IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received December 28, 2021, accepted February 1, 2022, date of publication February 7, 2022, date of current version February 16, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3149812

A Systematic Literature Review of Empirical Research on Epistemic Network Analysis in Education

RAMY ELMOAZEN [1], MOHAMMED SAQR [1], MATTI TEDRE [1], AND LAURA HIRSTO [2]
[1]School of Computing, University of Eastern Finland, 80100 Joensuu, Finland
[2]School of Applied Educational Science and Teacher Education, University of Eastern Finland, 80100 Joensuu, Finland

Corresponding author: Ramy Elmoazen (ramy.elmoazen@uef.fi)

ABSTRACT Over the past decade, epistemic network analysis (ENA) has emerged as a quantitative ethnography tool for modeling discourse in different types of human behaviors. This article offers a comprehensive systematic review of ENA educational applications in empirical studies (n=76) published between 2010 and 2021. We review the ENA methods that research has relied on, the use of educational theories, their method of application, comparisons across groups and the main findings. Our results show that ENA has helped visually model the coded interactions and illustrate the connection strength among elements of network models. The applications of ENA have expanded beyond discourse analysis to several new areas of inquiry such as modeling surveys, log files or game play. Most of the reviewed articles used ENA based on educational theories and frameworks (n=53, 69.7%), with one or more theories per article, while 23 articles (30.3%) did not report theoretical grounding. The implementation of ENA has enabled comparisons across groups and helped augment the insights of other methods such as process mining, however there is little evidence that studies have exploited the quantitative potential of ENA. Most of the reviewed studies used ENA on small sample size with manually coded interactions with few examples of large samples and automated coding.

INDEX TERMS Epistemic network analysis, quantitative ethnography, network analysis, learning analytics, systematic review.

I. INTRODUCTION

Epistemic network analysis (ENA) has emerged as a method for discourse modeling. The method builds on the notion that "the connections between ideas and actions are more significant to the learning process than either ideas or actions separately" [1]. ENA was developed to make sense of such connections using a repertoire of network-based methods that include visualizations and statistical modeling [2]. The network visualization of ENA models the co-occurrence of, for example, codes in discourse, activities in log files, or elements of interaction in a chat [3]. Whereas the method has been conceptualized in education research, it has been used in a wide variety of research questions and applications [4]. Recently, ENA became part of a growing new community of quantitative ethnography that extends to different types of human behavior and applications [4]. Shaffer [5] defines quantitative ethnography (QE) as a strategy that integrates statistical inference with the interpretive capability of qualitative, grounded analysis. In addition to ENA, other methods were developed, such as Shaffer's rho for measuring the inter-rater reliability to improve the level of calculation of agreement between data coders [6]. Moreover, to facilitate coding, the nCoder tool was developed [7]. The package was recently updated to ncoder+ with a semantic add-on to solve the low recall of nCoder [8]. Similarly, the Reproducible Open Coding Kit (ROCK) tool was developed to ease human coding [9].

Shaffer *et al.* [10] argued that although network analysis offers an alternative to traditional statistical methods for modeling collaborative interactions, many network analyses illustrate the nodes' connections of large networks using summary statistics. ENA can help solve this shortcoming by offering an – arguably – better visualization that better summarizes large numbers of nodes in a network. The authors also argued that with traditional network analysis, it is difficult to visually compare two networks if the nodes and edges are not in the same location in a visualization and, therefore,

The associate editor coordinating the review of this manuscript and approving it for publication was Mehdi Hosseinzadeh.

ENA could offer comparable layouts for networks [11]. ENA was then developed to address these issues by using several mathematical principles that aim to quantify the strength of connections and offer fixed layouts as well as several options for comparing networks across groups [2]. Another feature of ENA was its emphasis on modeling dynamic interactions, which showed how and when different codes were shared among collaborators [10]. Due to the flexibility of the method and the presence of free and simple tools, ENA has since been used to model a wide array of topics and problems across several fields [12]–[14].

Recently, the presence of vast amounts of quantitative and qualitative data about learners and their behavior kindled the interest in exploring ENA in diverse applications. For instance, ENA has been used to explore learners' collaborative interactions and their engagement in different learning activities [15]–[19]. Other common uses of ENA include the exploration of professional epistemic frame development among students [12]–[14], [20]–[22]. ENA has also been explored as a method to predict learning performance [23]. Comparison across groups is a common technique to determine how groups differ in their collaboration, thinking, and strategies or approach to learning [24]. Many studies have compared student activities, such as students' collaboration using chat data [15], interactions using online discussions [23], performance in online assignments [25] and students' progress [26].

A scoping review released in 2021 offered a brief overview of papers that studied quantitative ethnography (QE) [4]. Whereas the scoping review addressed the main threads of research, several issues require examination regarding the analysis of research methods, for example, the application of ENA in different educational levels and specialties, the use of educational theories, comparison across groups, and the implication of ENA use in interpretating educational context as well as research findings, for example, impact and contributions. This systematic literature review offers an in-depth review of the topic addressed by the scoping review and addresses the shortcomings thereof. We then offer a much-needed discussion about how ENA has fulfilled its promises. To that end, we present a systematic review that highlights the uses, methods, applications of ENA methods as well as the gaps. The article addresses these issues by answering the following research questions.

- What ENA methods have been used to address educational applications and how?
- What are the main findings [results] from research studies that have employed ENA methods in education?

II. BACKGROUND

Epistemic network analysis (ENA) offers techniques for individual and collaborative contexts that use the so-called epistemic frames theory to model acting and thinking or the behavior of, for example, a community of practice (CoP) [27]. A CoP is a group of people who share a repertoire of knowledge and approaches to similar problems and goals [28]. Individuals reframe their identities and interests in connection to such communities as a result of participation in their practices. The identity of a CoP can be described as epistemic frames with five primary elements: skills, knowledge, identity, values, and epistemology [29]. ENA can model individual and group learning characteristics, such as action, communication, and cognition, by representing them as nodes in an epistemic network. The nodes are connected by edges, and the relative weighting of the edges reflects the strength of association between the nodes [10].

An important step in ENA is to code the dataset of discourse or activities. Coding is the process of bridging two worlds: the world of events and the world of interpretation by investigating how the codes from a Discourse (upper-case D for community discourse) are systematically related to one another in the discourse (lower-case d signifying a person or group of people) [30], [31]. For example, in an urban planner epistemic game, Nash and Shaffer [20] evaluated the extent to which students imitated their mentor's Discourse by determining whether they made the same connections as their mentors over time and if they could develop the ability to think like professional urban planners.

In ENA, the data are divided into segments – called stanza – based on the nature of the data and the research question. Elements within the same stanza are connected and linked together in the ENA model, whereas elements in the different stanzas are not. Chasler et al [32] showed that the co-occurrence of elements in a stanza is important for understanding the meaning of the discourse and offers a good approach to model the cognitive interactions. The main process of ENA starts by creating a matrix that represents the links between codes created by each data line. These matrices are summed to construct the network that is placed in space, where each dimension corresponds to the association between unique pairs of codes to represent the connectivity and strength of the codes. Then the network visualization is generated by aligning the projected points in space [10]. The position of the nodes and the centroid of the network are computed to generate network visualizations. The resulting ENA model contains information about (i) Codes (nodes), which are the people/ concepts connected in the ENA model, (ii) Relations (edges), which is how codes relate to each other, (iii) Stanzas, which are the units of identification based on either time or process, and finally (iv) Evidence, which verifies the connection between codes [33]. ENA can be performed using the web tool or the R package rENA [34]. It is beyond the scope of this review to offer a comprehensive overview of ENA, and readers interested in reading more about the theory and methods are advised to refer to the tutorial by Shaffer *et al.* [10] or Shaffer's textbook on quantitative ethnography [5]. For more about the mathematical foundations, readers are advised to read the work of Bowman *et al.* [2].

