Keywords capturing educational outcomes suggest that much of the research in the area aims to understand how networks relate to learning, where network analysis appears alongside such keywords as *social capital, academic achievement,* and *knowledge-related keywords (e.g., knowledge building* and *knowledge construction*). Expectedly, keywords related to *centrality* measures are frequent, reflecting the prevalence of this research theme which studied whether individual network position links to educational outcomes. Keywords suggest that networks in educational research are predominantly studied in *higher education and online learning settings,* but also outside of formal educational structures *(MOOCs).* Studies of *teacher* networks forms a separate strand of research.

Keywords such as *learning management systems*, *social media,* and *forums* suggest different technological settings where networks are studied. Finally, the presence of keywords as *LA/EDM*, *visualization*, and *design-related* keywords suggests that certain research strands focused on networks in education take strong orientation towards informing practice, be it analytics, mirroring of network-based processes, or informing pedagogical choices.

### 1) KEYWORD TRENDS

Keyword trends explains the increase and decline of popularity in particular research areas (Figure 7). Figure 8 shows that all keywords—except for MOOCs— are on the rise due to the increasing number of publications. To gain additional view of the trends, we plotted the evolution of each keyword's share defined as the fraction of all the articles of a given year that contain each keyword. As presented in Figure S1, the share of keywords related to *SNA*, *LA/EDM*, and *academic achievement* have increased over the last decade, pointing to increasing reliance on network metric analysis to understand interactions and translate network metrics into indicators of academic achievement. Similarly, the *teacher*-related share of keywords is increasingly gaining prominence, indicating the growth of this research theme. *Social media* related keywords —although dropped in the last three years— show an overall increasing share. All the other keywords show a decreasing trend. Worth noting that a down trending keyword share is not synonymous with decreasing research: the keyword may have become so prevalent that it is no longer necessary to use it as a keyword, as is the case with keywords such as *collaborative learning* and *online learning*.

### 2) KEYWORD CO-OCCURRENCE

For an in-depth understanding of the focal thematic areas, we identified frequently co-occurring keywords. A network of co-occurrence was constructed, and keyword communities were identified using Louvain modularity to identify the main research themes. As visualized on Figure 9, nine clusters of interconnected keywords can be observed representing the main themes of SNA research in educational settings. As an interdisciplinary field, the clusters were strongly interconnected with each other. The nodes were interconnected and so were the clusters. A cluster dominated by keywords of *LA/EDM*, *collaboration*, *visualization* represents the emerging theme of using network methods within the field of LA/EDM [32], [68], [81]. Strongly connected with the LA/EDM cluster is the light blue cluster with the keywords *academic achievement, centrality, community detection, friendship*; a reflection of the corpus of research that used networks to study the influence of friendship on behavior and performance [33], [63]. Similarly, the LA/EDM is well connected to the *collaborative learning* cluster with the keywords *LMS* and *CMC* reflecting the technological settings in which networks were studied [76], [82]. Another close and well-connected cluster was the content/text analysis cluster with *ENA* and *assessment* keywords. This cluster reflects the up-trending theme of quantitative ethnography which attempts to offer an alternative method of analysis of coded discourse or interactions [92]. Another cluster is represented by the keywords *community of learners*, *online learning* and *design-related*, reflecting research around pedagogical practices fostering communities of inquiry and practice [93], [94]. We also observe a blue cluster with keywords related to popular contexts (e.g., *forums*, *MOOCs* and *interaction*) reflecting the interest of researchers in studying patterns of participation in MOOCs [95], [96]. A green cluster with keywords *teacher-related*, *leadership* and *professional development* reflecting research targeting different population, e.g., teacher and leaders [97], [98]. The last two clusters reflect dominance research on *social media* [99] and the *knowledge-building* research strands [30].

### 3) THE BUILDING BLOCKS

Citation analysis helped reveal literature strands cited by researchers working on the intersection of social networks and education. Most commonly co-cited groups of papers are described below. These strands have been identified via the application of a community detection algorithm in the network of co-cited papers. We have interpreted these frequently co-cited constellations of studies underlying research on networks and education, as follows (co-cited studies are visualized in Figure 10):

### 4) CLUSTER 1 - NETWORKS IN LEARNING AND CLASSROOM SETTINGS (PINK NODES IN FIGURE 10)

The studies in this cluster represent co-cited work likely to frame research on technology-mediated learning examined via network analysis. The cluster includes seminal texts as social learning theory of the communities of practice by Lave and Wenger [100]; Henri's [101] text referring to electronic records of student text as *a gold mine* for understanding computer-mediated communication unfolding in online environments; Haythornthwaite's [102] empirical and theoretical work on latent network ties in online

| Keyword | Frequency | Evolution | Example Keywords |
|---|---|---|---|
| SOCIAL NETWORK ANALYSIS | 813 | | SOCIAL NETWORK ANALYSIS, SNA, ONLINE SOCIAL NETWORK ANALYSIS |
| COLLABORATIVE LEARNING | 180 | | COLLABORATIVE LEARNING SCENARIOS, CSCL, |
| KNOWLEDGE-RELATED | 147 | | KNOWLEDGE BUILDING, KNOWLEDGE CONSTRUCTION, KNOWLEDGE CREATION, KNOWLEDGE EXCHANGE |
| LA/EDM | 145 | | SOCIAL LEARNING ANALYTICS, MULTIMODAL LEARNING ANALYTICS, DATA MINING, BIG DATA, DATA ANALYSIS |
| TEACHER-RELATED | 144 | | TEACHER EDUCATION, TEACHERS, TEACHER LEARNING, PRE-SERVICE TEACHERS, TEACHER DEVELOPMENT |
| ONLINE LEARNING | 128 | | ELEARNING, DISTANCE LEARNING, ONLINE LEARNING, ONLINE LEARNING ENVIRONMENT |
| SOCIAL NETWORK | 127 | | SOCIAL NETWORKING, SOCIAL NETWORK, SOCIAL NETWORK SITES, ONLINE SOCIAL NETWORKS, SOCIAL NETWORK THEORY |
| INTERACTION | 117 | | INTERACTION ANALYSIS, SOCIAL INTERACTION, PEER INTERACTION, ONLINE INTERACTION, INTERACTION PATTERNS |
| HIGHER EDUCATION | 103 | | HIGHER EDUCATION, INTERNATIONALIZATION OF HIGHER EDUCATION, HIGHER EDUCATION INSTITUTIONS |
| CONTENT/TEXT ANALYSIS | 93 | | CONTENT ANALYSIS, TEXT ANALYSIS, DISCOURSE ANALYSIS |
| FORUM | 89 | | DISCUSSION FORUM, ONLINE FORUMS, FORUM PARTICIPATION, |
| COMMUNITY OF LEARNERS | 86 | | LEARNING COMMUNITIES, COMMUNITY OF PRACTICE, COMMUNITY OF INQUIRY |
| CENTRALITY | 81 | | CENTRALITY MEASURES, DEGREE CENTRALITY, BETWEENNESS CENTRALITY, ETC. |
| SOCIAL MEDIA | 78 | | SOCIAL MEDIA, FACEBOOK, TWITTER, FACEBOOK GROUPS, FACEBOOK NETWORKS |
| LEARNING MANAGEMENT SYSTEM | 77 | | LEARNING MANAGEMENT SYSTEM, MOODLE |
| ACADEMIC ACHIEVEMENT | 73 | | ACADEMIC ACHIEVEMENT, STUDENT PERFORMANCE, ACADEMIC PERFORMANCE |
| MOOC | 57 | | MASSIVE OPEN ONLINE COURSES, HYBRID MOOCS, CMOOCS |
| VISUALIZATION | 54 | | VISUALIZATION, DATA VISUALIZATION, NETWORK VISUALIZATION |
| SOCIAL CAPITAL | 53 | | SOCIAL CAPITAL, SOCIAL CAPITAL THEORY, SOCIAL CAPITAL FORMATION, |
| DESIGN-RELATED | 51 | | LEARNING DESIGN, INSTRUCTIONAL DESIGN, DESIGN-BASED RESEARCH |