III. METHODOLOGY

The authors followed the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA 2020) which is a popular framework widely used across health, social and educational sciences for systematic reviews [35] and the eight essential steps of systematic review by Okoli [36]: (1) identify the purpose; (2) draft protocol; (3) apply practical screen; (4) conduct literature search; (5) extract data; (6) appraise quality; (7) synthesize studies; (8) write the review. The following section presents the main steps and how they were performed in the study.

A. IDENTIFY THE PURPOSE

The authors identified the lack of a comprehensive synthesis of ENA research and the clear need for a systematic review of epistemic network analysis in education based on research questions.

B. DRAFT PROTOCOL

After identifying the purpose and scope of the review, the second critical step prepares the protocol, which is a plan for the review steps to minimize researcher bias during study selection and data extraction [37]. Based on Fink [38], the authors frequently met to draft the protocol by writing down the strategy for conducting the review and practiced following the protocol to ensure complete reproducibility and consistency in the review execution. The protocol included generating the research question, the predefined strategy for the literature search, the search locations, the selection criteria, the assessment of studies, the data extraction strategy, and the planned timetable [37].

C. APPLY PRACTICAL SCREEN

The inclusion and exclusion criteria used in the selection of studies were based on the research questions and guided by previous reviews, for example, [39]. Articles that addressed an empirical ENA problem in education according to the following inclusion and exclusion criteria were included:

1. Articles are written in English.
2. The article is available as a full text and is peer reviewed. Thus, editorials, conference abstracts, and workshop proposals were excluded.
3. The research must be an empirical study that collects and analyzes empirical data with the appropriate methodology and results. Thus, reviews and theoretical or incomplete reports were excluded.

D. LITERATURE SEARCH

We identified five databases covering research at the intersection of social, educational and computer sciences relevant to our research questions: Scopus, Web of Knowledge, Springer, ERIC, and ACM Digital library as well as the first and second editions of international conferences of quantitative ethnography conferences (ICQE 2019 and ICQE 2020). The search formula was selected to cover all existing articles that are relevant to our research question. We used keywords with a wild card to capture all forms of the keyword. Thus, epistemic network* was selected to capture epistemic networks, epistemic networking, and epistemic network; quantitative ethnograph* was selected to capture quantitative ethnography and quantitative ethnographies. To capture keywords related to education, we used educat* to capture keywords based on the same stem, for example, education and educator. Similarly, learn (e.g., learning, learner), teach* (e.g., teaching, teacher), train* (e.g., learning, learner), collaborat* (e.g., collaboration, collaborative), cooperat* (e.g., cooperation, cooperative) and student* (e.g., student, students). Accordingly, the following search formula was used:

"epistemic network ∗" OR "quantitative ethnograph∗") AND ("educat∗" OR "learn∗" OR "teach∗" OR "train∗" OR "collaborat∗" OR "cooperat∗" OR student∗)

The search was conducted from 15 to 20 February 2021. The search yielded 395 articles from all selected databases (129 articles from Scopus, 48 articles from Web of Science, 176 articles from Springer, 23 articles from ERIC, and 19 articles from ACM digital library). All articles were imported into the Rayyan web-based system for analysis. Duplicates were removed, resulting in 291 articles. The abstracts, titles, and keywords of the first 100 manuscripts were independently scanned and reviewed by the first and second authors. The inter-rater agreement was 0.86, and manuscripts that had any conflict were discussed. The disagreements were resolved, and the first author proceeded with the filtering. The authors met to discuss and resolve uncertainties. The title and abstract scan resulted in 146 articles eligible for full-text review, which resulted in 82 eligible studies based on inclusion and exclusion criteria.

E. DATA EXTRACTION

Data extraction is the process by which authors captured the key information and categories of the included studies in the form of a codebook. To increase efficiency, minimize individual variation between reviewers, and reduce error in data analysis, the study adopted a codebook from a previous coding scheme of Kaliisa *et al.* [4] for data extraction and categorization. Furthermore, the authors adopted other categories that related to educational research, such as theory background [28], [40], participants' educational level, coding [41], comparisons, outcomes, and implications for education [42]. Accordingly, the extracted data included (1) year and publication type, (2) sample and population categories, (3) the raw data source in each study, (4) comparisons between study groups, (4) type of coding in each study and its method of application, (5) the theoretical background and (6) the main findings of each article. The coding was initially performed by two coders independently for 10 studies. Thereafter, they discussed the coding challenges and finalized the codebook. One of the coders then continued with coding the articles and met with the second coder to discuss and resolve uncertainties.

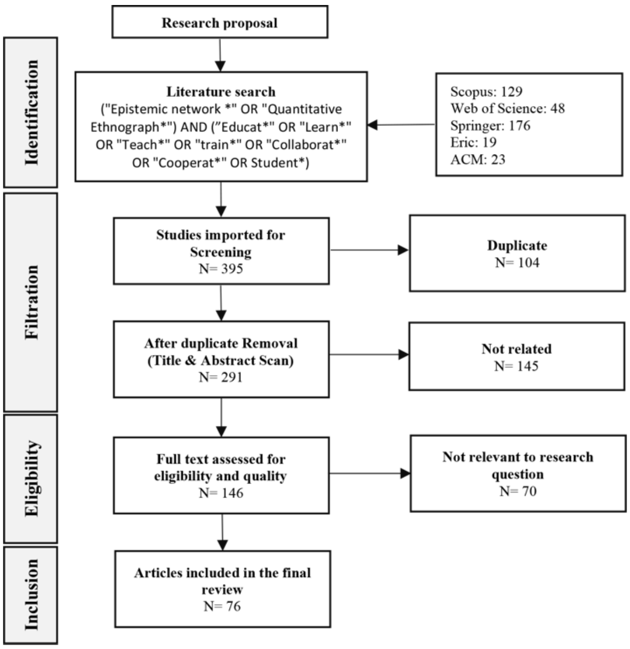

FIGURE 1. The study selection process.

F. QUALITY APPRAISAL

The extracted papers were examined more closely for quality. Four papers were excluded as they failed to meet the quality standards established by Fink [38] for presenting methodology, results, and conclusions. Two articles were excluded as they have the same sets of data published in a different journal and/ or conferences. Ultimately, 76 studies were included in the systematic review (Appendix). The flow diagram of the review process is shown in Figure 1.

G. SYNTHESIZE STUDIES

In this stage, the authors assembled, discussed, and analyzed the data to obtain a comprehensive sense of the collected data. The synthesis stage involves moving from an author-centric to a concept-centric perspective by mapping all data evaluation and incorporating it into the review's hypothesis and structure [43].

IV. RESULTS

In this section we will present the descriptive analysis of the reviewed studies, study populations, source of raw data, comparisons included in the studies, coding, theoretical underpinnings and the main research findings in the reviewed studies.

A. DESCRIPTIVE ANALYSIS OF THE REVIEWED STUDIES

The total number of studies was 76 (Figure 2). Most studies were journal articles (n=43, 56.6%), and the rest were conference papers (n=33, 43.4%). The studies were published between 2010 and 2021. The maximum number of studies in any year was 25 studies in 2019, when the first QE international conference was held in October 2019, followed by 2020 (21 studies) and 2021 (9 studies).

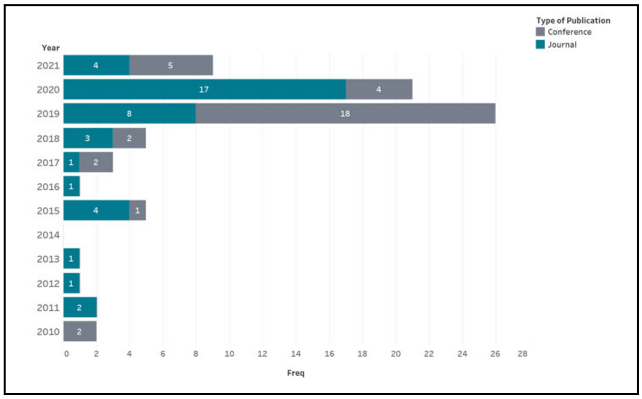

FIGURE 2. Types of publications in different years.

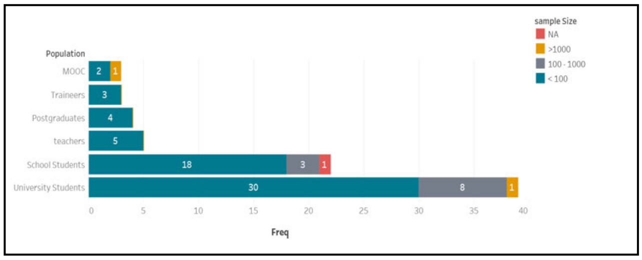

FIGURE 3. Sample size in different population categories.

B. STUDY POPULATION

Most of the included studies were conducted on university students (n=39, 51.3%), followed by 23 articles (30.3%) on school students in all levels, five studies (6.6%) on teachers, postgraduate students in four articles (5.3%), three articles (3.9%) on simulator trainees as a part of continuous medical education, and three articles (3.9%) based on Massive Open Online Courses (MOOCs) (Figure 3). Most studies included a small sample size of less than 100 (n=62, 81.6%), 11 studies (14.5%) had 100-900 participants, and two studies (2.6%) had more than 1000 participants [23], [44]. Only one study did not report its study sample size [45].

C. RAW DATA FORM

The coded data were extracted from a diverse range of sources, for example, online discussions, interviews, learner interactions, surveys, log files, and other miscellaneous forms (Figure 4). ***Online discussions*** were the source in 18 studies (23.7%), in which researchers explored, for example, virtual internship discussions [26], [46]–[50], online discourse [15], [18], [51]–[54] or community of inquiry [23], [55]–[59].

Transcripts of recorded ***interviews*** were used in 16 articles (21.1%) to explore the epistemic frames in the community of practice [13], [60]–[62], games and assignment interviews [14], [20], [21], [63], [64], students' sense-making of feedback [65]–[67] and reading strategies [68], [69].

Learners' interactions were coded in 13 articles (17.1%), other educational activities were coded in ten studies

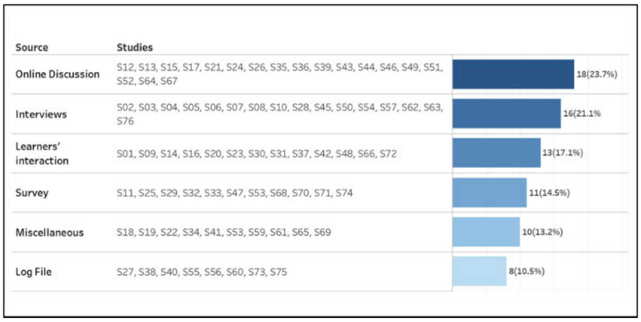

FIGURE 4. Sources of raw data.

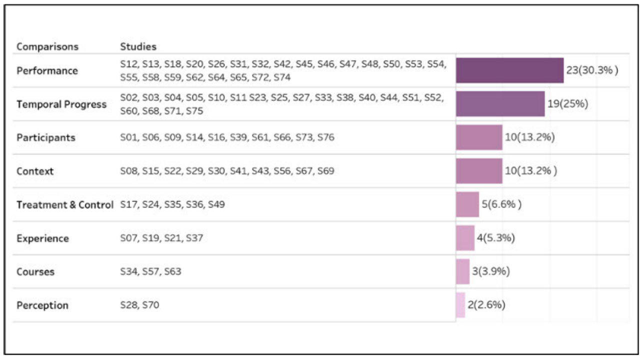

FIGURE 5. Comparisons categories.

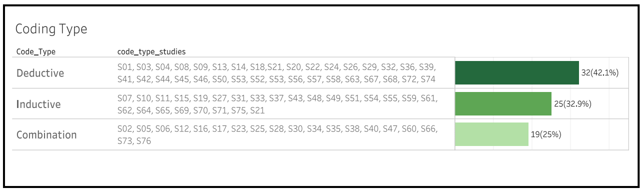

FIGURE 6. Types of coding.

[12], [16], [17], [19], [70]–[75]. Coded data of simulated practices were used in three studies [76]–[78]. ***Survey responses*** were coded in 11 articles (14.5%) either to assess students' gains [79]–[86] or to answer open-ended questionnaires [44], [87], [88].

Coding of data extracted from ***online log data*** was used in eight articles (10.5%) in the form of game-logged data [45], [89]–[92], students' actions in LMS log files [24], [93] and teacher's online logs [94].

Finally, 10 articles (13.2%) were coded based on ***diverse coding schemes***, for example, eye movements [95], [96], diagnostic activities [97], [98], written assignments [25], [99], and coding of online postings [100], [101].

D. COMPARISONS

All 76 reviewed articles included comparisons that can be grouped into one of eight categories (Figure 5). In 23 articles (30.3 %), the comparison was across levels of participants' ***performance***, for example, high vs. low performers [19], [23], [25], [50], [74], [76], [86], student competency level [69], [102], quality of ideas [46], [73], learning outcomes [80], [83], positive vs. negative gain [82], proficiency level [68], accuracy level [96] and correct graph solving [95].

Temporal progress was compared in 19 studies (25%), for example, to identify progress throughout the course [13], [79], [84], [90]–[92], students' epistemic frames after course completion [14], [20], [21], [64], [81], [87] and students' progress throughout different game levels [26], [45], [71], [89].

Context comparison occurred in 10 studies (13.2%) in the form of comparison between virtual and face-to-face mentoring [63], emotions towards online learning [101], 3D animation vs. video only teaching [72], self-performance and puppetry discourse [103], cognition in different multimedia materials [18], learning strategy models [24], [88], dimensions for learning progress diagnosis [97], scaffolding methods [51] and discourse segmentation methods [47]. Comparisons between different ***participant categories*** were reported in 10 studies (13.2%), for example, mentor vs. student epistemic frames [12], [60], [70], dyads vs. individuals' networks [16], [17], short readers vs. long readers [49], and medical vs. teacher students [98].

Comparisons between a ***control group*** with limited scaffolding and a ***treatment group*** with scaffolding instructions was performed in five studies (6.6%) [55]–[59]. Different ***levels of experience*** of participants were compared in four studies [48], [61], [77], [100]. Finally, only three studies (3.9%) compared ***different courses*** [66], [67], [104] and two studies (2.6%) compared ***students' perceptions*** of feedback [44], and LA dashboards [65].

E. CODING

1) CODING TYPE

The coding of the selected articles was deductive, inductive, or a combination of both (Figure 6). The ***deductive coding*** was used in 32 (42.1%) articles with a top-down approach based on codes derived from previous research. Another 25 (32.9%) articles used ***inductive coding*** in a bottom-up approach that was generated from the data with no predefined coding framework using, for example, the grounded theory approach [21], [25], [45], [49], [51], [54], [60], [61], [64], [68], [69], [101], [102]. The third type of coding, which was reported in 19 articles (25%), used ***a combination of coding methods***. Deductive codes were used first and were based on previous studies, followed by the inductive process of adding, extending, and refining codes.

2) CODING METHOD

The coding methods were reported in 67 articles (88.2%), whereas nine (11.8%) articles did not report their method of coding (Figure 7). Manual coding was conducted in the

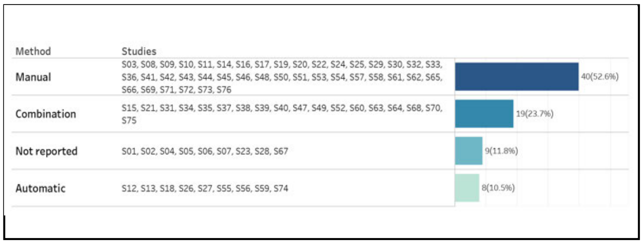

FIGURE 7. Methods of coding.