**FIGURE 7.** Top 20 keywords and examples of keywords that have been combined.

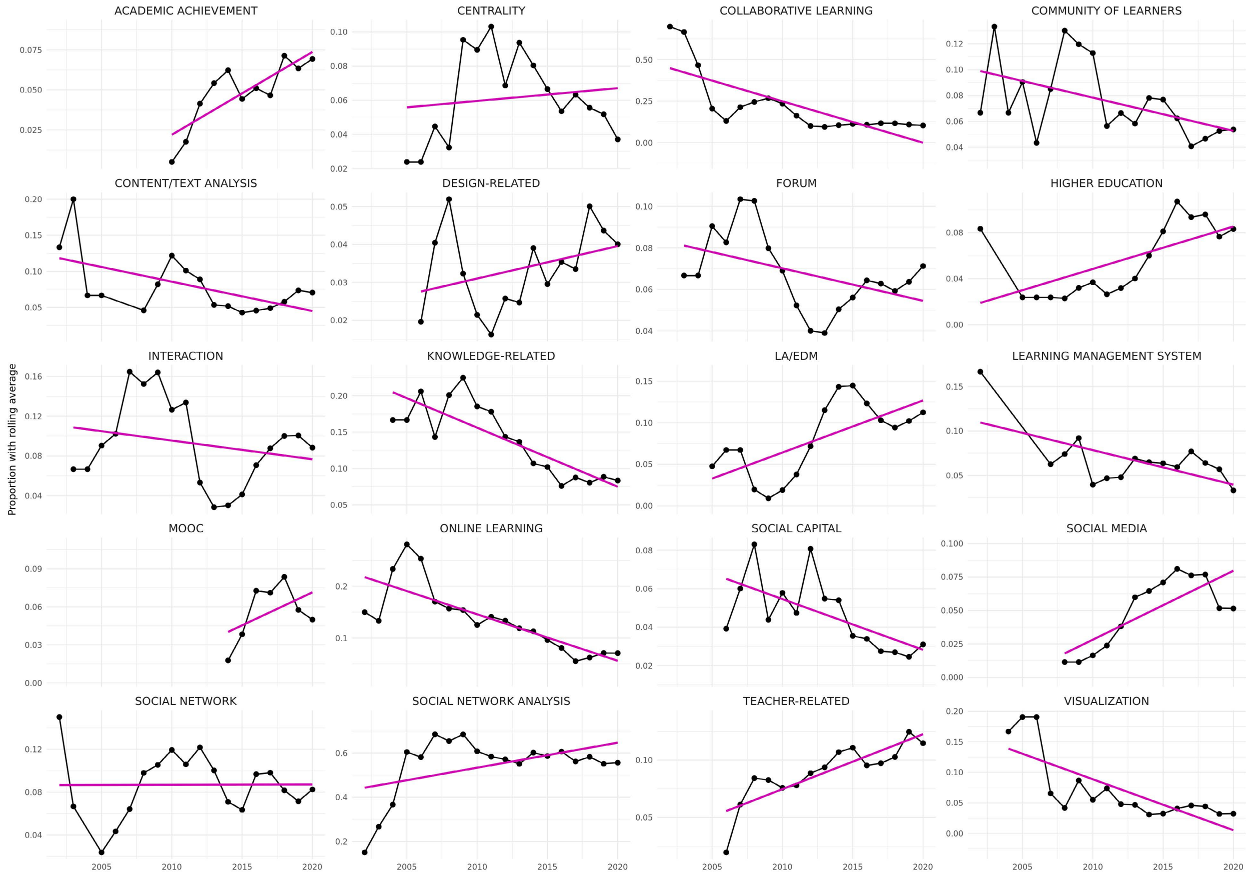



**FIGURE 8. Relative frequency (share) of the top 20 most used keywords throughout the years.**

environments, activated in diverse media use; and community of inquiry model by Garrison *et al.* [103]. These references enable researchers to build on the notions of community, communication, and network ties within the context of online learning. This cluster also contains early applications of network analysis to online environments. These foundational cases include de Laat's [82] early framework combining content analysis of online discussions and network analysis, Dawson's [78] early work on applying network analysis to LMS discussion data, research by Aviv *et al.* [90] focused on knowledge construction in online settings, and the work by De Wever *et al.* [104] that categorizes types of online communication for learning. Finally, this cluster contains references to software and methodological texts most often used within research on networks in learning and classroom settings, by including citations of network analysis software such as Gephi and Igraph, as well as to the seminal methodological [20], [25].

### 5) CLUSTER 2 - NETWORKS IN ORGANIZATIONS (GREEN NODES IN FIGURE 10)

This cluster includes seminal studies that explain how organizational social networks relate to knowledge processes in organizations, as captured through such authors as Hansen or Cross and Parker, among others. Other references in this cluster present foundational work that links social structures captured through networks of relationships with the notions of social capital, creativity, and diffusion of information (seminal work by scholars such as Burt, Lin, Coleman). These references enable researchers to link network positions and structure with examined knowledge processes and individual outcomes.

### 6) CLUSTER 3 - THEORIES OF SOCIAL PROCESSES AND SOCIAL LEARNING (BLUE NODES IN FIGURE 10)

This group of co-cited references contains a potpourri of classic sociological papers on fundamental processes related to peer influences in social networks, covering sociological concepts of homophily [105], selection and influence as mechanisms for tie formation [106], as well as social capital and weak ties [107]. The cluster also contains seminal work on the social cognitive learning theory [108]. This cluster of co-cited references largely refer to theories.