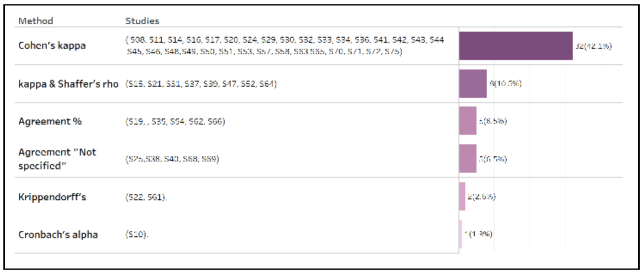

FIGURE 8. Reliability reported in the reviewed studies.

majority of articles (n=40, 52.6 %), most often by two coders (n=35). The remaining articles ranged from a single coder to nine coders. The ***manual coding was supported with software tools*** in 19 articles (25%), for example, ncoder automatic coding [26], [45], [47], [49], [73], [82], MAXQDA 2018 [84], [90]–[92] and NVivo [67], [104], automated coding algorithm [48], [77], Python [44], and linguistic tools (LIWC and Coh-Metrix) [59].

Automatic coding was reported in eight studies (10.5%). For instance, LDAtopic modeling was used to automatically extract topics from the text [15], [23]. Similarly, automatic text coding was performed by identifying the word stem and its conjugations [25], and automatically generating log files of students' actions from learning management system [24], [86], [93], or educational games [89]. Finally, an algorithm was used to analyze eye-tracking during the solving of graph tasks by determining areas of interest (AOIs) [95].

3) RELIABILITY

Fifty-three articles (69.7%) reported agreement statistics between the coders, whereas the remaining articles were either not reported or not applicable (n=23, 30.3%) (Figure 8). The reliability was tested in 32 (42.1%) articles using ***Cohen's kappa***. In addition, ***Shafer's rho statistic*** **alongside the kappa test** were reported in eight articles (10.5%) [26], [47]–[50], [73], [77], [82]. Other reliability tests were used less frequently, such as ***Krippendor?'s test*** in two articles [97], [98] and ***Cronbach's alpha*** in one article [64]. On the other hand, five articles reported only the percentage of agreement between coders [57], [68], [69], [75], [100], and five articles reported that an agreement was tested among coders, without specifying the type of agreement used [79], [84], [90], [91], [101].

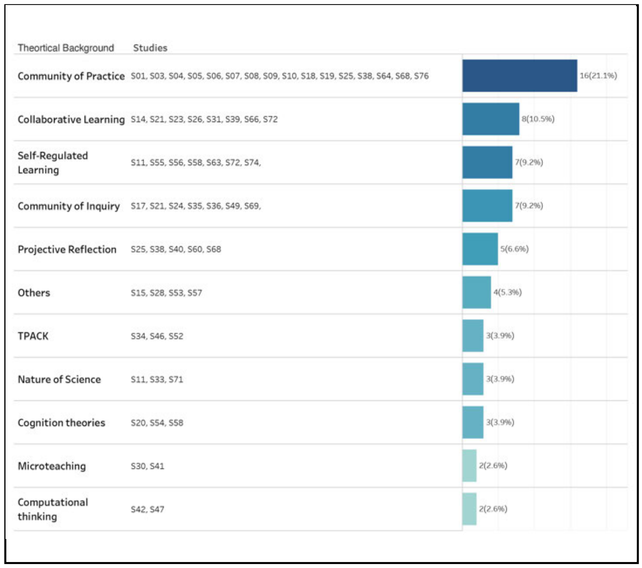

FIGURE 9. Theoretical background in the reviewed articles.

F. THEORETICAL BACKGROUND IN SELECTED ARTICLES

Most of the reviewed articles use ENA based on educational theories and frameworks (n=53, 69.7%), with one or more theories per article, whereas 23 articles (30.3%) did not report theoretical grounding (Figure 9).

1) COMMUNITIES OF PRACTICE (CoP)

CoP were the most commonly used theoretical framework. One of the oft-cited definitions by Lave and Wenger [28], describe CoP as "communities of people who share a common body of knowledge, a set of skills, a value system, and a set of decision-making processes". For any CoP, the elements of the epistemic frame theory, "the skills, knowledge, values, identity, and epistemology," are linked together and shared by CoP members in their professional context [29]. ENA was used to analyze networks of different CoP in 16 articles (21.1%), for example, urban planning [12], [20], [63], [70], [79], [84], [90], teaching [50], [62], [100], engineering [21], [61] and journalism [64].

2) COLLABORATIVE LEARNING (CL)

Johnson and Johnson [105] define collaborative learning as a "set of teaching and learning strategies promoting student collaboration in small groups in order to optimize their own and each other's learning". ENA was used to study students' collaboration with each other and their mentors to frame, examine, and solve complex problems in different settings, for example, collaborative problem-solving (CPS) [71], [73], [75] communities of inquiry [48], computer supportive collaborative learning (CSCL) [49], project-based engineering [19] and scientific reasoning processes [16].

3) SELF-REGULATED LEARNING THEORY (SRL)

An oft-cited definition by Panadero [106] describes SRL as "a core conceptual framework to understand the cognitive, metacognitive, behavior, motivational, and emotional aspects of learning.'' Learners effectively control their learning through internal and external feedback cycles of planning, performance, and reflection to control metacognitive and motivational behavior toward their goals in SRL. In the context of SRL, ENA was used in seven articles (9.2%) to analyze sequences of SRL processes [19], [93], identification of SRL in learning strategies [24], metacognition reflections [99], self-regulated behavior [86], feedback perception to adapt SRL processes [66] and finally, students' views in curriculum based on SRL [87].

4) COMMUNITY OF INQUIRY (CoI)

Garrison *et al.* [107] describe CoI as ''a pedagogical framework to examine the development of learning and cognition in online environments through three dimensional relationships called presences.'' ENA was used to analyze the relationship between one or more dimensions of CoI in seven articles (9.2%) [48], [55]–[59], [101].

5) PROJECTIVE REFLECTION (PR)

Foster *et al.* [79] define PR as ''a methodological and theoretical framework that considers learning as an exploration of identity.'' PR focuses on the integration between (**I**)dentity in a specific community of practice and self (**i**)dentity which reflects the personal goals. ENA explored identity in five articles (6.6%) through theoretical constructs of ''Knowledge, Interests/ Valuing, patterns of Self-organization/ Self-control, and Self-perceptions/Self-definitions'' [79], [84], [90]–[92].

6) TECHNOLOGICAL PEDAGOGICAL CONTENT KNOWLEDGE (TPACK)

Koehler & Mishra [108] define TPACK as ''a teaching knowledge base that integrates the teachers' technology practices with knowledge of content, general pedagogy, knowledge of learners and educational contexts.'' The connections of TPACK were investigated using ENA in three articles (3.9%) [26], [53], [104].

7) NATURE OF SCIENCE (NOS)

NOS is the main focus of science education and a key element of science literacy, expressing a way of knowing that integrates the features of scientific knowledge development [109]. The connections between NOS aspects were explored using ENA in two articles [81], [87] and views of the NOS surveys were used in a one article [85].

8) COGNITION THEORIES

Metacognition – or being aware of one's learning and improving the learning experience using one's own cognitive resources [110] – was used in two articles that focused on the metacognitive components of knowledge, goals, and actions [68], [99], whereas one article was based on ***distributed cognition theory***, which argues that cognition and knowledge are distributed among individuals and tools in the environment and are not confined to an individual [76].

9) COMPUTATIONAL THINKING (CT)

Wing [111] defines CT as ''a fundamental skill for solving problems, designing systems, and understanding human behavior that draws on concepts to computer science.'' The students' computational thinking was measured and analyzed using ENA in two articles [74], [82].

10) MICRO-TEACHING

Micro-teaching is an implementation method of dialogic pedagogical teaching. Student teachers are usually trained to play the roles of both teachers and students during practice micro-teaching. The teacher's role allows students to improve their teaching skills while the student's role helps them understand the psychology of the student [72], [103].

Other theoretical frameworks were reported in the reviewed studies, namely, knowledge building [83], the control-value theory of achievement emotion [66], and Language processing theory [47]. Finally, the achievement goal theory framework was reported [67].

G. THE MAIN RESEARCH FINDINGS IN THE ARTICLES

The most frequent theme in the research findings in the articles was related to learning interaction, which can be classified into learner-instructor, learner-contents, or learner-learner interactions [42]. ENA was used to evaluate***learner-instructor interaction***[63]. Although the teacher's communication with students might show initial resistance to the mentor's frame, it subsides with mentor facilitation to develop a more professional frame [60], [70]. When studying discourse, the authors reported that virtual mentoring is as effective as face-to-face mentoring [63]. The study of the participant and pedagogy reflections revealed that ''mentor-reflecting-on-student-action'' was related to ill-formed activity, whereas ''student-reflecting-on-mentor-action'' was related to more well-formed tasks [61].

1) LEARNER-CONTENTS INTERACTION

The learners' interactions with different discussion materials using ENA showed a more harmonious communication when students used interactive learning materials [18]. ENA analyzed the connections among the elements of professional skills and identity, for example, journalism [64], engineering [21], and urban planning [84]. Moreover, ENA was used to describe students' thinking in problem solving in a professional manner [64], the development of students' identity [91] and the stronger connections among the discourse elements of high performer students [78].

2) LEARNER-LEARNER INTERACTION (ANALYZING COLLABORATION)

ENA provided insight into CoI in online discussions for cognitive presence [55] and social relationships [57]. The temporal context of students' collaboration was explored in engineering design [61] and suggested that students with a more social exchange are more engaged in the planning and solving of collaborative tasks [19]. Similarly, ENA was used to investigate socio-cognitive activities and students' collaboration in problem solving [71]. In general, high performers had strong connections between adding concepts and creating links in the problem-solving process, which significantly differed from low performers' behavior [86]. Other analytical methods like process mining (PM) have been used with ENA to augment the insights [75], [88].

3) STUDENT EVALUATIONS

ENA was used to reveal the differences between high performers and low performers by modeling the connections between verbal codes [76], examining the connections between skills and decision-making [46], and evaluating the scientific reasoning [16]. The high-performing group create an actual open-ended learning experience and develop higher-order thinking whereas the low-performing group focus on the link between knowledge and learning activities [50]. Although ENA can function as an assessment tool for teachers to assess assignments and interpret their contents, Fougt *et al.* [25] reported that they were unable identify significant differences of different performance levels in assignments.

4) NATURE OF SCIENCE CONNECTIONS

ENA was used to explore learners' connections among NOS elements to identify the quality of students' understanding of NOS [81], [85]. Thus, ENA was shown to be useful for exploring pre-conceptions and post-conceptions by identifying the presence of ideas and the change of ideas [87].

V. DISCUSSION

ENA was conceptualized in 2009 to offer a quantitative method for studying coded discourse [3]. During the last three years, ENA witnessed increasing interest in the methods, a growing community, a yearly scientific event, and an expanding miscellany of applications that extended to other fields in addition to education. The ENA toolset has also expanded to include web applications and R packages, for example, rENA [34] and ncodeR [7]. The current pace of growth suggests that ENA adoption will result in more and diverse scientific output. Therefore, our systematic review sought to analyze the emerging field to highlight current achievements and future opportunities. Our study analyzed 76 empirical studies published between 2010 and 2021, and 2019 had the most papers. In the last two years, 30 papers were published.

Sixty-two studies (82%) in our review analyzed samples of less than 100 students, with a median of 32 students. Fourteen studies (18%) analyzed more than 100 students, and only two studies had more than 1000 students [23], [44]. Whereas this limited number of study participants could be explained – at least partially – by the need to code the data, it indicates that ENA has been a tool for research rather than practice with limited scalability. Recently, automatic coding using, for example, topic modeling [15], [23] has emerged, as well as ‘‘semi-automated coding'' where part of the data is coded manually (as a training dataset) and the remainder is automatically coded based on the training data [47]. Progress toward automatic coding would accelerate the uptake of ENA as a scalable tool that can be embedded in, for example, dashboards to help teachers or students.

Many of the studies have analyzed online discussions or interactions among learners, and they constitute almost half of the studies in our review. Interestingly, the other half used data that was not related to students' interactions or conversations (interviews, assessment tools, log files, and observations), which signals the expanding repertoire of the applications of ENA beyond discourse analysis and to new areas of inquiry. Theories related to collaborative learning (CoP, CL, CoI) constituted almost half of the included studies, and SRL – as a theoretical underpinning – came a distant second with around 10%.The remainder consisted of other theoretical frameworks (e.g., projective reflection, TPACK, and computational thinking) or no clear theoretical framework in 23 studies (30%). Such a picture reflects the interest in exploring the relational insights of ENA to model the relations and the structure of connections between the studied elements in new areas.

The prevalence of diverse types of data, inductive coding and the use of various theoretical frameworks point to many papers that tried ENA to solve existing problems, shed light on new aspects, or explore new areas of inquiry. Similarly, several studies have explored ENA as a complementary method to existing approaches. For instance, [23] explored the value of combining SNA and ENA in what they referred to as a social epistemic network signature (SENS) to predict students' performance. Similar combinations with SNA have also been explored to examine teachers' agency [94] and participation in knowledge building [80], [83]. Recently, ENA was used in tandem with process and sequence mining to augment the derived insights and reveal the strength and magnitude of connections between students' SRL elements [24], [88], [93].

Perhaps the most important objective of this study is how ENA has contributed to our research and practice. The largest number of the studies used ENA to evaluate, understand, or compare aspects related to students' interactions with each other, with content, or with teachers. Studies evaluating students and mentor interactions reported that ENA helped to assess mentoring and examine teacher facilitation and learners' reflection on teacher mentoring [12], [63]. Similarly, several researchers have reported the utility of ENA to evaluate the development of professional abilities in journalism, engineering, urban planning, script writing, and surgery, for example [21], [52], [64], [77], [84]

The largest group in the studies have assessed collaborative settings. The most common findings relate to revealing the types and strengths of presence in CoI [101], assessing patterns of knowledge exchange in collaborative learning [80], [83], or understanding problem solving behavior [71]. Less commonly, ENA was utilized to assess the nature of scientific thinking [81], computational thinking [82] and TPACK [26]. A small subset of the studies explored students' perception and sense-making of feedback in terms of value, impact, and quality [67]. Comparing and contrasting high and low achievers to infer differences and similarities has also been a common theme for evaluating differences in verbal discourse [76], scientific reasoning [16] and higher-order thinking [50]. However, studies assessing the utility of ENA as an assessment tool were unable to identify significant performance differences in assignments [25].

The studies in our review included a type of comparison, and the most common comparisons (23) addressed differences in performance (30%), 19 examined differences in temporal progress (25%), and 10 focused on differences between participant subgroups (13%). However, this comparison has been performed visually in many studies (43%). Studies comparing the differences among groups have mostly used pairwise comparison along the X-axis or Y-axis. A drawback of such comparison is the difficulty of communicating the results in an easy-to-understand way or translating differences along coordinates into interpretable conclusions. Furthermore, most of the existing comparisons have not compared across several groups (more than two), and they use pre-post design or report effect sizes to quantify the magnitude of such differences with effect size.