### 7) CLUSTER 4 - SOCIOLOGICAL METHODOLOGIES AND TOOLS (ORANGE NODES IN FIGURE 10)

This cluster contains methodological textbooks and handbooks on SNA, such as those authored by Borgatti, Everett, Freeman; Hanneman and Riddle; Scott. The cluster also contains co-cited references to centrality measures in

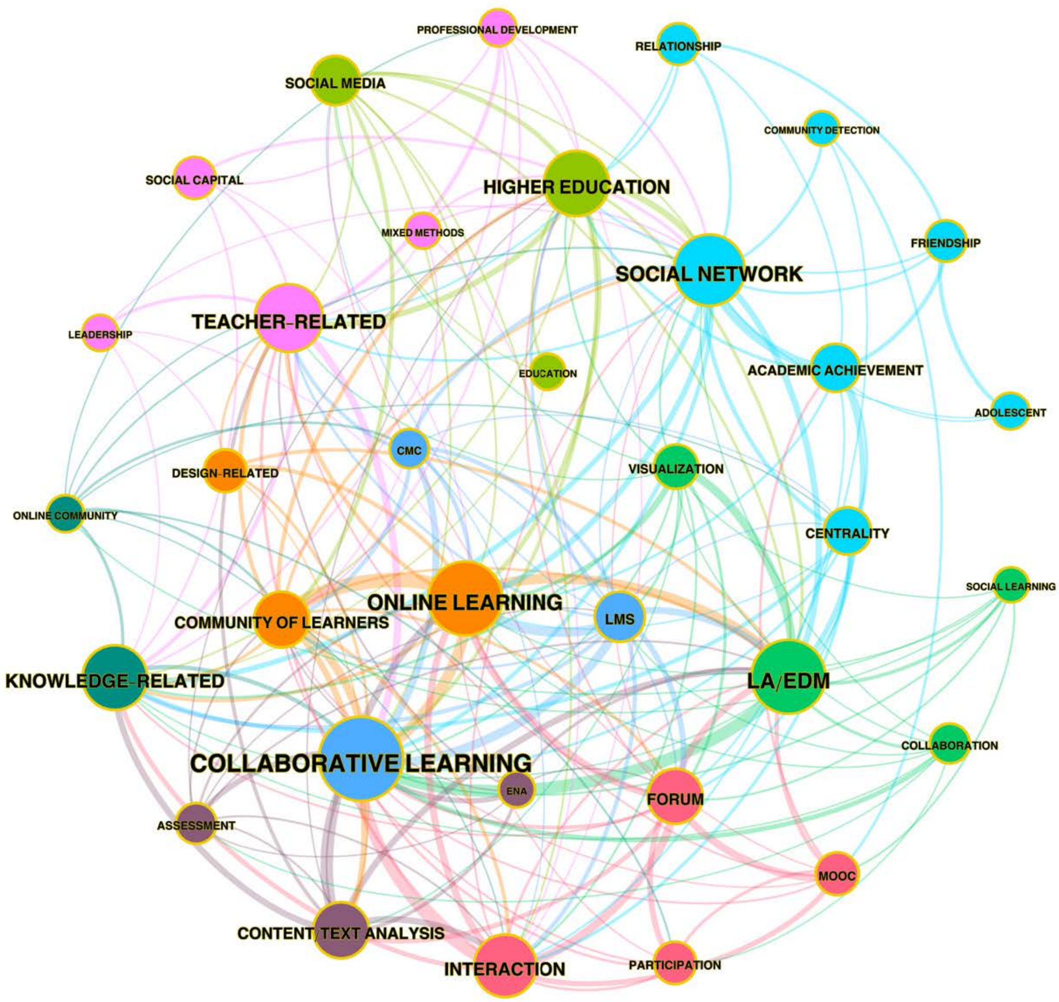



**FIGURE 9.** Network of co-occurrences for most frequent keywords, nodes colored by cluster.

networks and references to some network analysis tools (UCINET, NetDraw) [109]. Given the methodological flavor of this cluster, these studies are likely to be co-cited to support methods sections of existing studies on networks and education.

## VII. DISCUSSION

There is a long and rich history of research on networks in education. Recently, an accelerating pace of network research (with an average growth rate of 20%) has led network scientific production to double in volume in the last five years. Platts and Wyant—authors of the oldest article in our dataset— had high hopes that using networks would bring excellence to education by, e.g., better planning, effective implementations, and efficient usage of resources [9]. Having reviewed this large corpus of research, we see a field that has grown far more diverse in applications and research traditions yet falls short of Platts and Wyant's aspirations.

To provide an overview of research located on the intersection of networks and education, we combined a scientometric approach with qualitative insights. We analyzed relevant studies in relation to major knowledge producers, namely authors and publication venues where research on networks in education can be found. We further analyzed the most cited papers to present important research directions, as well as most co-cited papers to describe foundational research work. Finally, we studied the keywords to describe research themes and their evolution over time.

Our findings indicate that network research in education is diverse, widely used, and interdisciplinary. In our dataset, the majority of articles were collaborative as they were authored by multiple researchers. Publication venues were spread over a large number of journals and conferences in different domains. Despite this seeming disciplinary diversity, the research was predominantly authored by the scholars working in developed countries, and many author affiliations were —unsurprisingly— based on geographic proximity. Therefore, our knowledge about the generality of findings as applied to the remainder two thirds of the world is thin [110]. Given that social traditions, relationships, and interactions vary across cultures and societies, it remains

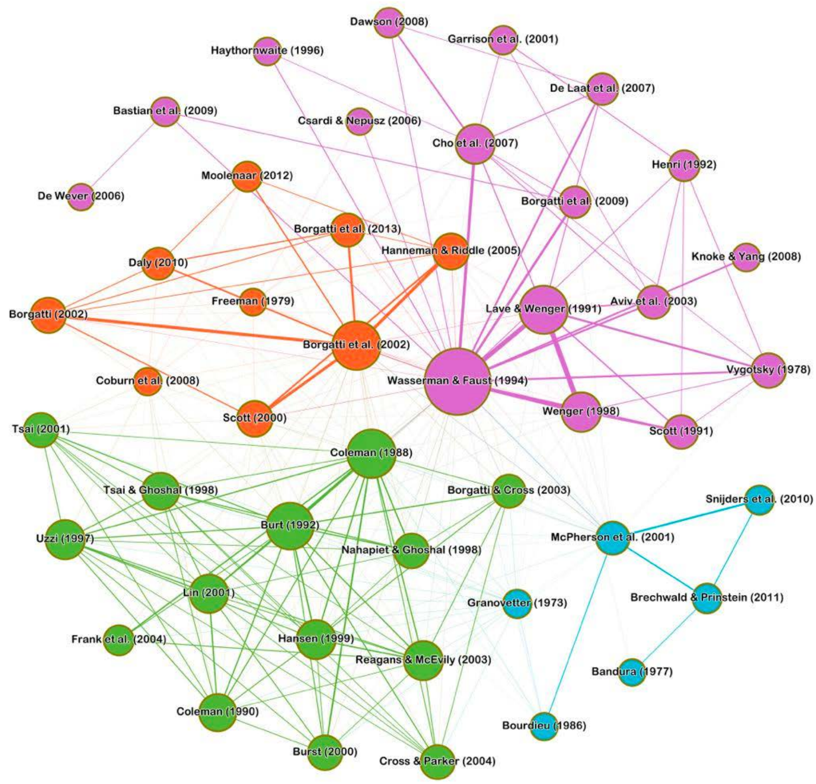



**FIGURE 10. Analysis of most often co-occurring papers cited in the research on social networks and education.**

unclear if today's network research applies beyond its narrow origins [111]. Diversity, reach and extending networks of collaboration could help us gain a better understanding of the full breadth of human behavior [110], [111].

Our findings also suggest fragmentation of research focus on networks and education. Most of the prolific and highly cited researchers in our dataset were not exclusively devoted to network research, few affiliated with network laboratories or specialized centers. The sparse authorship landscape with few devoted researchers could lead to the situation where the few proliferative researchers have an outsized influence on the field. In fact, the analysis of keywords and seminal papers have shown that this may be partially true where most research themes were reflections of top authors' interests. Such a trend however limits the advancement of the research area, which is propelled by the diversity of perspectives and novel ideas [111], [112].

The analysis of research venues mirrors the trend observed in author and keywords: a highly fragmented dissemination landscape with no specialized venue or research outlet. The publication venues in our dataset belonged to 81 different fields according to Scopus, dominated by computer science and its sub-categories (57%), education (41%), social sciences, psychology, e-learning, and human-computer interaction.[1] Whereas the top domains belonged to computer science in which conferences take the upper hand, network research was more likely to be published in journals [69]. The fragmentation of research practices and the lack of agreement on focal elements deemed valuable within the findings can become an obstacle against a coherent understanding of what is known in this research domain. We envision that, moving forward, the presence of a conference representing a community of interest, or a specialized research venue would help drive the field forward and may help settle some grand perspectives of the field or research directions, such as standardized reporting methods, focus on impact or improving replicability [45].

The analysis of research themes through keywords shows a tripartite research structure that connects pedagogical approaches (collaborative learning, knowledge building, peer-based interactions, communities of learners, design-related) with contexts (forums, LMSs, online learning, MOOCs, social media) and network methods (SNA and visualization). Examining keywords' evolution showed that most themes have been increasing in frequency except for MOOCs (that were down-trending both in quantity and ratio). It remains to be seen whether that downtrend of MOOCs continues, if this is a temporary phenomenon, or a trend that simply reflects declining interest from committed proliferative researchers in our dataset. Despite the increasing number of articles for the most keywords, ratios in most of them are decreasing, which is indicative of a growing diversity of themes. Keywords such as *LA/EDM*, *academic*

[1] Scopus assigns multiple categories to the same journal such as medicine and education for a medical education journal.

*achievement*, *higher education*, and *centrality* were strongly interlinked showing the growing interest in capitalizing on network metrics to understand and possibly help predict academic achievement [32], [80], [81]. Analysis also shows growing interest in teacher-related research as well as social media which could be partially attributed to interested sub communities of researchers (see our review of authors) and the use of social media for professional development in communities of practice.

Equally important are the keywords that —against our expectations— showed limited presence in our dataset, despite their overall popularity in other network scientific domains. Appearances of keywords reflecting network inference methods (e.g., Exponential Random Graphs or Stochastic Actor Oriented Models) was lower than one per cent. Network inference methods go beyond describing the network and offer a powerful analytical framework for causal inference and theory formulation that can help understand the processes behind network generation while accounting for the complex dependencies between network elements [106], [113]. In doing so, it allows us to explain why a phenomenon occurs, i.e., why a student chose to engage in an interaction, why and how a collaborative group formed, or why there is an association between an element of discourse and another [37]. Given the complexity, the relational nature, and the multiple dependencies between learners and learning processes, network inference methods seem to offer a much-needed solution that could advance our understanding of learning [114]. Similarly, our data has shown rare occurrences of some of the latest trends in network research, such as psychological networks and graphical gaussian models (less than 0.2%). Temporality has witnessed an increased attention in the last decade, however, "temporal" as a keyword has appeared in fewer than 1% of the articles in our dataset, and the appearance of *temporal networks* is lower than 0.3%. In the same way, bipartite, two-mode, and multi-layer network analysis, and link prediction have all appeared even less, despite being a rising area in network science. Finally, although network science has been around for two decades, fewer than 1% of the articles in our dataset identified as network science [1]. This may indicate either that those educational researchers do not subscribe to the epistemologies associated with network science, or rather that the techniques and approaches used in network science require more advanced disciplinary knowledge and require specialized expertise from educational researchers.

Centrality measures have appeared frequently within our reviewed papers —keywords and abstracts— with strong connection to data-driven applications, friendships, and social capital. However, the reported centralities were limited to the traditional measures, e.g., degree, closeness and betweenness centralities with rare appearance of Katz centrality as well as some others. Some degree centralities are local measures that can be calculated from counting direct edges or neighbors (i.e., degree centrality). Therefore, they require no knowledge of the network or the relational structure [115]. Furthermore, many of centrality measures that are commonly operationalized in educational research are limited in leveraging the insight of the relational information encoded in the network structure. Improving the measurement and interpretation of existing centralities or exploring novel ones within carefully constructed network representations could be helpful in understanding under-examined learning processes, such as diffusion of learning or group cohesion [116].

Our analysis shows that the research intersecting areas of networks and education is slow-moving. Studies lag in methodological diversity, with slow adoption from network science, and rely heavily on descriptive methods and classical metrics developed for social science context, such as traditional centrality measures. This may be understood –at least partially– in the context of fragmentation of knowledge producers within this area of research. We show the lack of research centers and researchers specialized in the area of relational analysis of educational phenomena. This suggests that knowledge of the domain is scattered, and that methodological adoption is ad-hoc.

In conclusion, compared to previous reviews of network-related literature, our study offered a comprehensive overview. Such an overview is inclusive of research strands but offers a less detailed description of the fine-grained details of the entirety of included papers. Such is a trade-off between depth and breadth of coverage that makes both types of synthesis (systematic reviews and scientometrics) rather complementary. Here, our contribution is focused on bringing out aspects unexamined in prior work, such as the state of the field, the authors, countries and their collaboration, the venues, the trends of keywords, and the theoretical underpinning. These higher-level temporal and relational aspects between knowledge producers and research themes broadened our understanding of the current state of this vibrant area of research, enabling to see future gaps.

Our results concur with earlier work highlighting the of organized academic communities [45]. Themes of network research that we report are similar to those of Sie *et al.* [46]. Analysis shows under-studied areas of research, such as the use of network intervention and network simulation is educational research. Similarly, we report similar findings to those of Cela *et al.* [17] regarding research themes and similar to those of Dado *et al.* [26] regarding methodological approaches. Keyword analysis demonstrates novel insights such as the increasing interest in teacher-related research, the decline of MOOCs research, and surging studies of social media and LA/EDM focus. Yet, the conclusions of Dado *et al.* still apply to this larger dataset: descriptive network methods are go-to methods in educational research [26].

The present study has a number of limitations. Bibliographic databases, where we draw metadata analyzed in the study, have deficiencies, inaccuracies, and missing data especially before 2000. To minimize the influence of such problems and improve the accuracy of our research we cleaned author names, the keywords, and the publication venues manually. We have also filtered articles manually to avoid non-relevant research. Bibliometric research is commonly used to rank researchers and assign scores to authors or institutions, which has been criticized [13], [117]. Instead, we have opted out of using popular ranking indices and shifting our focus to a nuanced view of the field, describing most prominent authors, papers, keywords, and trends. These qualitative insights do not describe the entirety of studies as our dataset includes thousands of articles and authors.