5) HAVE ENA FULFILLED ITS PROMISES?

We offer a concise overview based on the literature covered in this study and the seminal papers by the founders of the field to answer the question of how ENA has realized the aspirations behind the establishment of the field. One of the main promises of ENA was to offer quantitative analysis of coded data, which Shaffer *et al.* [10] stated as follows: "ENA identifies and measures connections among elements ...and measures the strength of association among elements in a network." However, the reviewed studies have barely reported the quantified structure of discourse or the strength of association between codes. Similarly, ENA was proposed as a method to "assess learner performance." However, the reviewed research has reported aggregated results of groups in most studies, with just one study in which the authors concluded that using ENA did not help assess students' performance [25]. ENA was built with temporality in mind. According to Shaffer *et al.* [10], "using ENA to create a trajectory model, which indicates changes in structures of connections over time." We have not seen such trajectories in the reviewed papers, and the papers that examined progression have compared different aggregated networks of certain periods – an approach that strips the network of its longitudinal aspect. Whereas many studies harness the co-temporal nature of interactions, these papers display the results in aggregated networks and, therefore, the aspect of temporality is lost.

Furthermore, ENA promised to enable "the analysis of networks too large for multivariate parametric techniques" [10]. However, the majority of the networks in our study were small with a limited number of nodes (median =8). Studies that used ENA with a larger number of nodes were difficult to interpret or visualize with several overlapping edges and nodes [81], [87]. Another promise behind ENA was that the "network graphs allow us to interpret the significance of the locations of the points in the ENA model" [10]. This promise was achieved to a degree in the reviewed papers, as visual comparisons were prevalent in all our studies and easily showed the networks and their connections. However, it was unclear how the reported results by the reviewed papers help explain such differences to the reader or practitioner. In other words, if a study found a significant difference on the Y-axis between group A more than group B, how can these results be communicated simply and clearly to practitioners? Moreover, how are such differences translated into practice?

The literature review has left us pondering the comparison between ENA and SNA and multivariate analysis which are methods that ENA was built to address some of their shortcomings [2], [10], [112]. SNA offers several advantages for quantifying the overall networks (e.g., density, reciprocity, and efficiency), node positions (e.g., centrality measures), and connection strength (e.g., edge weights and edge centralities) [113], [114]. Such quantitative analysis has been immensely useful across vast fields of research [115]. ENA methods have no comparable quantitative measures and researchers cannot, for instance, report which was the most central node that bridges others or the node that was close to other codes. SNA offers several null models that allow researchers to compare networks under study to random network models to determine whether their networks are statistically meaningful [116]. Similarly, SNA offers several inferential network methods and robust confirmatory tests that help understand why edges form and, thus, help researchers build or contribute to hypotheses [116]. SNA has a vast community with several threads of research, a large repertoire of methods, open software tools, and solid theoretical foundations. However, ENA is maintained by a small group of researchers concentrated in a single institution with limited contribution from the wider community to the theory, statistical foundations, or development. In comparison, the network psychometric field has emerged to analyze multivariate networks, and it has a large community that contributes to the tools, methods, and theory, offers several network confirmatory tests, for example, network bootstrapping, and it has vibrant discussions regarding the theoretical and mathematical foundations of the methods [117]. It is unclear whether ENA has addressed the said shortcomings of the two methods. ENA may have offered alternative methods for implementing or improving some existing functions currently offered by network analysis methods. However, these alternative functions came at the expense of losing access to a wealth of potentials offered by the network ecosystem shared across several fields, for example, network science and SNA. The lack of ENA networks' inter-operability with existing methods or an export file format makes integration with other tools impossible.

TABLE 1. Primary studies references.

ID	Authors (Date)	Title	Citation	Aim
S1	Bagley (2010) [12]	The epistemography of an urban and regional planning 912: appropriation in the face of resistance.	9	Uncover the learning process within a graduate urban planning practicum
S2	Hatfield & Shaffer (2010) [13]	The epistemography of journalism 335: complexity in developing journalistic expertise	10	Organize the development of important epistemic sub-frames in Journalism practicum and quantitative measure of linkage between participant structure and epistemic sub-frame Using ENA
S3	Nash & Shaffer (2011) [20]	Mentor modeling: the internalization of modeled professional thinking in an epistemic game	73	1. Did the players of Urban Science develop planning epistemic frames? 2. Did the players imitate the epistemic frame that the in-game mentors modeled? 3. Did the epistemic frames that players demonstrated during the game with the mentors persist when the mentors were not present?
S4	Svarovsky (2011) [21]	Exploring complex engineering learning over time with epistemic network analysis	28	1. Do middle school girls develop their understanding of engineering epistemic frame elements by playing Digital Zoo? 2. Is there specific participant structure that reflect or connect to the epistemic frame elements of values and epistemology?
S5	Bodin (2012) [14]	Mapping university students' epistemic framing of computational physics using network analysis	41	1. What are the students focusing on, in terms of knowledge and beliefs in problem-solving task? 2. What role does physics knowledge take in problem-solving situation? 3. Are the epistemic networks useful for describing epistemic framing?
S6	Nash & Shaffer (2013) [60]	Epistemic trajectories: mentoring in a game design practicum	40	1. Does the team ''follow'' the mentor's frames during meetings become more similar over time? 2. Where does the team follow the mentor? 3. What is the explanation for how the mentor and team's frames change in distance across their trajectories?
S7	Arastoopour & Shaffer (2015) [61]	Epistemography and professional CSCL environment design	4	1. What are the significant activities in the engineering co-op program? 2. What are the participant structures of reflection that are used in those activities? 3. How do the participant structures of reflection relate to the epistemologies of the engineering co-op?
S8	Bagley & Shaffer, (2015b) [63]	Stop talking and type: comparing virtual and face-to-face mentoring in an epistemic game	29	1. Were there differences in mentors' and students' reflection meeting discourse between the two conditions? 2. Were there differences in students' learning outcomes and level of engagement between the two conditions?
S9	Bagley & Shaffer (2015a) [70]	Learning in an urban and regional planning practicum: the view from educational ethnography	21	1. observe students learning to become planners through participation in a practicum. 2. explore the difference between the students' and teacher's epistemic frames.
S10	Hatfield (2015) [64]	The right kind of telling: an analysis of feedback and learning in a journalism epistemic game	12	1. Does ENA show similarities or differences between (a) mentor feedback from a professional practicum and (b) mentor feedback from a game? 2. Does ENA show that players in science.net learn to think more like professional journalists?
S11	Peters-Burton (2015) [87]	Outcomes of a self-regulated learning curriculum model: network analysis of middle school students' views of nature of science	19	1. What are the shifts in the connections among ideas about NOS for the student group before and after a yearlong course that used a curricular model based on self-regulated learning? 2. What are the influences of the yearlong curriculum based on self-regulated learning theory on the ways student ideas about NOS shifted before and after the course?
S12	Arastoopour et al., (2016) [46]	Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis	62	Explore virtual internships, online simulations of 21st-century engineering design practice, as one method for teaching engineering design thinking using ENA.
S13	Cai et al. (2017) [15]	Epistemic network analysis and topic modeling for chat data from collaborative learning environment	20	1. Explore collaborative interaction chat data using the combination of topic modeling and epistemic network analysis. 2. Predict learning performance by semantic network connections students make in chats.
S14	Csanadi et al. (2017) [16]	Collaborative and individual scientific reasoning of pre-service teachers: new insights through epistemic network analysis (ENA)	8	1. Do collaborative and individual reasoners exhibit different epistemic networks of scientific reasoning while solving a professional problem? 2. Do the epistemic networks we detect investigating differ from epistemic networks based on the same data set that has been randomly resorted (i.e. with the same frequency information)?