## REFERENCES

[1] A. Vespignani, "Twenty years of network science," *Nature*, vol. 558, no. 7711, pp. 528–529, Jun. 2018.

[2] M. Gerlach, T. P. Peixoto, and E. G. Altmann, "A network approach to topic models," *Sci. Adv.*, vol. 4, no. 7, Jul. 2018, Art. no. eaaq1360, doi: 10.1126/sciadv.aaq1360.

[3] K. M. Petersson, A. Reis, A. Castro-Caldas, and M. Ingvar, "Language processing modulated by literacy: A network analysis of verbal repetition in literate and illiterate subjects," *J. Cognit. Neurosci.*, vol. 12, no. 3, pp. 364–382, 2000.

[4] D. Zhang, T. He, F. Zhang, and C. Xu, "Urban-scale human mobility modeling with multi-source urban network data," *IEEE/ACM Trans. Netw.*, vol. 26, no. 2, pp. 671–684, Apr. 2018.

[5] F. Riquelme, A. Aguilera, and A. Inostrosa-Psijas, "Contagion modeling and simulation in transport and air travel networks during the COVID-19 pandemic: A survey," *IEEE Access*, vol. 9, pp. 149529–149541, 2021, doi: 10.1109/access.2021.3123892.

[6] C. S. Q. Siew, "Using network science to analyze concept maps of psychology undergraduates," *Appl. Cognit. Psychol.*, vol. 33, no. 4, pp. 662–668, Nov. 2018.

[7] Z. Boda, T. Elmer, A. Vörös, and C. Stadtfeld, "Short-term and long-term effects of a social network intervention on friendships among university students," *Sci. Rep.*, vol. 10, no. 1, Feb. 2020, Art. no. 2889.

[8] N. M. Moolenaar, A. J. Daly, and P. J. C. Sleegers, "Occupying the principal position: Examining relationships between transformational leadership, social network position, and schools' innovative climate," *Educ. Admin. Quart.*, vol. 46, no. 5, pp. 623–670, Dec. 2010.

[9] C. V. Platts and G. Wyant, "Network analysis and the possibility of its use in education," *Educ. Rev.*, vol. 21, no. 2, pp. 109–119, Feb. 1969.

[10] L. Ivancheva, "Scientometrics today: A methodological overview," *Collnet J. Sci. Inf. Manage.*, vol. 2, no. 2, pp. 47–56, Dec. 2008.

[11] J. Mingers and L. Leydesdorff, "A review of theory and practice in scientometrics," *Eur. J. Oper. Res.*, vol. 246, no. 1, pp. 1–19, Oct. 2015.

[12] S.-U. Hassan, M. Imran, S. Iqbal, N. R. Aljohani, and R. Nawaz, "Deep context of citations using machine-learning models in scholarly full-text articles," *Scientometrics*, vol. 117, no. 3, pp. 1645–1662, Dec. 2018.

[13] D. Hicks, P. Wouters, L. Waltman, S. de Rijcke, and I. Rafols, "Bibliometrics: The Leiden manifesto for research metrics," *Nature*, vol. 520, no. 7548, pp. 429–431, Apr. 2015.

[14] M. Jacomy, "Epistemic clashes in network science: Mapping the tensions between idiographic and nomothetic subcultures," *Big Data Soc.*, vol. 7, no. 2, Jul. 2020, Art. no. 205395172094957.

[15] S. P. Borgatti, A. Mehra, D. J. Brass, and G. Labianca, "Network analysis in the social sciences," *Science*, vol. 323, no. 5916, pp. 892–895, 2009.

[16] J. Scott and P. Carrington, *The SAGE Handbook of Social Network Analysis*. London, U.K.: SAGE, 2014.

[17] K. L. Cela, M. Á. Sicilia, and S. Sánchez, "Social network analysis in E-Learning environments: A preliminary systematic review," *Educ. Psychol. Rev.*, vol. 27, no. 1, pp. 219–246, Mar. 2015.

[18] E. Erikson, "Formalist and relationalist theory in social network analysis," *Sociol. Theory*, vol. 31, no. 3, pp. 219–242, Sep. 2013.

[19] C. Haythornthwaite, "Collaborative work networks among distributed learners," in *Proc. 32nd Annu. Hawaii Int. Conf. Syst. Sci.*, 2016, p. 16.

[20] S. Wasserman and K. Faust, *Social Network Analysis: Methods and Applications*. New York, NY, USA: Cambridge Univ. Press, 1994.

[21] S. P. Borgatti and M. G. Everett, "A graph-theoretic perspective on centrality," *Soc. Netw.*, vol. 28, no. 4, pp. 466–484, 2006.

[22] L. C. Freeman, "Centrality in social networks conceptual clarification," *Soc. Netw.*, vol. 1, no. 3, pp. 215–239, Jan. 1978.

[23] M. Newman, *Networks*. Oxford, U.K.: Oxford Univ. Press, 2018.

[24] K. Bürner, S. Sanyal, and A. Vespignani, "Network science," *Annu. rev. inf. sci. technol.*, vol. 41, no. 1, pp. 537–607, 2007.

[25] D. Knoke and S. Yang, *Social Network Analysis*, 3rd ed. Thousand Oaks, CA, USA: SAGE, 2020.

[26] M. Dado and D. Bodemer, "A review of methodological applications of social network analysis in computer-supported collaborative learning," *Educ. Res. Rev.*, vol. 22, pp. 159–180, 2017.

[27] S. T. Ennett and K. E. Bauman, "Adolescent social networks: School, demographic, and longitudinal considerations," *J. Adolescent Res.*, vol. 11, no. 2, pp. 194–215, Apr. 1996.

[28] R. Rabbany, S. Elatia, M. Takaffoli, and O. R. Zaïane, "Collaborative learning of students in online discussion forums: A social network analysis perspective," in *Educational Data Mining* (Studies in Computational Intelligence), vol. 524, A. Peña-Ayala, Ed. Cham, Switzerland: Springer, 2014, doi: 10.1007/978-3-319-02738-8_16.

[29] J.-A. Marcos-García, A. Martínez-Monés, and Y. Dimitriadis, "DESPRO: A method based on roles to provide collaboration analysis support adapted to the participants in CSCL situations," *Comput. Educ.*, vol. 82, pp. 335–353, Mar. 2015.

[30] J. Oshima, R. Oshima, and Y. Matsuzawa, "Knowledge building discourse explorer: A social network analysis application for knowledge building discourse," *Educ. Technol. Res. Develop.*, vol. 60, no. 5, pp. 903–921, Oct. 2012.

[31] P. A. Haya, O. Daems, N. Malzahn, J. Castellanos, and H. U. Hoppe, "Analysing content and patterns of interaction for improving the learning design of networked learning environments," *Brit. J. Educ. Technol.*, vol. 46, no. 2, pp. 300–316, Mar. 2015.

[32] C. Romero, M.-I. López, J.-M. Luna, and S. Ventura, "Predicting students' final performance from participation in on-line discussion forums," *Comput. Educ.*, vol. 68, pp. 458–472, Oct. 2013.

[33] H. Cho, G. Gay, B. Davidson, and A. Ingraffea, "Social networks, communication styles, and learning performance in a CSCL community," *Comput. Educ.*, vol. 49, no. 2, pp. 309–329, Sep. 2007.

[34] M. Dascalu, S. Trausan-Matu, P. Dessus, and D. S. McNamara, "Discourse cohesion: A signature of collaboration," in *Proc. 5th Int. Conf. Learn. Anal. Knowl.*, Mar. 2015, pp. 350–354.

[35] D. W. Shaffer, W. Collier, and A. R. Ruis, "A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data," *J. Learn. Anal.*, vol. 3, no. 3, pp. 9–45, Dec. 2016.

[36] D. Suthers, "From contingencies to network-level phenomena: Multilevel analysis of activity and actors in heterogeneous networked learning environments," in *Proc. 5th Int. Conf. Learn. Analytics Knowl.*, Poughkeepsie, NY, YSA, Mar. 2015, pp. 368–377.

[37] B. Chen and O. Poquet, "Socio-temporal dynamics in peer interaction events," in *Proc. 10th Int. Conf. Learn. Analytics Knowl.*, Mar. 2020, pp. 203–208.

[38] M. Saqr and J. Nouri, "High resolution temporal network analysis to understand and improve collaborative learning," in *Proc. 10th Int. Conf. Learn. Analytics Knowl.*, Frankfurt Germany, Mar. 2020, pp. 314–319, doi: 10.1145/3375462.3375501.

[39] S. Li, H. Du, W. Xing, J. Zheng, G. Chen, and C. Xie, "Examining temporal dynamics of self-regulated learning behaviors in STEM learning: A network approach," *Comput. Educ.*, vol. 158, Dec. 2020, Art. no. 103987.

[40] M. Saqr and S. Lopez-Pernas, "Idiographic learning analytics: A definition and a case study," in *Proc. Int. Conf. Adv. Learn. Technol. (ICALT)*, Tartu, Estonia, Jul. 2021, pp. 163–165.

[41] V. Kassarnig, A. Bjerre-Nielsen, E. Mones, S. Lehmann, and D. D. Lassen, "Class attendance, peer similarity, and academic performance in a large field study," *PLoS ONE*, vol. 12, no. 11, Nov. 2017, Art. no. e0187078, doi: 10.1371/journal.pone.0187078.

[42] N. Mirriahi, D. Liaqat, S. Dawson, and D. Gašević, "Uncovering student learning profiles with a video annotation tool: Reflective learning with and without instructional norms," *Educ. Technol. Res. Develop.*, vol. 64, no. 6, pp. 1083–1106, Dec. 2016.

[43] R. Baker, "Understanding college students' major choices using social network analysis," *Res. Higher Educ.*, vol. 59, no. 2, pp. 198–225, Mar. 2018.

[44] A. Whitelock-Wainwright, Y.-S. Tsai, K. Lyons, S. Kaliff, M. Bryant, K. Ryan, and D. Gašević "Disciplinary differences in blended learning design: A network analytic study," in *Proc. 10th Int. Conf. Learn. Anal. Knowl.*, Mar. 2020, pp. 579–588.

[45] S. Biancani and D. A. McFarland, "Social networks research in higher education," in *Higher Education: Handbook Theory Research*. Dordrecht, The Netherlands: Springer, 2013, pp. 151–215.

[46] R. L. L. Sie, D. T. Ullmann, K. Rajagopal, K. Cela, M. Bitter-Rijpkema, and P. B. Sloep, "Social network analysis for technology-enhanced learning: Review and future directions," *Int. J. Technol. Enhanced Learn.*, vol. 4, nos. 3–4, pp. 172–190, 2012.

[47] D. E. Froehlich, S. Van Waes, and H. Schäfer, "Linking quantitative and qualitative network approaches: A review of mixed methods social network analysis in education research," *Rev. Res. Educ.*, vol. 44, no. 1, pp. 244–268, Mar. 2020.

[48] S. K. Jan, P. Vlachopoulos, and M. Parsell, "Social network analysis and online learning communities in higher education: A systematic literature review," *Online Learn.*, vol. 23, no. 1, pp. 249–264, Mar. 2019.

[49] E. Wenger, *Communities of Practice: Learning, Meaning, and Identity*. Cambridge, U.K.: Cambridge Univ. Press, 1998.

[50] M. Cleveland-Innes, D. R. Garrison, and N. Vaughan, "The community of inquiry theoretical framework," in *Handbook Distance Education*. New York, NY, USA: Routledge, 2018, pp. 67–78.

[51] M. L. Rethlefsen, P.-S. Group, S. Kirtley, S. Waffenschmidt, A. P. Ayala, D. Moher, M. J. Page, and J. B. Koffel, "PRISMA-S: An extension to the PRISMA statement for reporting literature searches in systematic reviews," *Syst. Rev.*, vol. 10, no. 1, p. 39, Jan. 2021.

[52] M. Norris and C. Oppenheim, "Comparing alternatives to the web of science for coverage of the social sciences' literature," *J. Informetrics*, vol. 1, no. 2, pp. 161–169, Apr. 2007.

[53] J. Baas, M. Schotten, A. Plume, G. Câté, and R. Karimi, "Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies," *Quant. Sci. Stud.*, vol. 1, no. 1, pp. 377–386, Feb. 2020.

[54] V. K. Singh, P. Singh, M. Karmakar, J. Leta, and P. Mayr, "The journal coverage of web of science, scopus and dimensions: A comparative analysis," *Scientometrics*, vol. 126, no. 6, pp. 5113–5142, Jun. 2021.

[55] M. Aria and C. Cuccurullo, "Bibliometrix : An R-tool for comprehensive science mapping analysis," *J. Inf.*, vol. 11, no. 4, pp. 959–975, Nov. 2017.

[56] R Core Team. (2018). *R: A Language and Environment for Statistical Computing*. Accessed: Nov. 7, 2021. [Online]. Available: https://www.r-project.org

[57] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Proc. 3rd Int. AAAI Conf. weblogs Social Media*, San Jose, CA, USA, 2009, pp. 361–362.

[58] A. Perianes-Rodriguez, L. Waltman, and N. J. van Eck, "Constructing bibliometric networks: A comparison between full and fractional counting," *J. Informetrics*, vol. 10, no. 4, pp. 1178–1195, Nov. 2016.

[59] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti, "Generalized Louvain method for community detection in large networks," in *Proc. 11th Int. Conf. Intell. Syst. Design Appl.*, Nov. 2011, pp. 88–93.

[60] T. M. J. Fruchterman and E. M. Reingold, "Graph drawing by force-directed placement," *Softw. Pract. Exp.*, vol. 21, no. 11, pp. 1129–1164, 1991.

[61] B. Rienties, D. Tempelaar, P. Van den Bossche, W. Gijselaers, and M. Segers, "The role of academic motivation in computer-supported collaborative learning," *Comput. Hum. Behav.*, vol. 25, no. 6, pp. 1195–1206, Nov. 2009.