TABLE 1. ***(Continued.)*** **Primary studies references.**

S15	Siebert-Evenstone et al. (2017) [47]	In search of conversational grain size: modelling semantic structure using moving stanza windows	96	Does the moving stanza window method provide information about group discourse that the conversation method does not?
S16	Csanadi et al. (2018) [17]	When coding-and-counting is not enough: using epistemic network analysis (ENA) to analyze verbal data in CSCL research	70	1. Which technique provides the best explanation of group differences with respect to learners' engagement in different learning actions? 2. To what extent are the results from Q1 due to systematic temporal co-occurrences between learning actions?
S17	R. Ferreira et al. (2018) [55]	Towards combined network and text analytics of student discourse in online discussions	22	1. What is the effect of instructional scaffolding intervention on students' cognitive presence of course topics? 2. Can we use the proposed model to assess the effectiveness of role assignment intervention on the development of students' cognitive presence?
S18	Fougt et al. (2018) [25]	Epistemic network analysis of students' longer written assignments as formative/summative evaluation	10	1. How, and to what extent, can ENA predict the quality of longer student assignments? 2. What characterizes the epistemic network of low, middle, and high performers?
S19	Hu et al. (2018) [100]	What do teachers share within socialized knowledge communities: a case of Pinterest	57	1. What do teachers share within Pinterest? 2. How do they make sense of curated curricula?
S20	Sullivan et al. (2018) [76]	Using epistemic network analysis to identify targets for educational interventions in trauma team communication	32	Model and Compare connections among trauma team discourse elements a modeled by ENA would predict the quality of team performance in trauma simulation.
S21	Ruis, Siebert-Evenstone, et al., (2019) [48]	Finding common ground: a method for measuring recent temporal context in analyses of complex, collaborative thinking	12	Measure recent temporal context in analyses of complex, collaborative thinking
S22	Bauer et al. (2019) [97]	Using ENA to analyze pre-service teachers' diagnostic argumentations: a conceptual framework and initial applications	3	Analyze Pre-service Teachers' Diagnostic Argumentations.
S23	Bressler et al. (2019) [71]	Using epistemic network analysis to examine discourse and scientific practice during a collaborative game	13	1. How does scientific practice evolve during gameplay? 2. Which elements of collaborative discourse support the development of scientific practice?
S24	R. Ferreira & Gašević (2019) [56]	What is the effect of a dominant code in an epistemic network analysis?	10	1. What is the effect of using different dimensions of SVD on the final network when dealing with the problem of a dominant code? 2. Does the exclusion of a dominant code impact the information provided by the ENA?
S25	Foster et al., (2019) [79]	High school students' role-playing for identity exploration: findings from virtual city planning	11	What is the nature of high school students' identity exploration as a result of exploring role-specific possible selves of an environmental scientist and urban planner in a play-based course?
S26	Gašević et al., (2019) [23]	SENS: network analytics to combine social and cognitive perspectives of collaborative learning	110	1. Does an ENA analysis of the content of collaborative discourse predict the structure of social network ties? 2. Does an ENA analysis of individual students' discourse predict students' centrality in the social network? 3. Does an SNA analysis of high- and low-performing sub-communities of students predict differences in the content of students' discourse? 4. Does a combined SNA and ENA model predict student outcomes better than an SNA or ENA model alone
S27	Karumbaiah et al., (2019) [89]	Using epistemic networks with automated codes to understand why players quit levels in a learning game	7	Why do Players quit Levels in a Learning Game?
S28	Lim et al. (2019) [65]	Exploring students' sensemaking of learning analytics dashboards: does frame of reference make a difference?	25	1. How does the salience of graphics in LADs differ according to the frame of reference? 2: Is there a difference in response to LADs according to the frame of reference? 3. Do differences in self-regulation influence the outcomes of research question 2? 4. How does an individual's sensemaking of LADs differ according to the frame of reference?

TABLE 1. *(Continued.)* Primary studies references.

S29	Melzner et al. (2019) [88]	Using process mining (PM) and epistemic network analysis (ENA) for comparing processes of collaborative problem regulation	7	How do students regulate collaborative learning when faced with either motivational or comprehension-related problems using PM and ENA?
S30	Mochizuki et al. (2019) [72]	Effects of perspective-taking through tangible puppetry in microteaching and reflection on the role-play with 3d animation	0	Examine the effectiveness of 3D animation for reflection that the system generates to foster perspective-taking.
S31	Nachtigall & Sung, (2019) [73]	Students' collaboration patterns in a productive failure setting: an epistemic network analysis of contrasting cases	5	1. Compare students' collaborative problem-solving processes between PF groups that had generated a solution with a high quality (HQ) and a low quality (LQ) groups. 2. Compare the relationship between PF groups' solution quality and their learning outcome.
S32	Oshima et al. (2019) [80]	Collective knowledge advancement through shared epistemic agency: socio-semantic network analyses	5	1. how high-school students engage in collective knowledge advancement through shared epistemic agency during jigsaw instruction 2. how their collective knowledge advancement relates to learning outcomes
S33	Peters-Burton et al. (2019) [81]	Extending the utility of the views of nature of science assessment through epistemic network analysis	5	1. To what extent did the group understand selected NOS concepts pre- and post-course? 2. In what ways did the group make connections among NOS concepts? 3. How do the results for traditional VNOS analysis and ENA compare? 4. What are the benefits and drawbacks of administering the VNOS assessment with an ENA extension?
S34	Phillips et al. (2019) [104]	The influence of discipline on teachers' knowledge and decision making	1	What are the connections between teachers' knowledge and their decision-making as defined?
S35	Rolim, Ferreira, Kovanović, et al. (2019) [57]	Analysing social presence in online discussions through network and text analytics	11	1. What is the relationship between students' social presence and different course topics? 2. What is the effect of instructional scaffolding on students' social presence within different course topics?
S36	Rolim, Ferreira, Lins, et al. (2019) [58]	A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry	37	1. What are the association between the individual phases of cognitive presence and the indicators of social presence? 2. What are the effects of the instructional scaffolds aimed at promoting cognitive presence on the association between the phases of cognitive presence and indicators of the social presence? 3. How do the associations between the phases of cognitive presence and the indicators of social presence evolve over time, under the use of the different instructional scaffolds?
S37	Ruis, Rosser, et al. (2019) [77]	Multiple uses for procedural simulators in continuing medical education contexts	1	How do surgeons with different levels of experience use procedural simulators?
S38	Shah et al. (2019) [90]	Examining the impact of virtual city planning on high school students' identity exploration	0	What is the nature of high school students' identity exploration as a result of exploring the role-possible selves of an environmental scientist and urban planner in a play-based course?
S39	Sung et al. (2019) [49]	Reading for breadth, reading for depth: understanding the relationship between reading and complex thinking using epistemic network analysis	2	1. Do long readers and short readers show different patterns of interaction in collaborative discussions? 2. Do these differences (if any) reflect a difference in the depth to which long and short readers are engaging in the text?
S40	Talafian et al. (2019) [91]	Promoting and tracing high school students' identity change in an augmented virtual learning environment	1	To what extent did participants engage in an integrated identity exploration in Virtual City Planning as a result of exploring the roles of an environmental scientist and urban planning?
S41	Wakimoto et al., (2019) [103]	Student teachers' discourse during puppetry-based microteaching	3	How the discourse patterns within the EduceBoard puppetry-based microteaching differ from normal microteaching to identify the former's specific effects?
S42	B. Wu, Hu, Ruis, et al. (2019) [74]	Analysing computational thinking in collaborative programming: a quantitative ethnography approach	24	1. What CT patterns do novice programmers exhibit when they collaborate on the development of a software application? 2. Do novice programmers follow different trajectories of CT competence development based on the collaborative programming activities of their group?
S43	B. Wu, Hu, & Wang (2019)[51]	Scaffolding design thinking in online stem preservice teacher training	18	How do STEM preservice teachers develop their design thinking competence collaboratively under two scaffolding modes, i.e., static scaffolding and adaptive scaffolding?
S44	Yi et al. (2019) [52]	Exploring the development of reflection among pre-service teachers in online collaborative writing: an epistemic network analysis	0	1. What are the characteristics in teachers' reflection journals while engaging them in online collaborative script writing? 2. Is there any difference in the epistemic networks of teachers' reflection between different phases of the online activity phases? If yes, what are the different epistemic characteristics of teachers in the two phases?