[62] Z. Swiecki and D. W. Shaffer, "ISENS: An integrated approach to combining epistemic and social network analyses," in *Proc. 10th Int. Conf. Learn. Anal. Knowl.*, Mar. 2020, pp. 305–313.

[63] D. Gašević, A. Zouaq, and R. Janzen, "'Choose your classmates, your GPA is at stake!' The association of cross-class social ties and academic performance," *Amer. Behav. Sci.*, vol. 57, no. 10, pp. 1460–1479, Oct. 2013.

[64] H. U. Hoppe, "Computational methods for the analysis of learning and knowledge building communities," in *Handbook Learing Analytics*. Alberta, BC, Canada: Society for Learning Analytics Research, 2017, pp. 23–33.

[65] M. Dascalu, P. Dessus, Ş. Trausan-Matu, M. Bianco, and A. Nardy, "ReaderBench, an environment for analyzing text complexity and reading strategies," in *Artificial Intelligence in Education*, vol. 7926. Berlin, Germany: Springer, 2013, pp. 379–388.

[66] D. Sirbu, M. Panaite, A. Secui, M. Dascalu, N. Nistor, and S. Trausan-Matu, "ReaderBench: Building comprehensive sociograms of online communities," in *Proc. 19th Int. Symp. Symbolic Numeric Algorithms Sci. Comput. (SYNASC)*, Timisoara, Romania, Sep. 2017, pp. 225–231.

[67] A. Munniksma, T. H. Stark, M. Verkuyten, A. Flache, and R. Veenstra, "Extended intergroup friendships within social settings: The moderating role of initial outgroup attitudes," *Group Processes Intergroup Relations*, vol. 16, no. 6, pp. 752–770, Nov. 2013.

[68] M. Saqr, O. Viberg, and H. Vartiainen, "Capturing the participation and social dimensions of computer-supported collaborative learning through social network analysis: Which method and measures matter?" *Int. J. Comput.-Supported Collaborative Learn.*, vol. 15, no. 2, pp. 227–248, Jun. 2020.

[69] M. Saqr, K. Ng, S. S. Oyelere, and M. Tedre, "People, ideas, milestones: A scientometric study of computational thinking," *ACM Trans. Comput. Educ.*, vol. 21, no. 3, pp. 1–17, Mar. 2021.

[70] V. A. Berkeley, "Audio-visual materials in the library: The relevance of some learning theories to the development of a new syllabus," *J. Librarianship Inf. Sci.*, vol. 8, no. 1, pp. 33–46, 1976.

[71] C. L. Walck and A. T. Jordan, "Using ethnographc techniques in the organizational behavior classroom," *J. Manage. Educ.*, vol. 17, no. 2, pp. 197–217, May 1993.

[72] K. Raghavan and R. Glaser, "Model-based analysis and reasoning in science: The MARS curriculum," *Sci. Educ.*, vol. 79, no. 1, pp. 37–61, 1995.

[73] A. M. von Minden, R. T. Walls, and A. H. Nardi, "Charting the links between mathematics content and pedagogy concepts: Cartographies of cognition," *J. Exp. Educ.*, vol. 66, no. 4, pp. 339–358, Jan. 1998.

[74] J.-S. Choi and J. Song, "Students' preferences for different contexts for learning science," *Res. Sci. Educ.*, vol. 26, no. 3, pp. 341–352, Sep. 1996.

[75] H. Astleitner and D. Leutner, "Applying standard network analysis to hypermedia systems: Implications for learning," *J. Educ. Comput. Res.*, vol. 14, no. 3, pp. 285–303, Apr. 1996.

[76] L. Lipponen, M. Rahikainen, J. Lallimo, and K. Hakkarainen, "Patterns of participation and discourse in elementary students' computer-supported collaborative learning," *Learn. Instruct.*, vol. 13, no. 5, pp. 487–509, Oct. 2003.

[77] J. Zhang, M. Scardamalia, R. Reeve, and R. Messina, "Designs for collective cognitive responsibility in knowledge-building communities," *J. Learn. Sci.*, vol. 18, no. 1, pp. 7–44, Jan. 2009.

[78] S. Dawson, "A study of the relationship between student social networks and sense of community," *Educ. Technol. Soc.*, vol. 11, no. 3, pp. 224–238, 2008.

[79] A. Martínez, Y. Dimitriadis, B. Rubia, E. Gómez, and P. de la Fuente, "Combining qualitative evaluation and social network analysis for the study of classroom social interactions," *Comput. Educ.*, vol. 41, pp. 353–368, Dec. 2003.

[80] N. Gillani and R. Eynon, "Communication patterns in massively open online courses," *Internet Higher Educ.*, vol. 23, pp. 18–26, Oct. 2014.

[81] L. P. Macfadyen and S. Dawson, 'Mining LMS data to develop an 'early warning system' for educators: A proof of concept," *Comput. Educ.*, vol. 54, no. 2, pp. 588–599, Feb. 2010.

[82] M. De Laat, V. Lally, L. Lipponen, and R. J. Simon, "Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for social network analysis," *J. Comput. Supported Collaborative Learn.*, vol. 2, no. 1, pp. 87–103, 2007.

[83] T. Postmes, R. Spears, and M. Lea, "The formation of group norms in computer-mediated communication," *Human Commun. Res.*, vol. 26, no. 3, pp. 341–371, Jul. 2000.

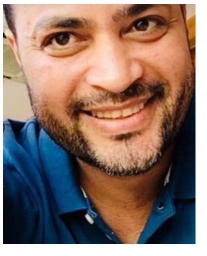
[84] E. Zhu, "Interaction and cognitive engagement: An analysis of four asynchronous online discussions," *Instructional Sci.*, vol. 34, no. 6, pp. 451–480, Nov. 2006.

[85] S. Dawson, "'Seeing' the learning community: An exploration of the development of a resource for monitoring online Student networking," *Brit. J. Educ. Technol.*, vol. 41, no. 5, pp. 736–752, Sep. 2010.

[86] C. Haythornthwaite, "Exploring multiplexity: Social network structures in a computer-supported distance learning class," *Inf. Soc.*, vol. 17, no. 3, pp. 211–226, Jul. 2001.

[87] C. E. Coburn and J. L. Russell, "District policy and Teachers' social networks," *Educ. Eval. Policy Anal.*, vol. 30, no. 3, pp. 203–235, Sep. 2008.

[88] A. J. Daly and K. S. Finnigan, "A bridge between worlds: Understanding network structure to understand change strategy," *J. Educ. Change*, vol. 11, no. 2, pp. 111–138, May 2010.
[89] E. L. Paluck, H. Shepherd, and P. M. Aronow, "Changing climates of conflict: A social network experiment in 56 schools," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 3, pp. 566–571, Jan. 2016.
[90] R. Aviv, Z. Erlich, G. Ravid, and A. Geva, "Network analysis of knowledge construction in asynchronous learning networks," *Online Learn.*, vol. 7, no. 3, pp. 1–23, Mar. 2019.
[91] N. M. Moolenaar, P. J. C. Sleegers, and A. J. Daly, "Teaming up: Linking collaboration networks, collective efficacy, and Student achievement," *Teach. Teacher Educ.*, vol. 28, no. 2, pp. 251–262, Feb. 2012.
[92] B. R. Eagan and D. W. Shaffer, "SENS: Network analytics to combine social and cognitive perspectives of collaborative learning," *Comput. Hum. Behav.*, vol. 92, pp. 562–577, Jul. 2019.
[93] A. Cocciolo, H. S. Chae, and G. Natriello, "Using social network analysis to highlight an emerging online community of practice," in *Proc. 8th Int. Conf. Comput. Supported Collaborative Learn.*, 2007, pp. 145–149.
[94] P. Shea, S. Hayes, J. Vickers, and M. Gozza-Cohen, "A re-examination of the community of inquiry framework: Social network and content analysis," *Internet Higher Educ.*, vol. 13, nos. 1–2, pp. 10–21, 2010.
[95] S. Kellogg, S. Booth, and K. Oliver, "A social network perspective on peer supported learning in MOOCs for educators," *Int. Rev. Res. Open Distrib. Learn.*, vol. 15, no. 5, pp. 263–289, Oct. 2014.
[96] P. Oleksandra and D. Shane, "Untangling MOOC learner networks," in *Proc. 6th Int. Conf. Learn. Anal. Knowl.*, 2016, pp. 208–212.
[97] K. J. Baker-Doyle, "Stories in networks and networks in stories: A tri-modal model for mixed-methods social network research on teachers," *Int. J. Res. Method Educ.*, vol. 38, no. 1, pp. 72–82, 2015.
[98] A. J. Daly, N. M. Moolenaar, C. Der-Martirosian, and Y.-H. Liou, "Accessing capital resources: Investigating the effects of teacher human and social capital on student achievement," *Teachers College Rec., Voice Scholarship Educ.*, vol. 116, no. 7, pp. 1–42, Jul. 2014.
[99] A. Aborujilah, R. M. Nassr, S. B. Musa, and M. Abdullah, "IT students' sentiment of faculty evaluations and posting opinions on social media networks," in *Proc. 12th Int. Conf. Ubiquitous Inf. Manage. Commun.*, Jan. 2018, pp. 1–5, doi: 10.1145/3164541.3164579.
[100] J. Lave and E. Wenger, *Learning in Doing: Social, Cognitive, and Computational Perspectives* (Situated Learning: Legitimate Peripheral Participation). Cambridge, U.K.: Cambridge Univ. Press, 1991.
[101] F. Henri, "Computer conferencing and content analysis," in *Collaborative Learning Through Computing Conferencing*. Berlin, Germany: Springer, 1992, pp. 117–136.
[102] C. Haythornthwaite, "Social network analysis: An approach and technique for the study of information exchange," *Library Inf. Sci. Res.*, vol. 18, no. 4, pp. 323–342, 1996.
[103] D. R. Garrison, T. Anderson, and W. Archer, "Critical thinking, cognitive presence, and computer conferencing in distance education," *Amer. J. Distance Educ.*, vol. 15, no. 1, pp. 7–23, Jan. 2001.
[104] B. De Wever, T. Schellens, M. Valcke, and H. Van Keer, "Content analysis schemes to analyze transcripts of online asynchronous discussion groups: A review," *Comput. Educ.*, vol. 46, no. 1, pp. 6–28, 2006.
[105] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annu. Rev. Sociol.*, vol. 27, no. 1, pp. 415–444, 2001.
[106] T. A. B. Snijders, G. G. van de Bunt, and C. E. G. Steglich, "Introduction to stochastic actor-based models for network dynamics," *Social Netw.*, vol. 32, no. 1, pp. 44–60, Jan. 2010.
[107] M. Granovetter, "The strength of weak ties," *Amer. J. Sociol.*, vol. 78, no. 6, pp. 1360–1380, 1973.
[108] A. Bandura, "Self-efficacy: Toward a unifying theory of behavioral change," *Psychol. Rev.*, vol. 84, no. 2, pp. 191–215, 1977.
[109] S. P. Borgatti and R. Cross, "A relational view of information seeking and learning in social networks," *Manage. Sci.*, vol. 49, no. 4, pp. 432–445, Apr. 2003.
[110] A. Ma, R. J. Mondragân, and V. Latora, "Anatomy of funded research in science," *Proc. Nat. Acad. Sci. USA*, vol. 112, no. 48, pp. 14760–14765, 2015.
[111] M. S. Rad, A. J. Martingano, and J. Ginges, "Toward a psychology of Homo sapiens: Making psychological science more representative of the human population," *Proc. Nat. Acad. Sci. USA*, vol. 115, no. 45, pp. 11401–11405, Nov. 2018.
[112] V. C. Plaut, "Diversity science: Why and how difference makes a difference," *Psychol. Inquiry*, vol. 21, no. 2, pp. 77–99, May 2010.
[113] G. Robins, P. Pattison, Y. Kalish, and D. Lusher, "An introduction to exponential random graph (p*) models for social networks," *Social Netw.*, vol. 29, no. 2, pp. 173–191, May 2007.
[114] S. J. Cranmer and B. A. Desmarais, "Inferential network analysis with exponential random graph models," *Political Anal.*, vol. 19, no. 1, pp. 66–86, 2011.
[115] S. P. Borgatti, M. G. Everett, and J. C. Johnson, *Analyzing Social Networks*. Thousand Oaks, CA, USA: SAGE, 2013.
[116] M. Saqr and S. López-Pernas, "Modelling diffusion in computer-supported collaborative learning: A large scale learning analytics study," *Int. J. Comput.-Supported Collaborative Learn.*, vol. 16, no. 4, pp. 441–483, Dec. 2021.
[117] F. Franceschini, D. Maisano, and L. Mastrogiacomo, "Empirical analysis and classification of database errors in scopus and web of science," *J. Inf.*, vol. 10, no. 4, pp. 933–953, Nov. 2016.



**MOHAMMED SAQR** received the Ph.D. degree in learning analytics from Stockholm University, Sweden. He is currently a Senior Researcher at the University of Eastern Finland. His research interests include network science, big data in education, and science of science. He is particularly focused on social and temporal networks, machine learning, process, and sequence mining as well as temporal processes in general. He is also an active member of several scientific organizations and acts as an academic editor in leading academic publications.

**OLEKSANDRA POQUET** has a background in pedagogy, sociology of education, and learning analytics. She is currently a Researcher at the Centre for Research and Interdisciplinarity (CRI), University of Paris, and an external affiliate of the Centre for Change and Complexity in Learning (C3L), University of South Australia. Her research interests include how human, technological, and artificially intelligent agents can change peer learning behavior and knowledge processes in digital settings.

**SONSOLES LÓPEZ-PERNAS** (Graduate Student Member, IEEE) received the Ph.D. degree in telematics engineering from the Universidad Politécnica de Madrid. She is currently an Assistant Professor at the Universidad Politécnica de Madrid. Her research interests include game-based learning, educational technology, learning analytics, computer science education, bibliometrics, and big data.