TABLE 1. ***(Continued.)*** **Primary studies references.**

S45	Yue et al. (2019) [118]	Applying epistemic network analysis to explore the application of teaching assistant software in classroom learning	1	Reveal the differences and changes in the use of teaching assistant software among students of high-score and low-score learning level, to explore the relationship between the application of teaching assistant software and students' classroom learning.
S46	Zhang et al. (2019) [53]	Exploring primary school teachers' technological pedagogical content knowledge (TPACK) in online collaborative discourse: an epistemic network analysis	36	1. What are the categories, frequency distribution and time series characteristics of teachers' knowledge domains in online discourse? 2. What are the differences between the characteristics of teachers in the higher-score and lower-score groups? 3. What are the differences among the characteristics of teachers with different? ages? 4. What are the differences between the characteristics of teachers in the post and reply groups?
S47	Arastoopour et al. (2020) [82]	Modeling and measuring high school students' computational thinking practices in science	18	1. Do students demonstrate gains on a pre-post CT STEM assessment after participating in From Ecosystems to Speciation? 2. How do students' CT-STEM practices change over time when participating in From Ecosystems to Speciation as represented by ENA discourse networks? 3. Are students' pre and post scores associated with CT-STEM practices as represented by ENA discourse networks?
S48	D'Angelo et al. (2020) [78]	Evaluating how residents talk and what it means for surgical performance in the simulation lab	9	Explore a method for assessing intraoperative performance by modeling how surgeons integrate skills and knowledge through discourse.
S49	M. Ferreira et al. (2020) [59]	Towards automatic content analysis of social presence in transcripts of online discussions	17	1. To what extent can accurately text mining methods automatically code online discussion messages according to the categories of social presence? 2. Which features do best predict each category of social presence? 3. Do automatically coded messages preserve similar structural properties in the analysis of associations between the categories of social presence and discussion topics to the results of the analysis performed with manual manually coded messages?
S50	Gamage et al. (2020) [102]	Exploring MOOC user behaviors beyond platforms	20	Identify if there are behavioral similarities between extreme users and novice users.
S51	Hod et al (2020) [54]	Refining qualitative ethnographies using epistemic network analysis: a study of socioemotional learning dimensions in a humanistic knowledge building community	5	1. Test the relation between qualitative analysis and ENA. 2. Investigate what ENA can tell us about the way socioemotional dimensions of learning are expressed in learning communities in relation to stages of group development.
S52	Oner (2020) [26]	A virtual internship for developing technological pedagogical content knowledge	16	How did participants' TPACK, represented as epistemic network graphs, change character over the eight sessions of School of the Future?
S53	Oshima et al. (2020) [83]	Analysis of students' ideas and conceptual artifacts in knowledge-building discourse	5	1. How are learners' ideas improved through their collaborative discourse? 2. How do learners engage in collectively improving their ideas by using their conceptual artifacts?
S54	Pratt (2020) [68]	A mixed methods approach to exploring the relationship between beginning readers' dialog about their thinking and ability to self-correct oral reading	4	1. What is the relationship between beginning readers' verbalizations of their thought processes when self-monitoring oral reading and their ability to self-correct reading miscues? 2. What themes of response do beginning-readers who are more proficient in self-correcting oral reading give to explain their self-monitoring strategies? 3. What themes of response do beginning-readers who are less proficient in self-correcting oral reading give to explain their self-monitoring strategies?
S55	Saint et al. (2020) [93]	Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning	22	1. To what extent can we qualitatively and quantitively characterize students' sequences of SRL micro-level processes, using frequency learning behaviors from event measures, network analysis and Process mining? 2. To what extent can we articulate contrasting patterns of SRL behaviors across different student groups, based on assessment performance, by using frequency measures, network analysis, and process mining? 3. To what extent can we consolidate these analytical methods to provide a coherent temporal/sequential narrative on SRL, as enacted in a blended-learning environment?

TABLE 1. ***(Continued.)*** **Primary studies references.**

S56	Uzir et al. (2020) [24]	Analytics of time management and learning strategies for effective online learning in blended environments	21	1. To what extent can a combination of data analytic methods provide a holistic view to theoretically meaningful learning strategies composed of time management and learning tactics? 2. To what extent a combination of network and process analytics techniques, proposed in this study, can be used to explain the critical dimensions (i.e., time, ordering, frequency, and strength of connections in tactic use) of learning strategies extracted from trace data?
S57	Lim, Dawson, Gašević, Joksimović, Pardo, et al. (2020) [66]	Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: an exploratory study of four courses	10	1. What are students' perceptions of their personalized, learning analytics-based feedback? 2. What are students' affective responses to their personalized, learning analytics-based feedback? 3. What is the relationship between students' perceptions of personalized, learning analytics-based feedback and their affective responses? 4. How does the relationship identified in Q3 differ according to course context?
S58	L. Wu et al. (2020) [99]	Using epistemic network analysis and self-reported reflections to explore students' metacognition differences in collaborative learning	6	1. What is the distribution of the metacognition phenomena in students' self-reported reflections in collaborative learning? 2. What are the differences in students' metacognitive patterns in high score group versus low-score group? 3. What are the differences in students' metacognitive patterns across different disciplines?
S59	Brückner et al. (2020) [95]	Epistemic network analyses of economics students' graph understanding: an eye-tracking study	1	1. How do students who solved a task correctly (correct solvers) differ from students who solved a task incorrectly (incorrect solvers) in their proportion of fixations on different Area of interest (AOIs) of the task graphs? 2. How do correct and incorrect solvers differ in their transition patterns between different AOIs of the task graphs?
S60	Shah et al. (2021) [92]	Facilitating and interpreting high school students' identity exploration trajectories in stem	2	What is the nature of high school students' identity exploration trajectories as a result of participating in a 9-week augmenting virtual learning environment?
S61	Bauer et al. (2020) [98]	Diagnostic activities and diagnostic practices in medical education and teacher education: an interdisciplinary comparison	5	1. To what extent do learners' diagnostic activities differ between medical education and teacher education? 2. To what extent do learners' diagnostic practices differ between medical education and teacher education?
S62	Pratt & Coleman (2020) [69]	Using epistemic network analysis to visually map a metacognitive continuum of urban fourth graders' strategies for navigating multimodal science texts	0	1. What strategies do urban fourth-grade students report for how they navigate multimodal science texts? 2. How does students' self-report of strategy use vary by text topic (e.g., physical or life science)? 3. What are the affordances of ENA in visually mapping the self-reported strategies of urban fourth graders for reading multimodal science books?
S63	Lim, Dawson, Gašević, Joksimović, Fudge, et al. (2020) [67]	Students' sense-making of personalised feedback based on learning analytics	3	1. How do students describe their SRL adaptations in their sense-making of personalized feedback? 2. Does (and if so, how) students' sense-making of personalized feedback—that is, the association between their perceptions and self-described SRL adaptations— differ across contexts?
S64	Peng et al. (2020) [50]	Modeling stem learning design competence through discourse analysis	5	1. How do high and low-performing pre-service teachers develop learning design competence during STEM learning design meetings? 2. What relationship, if any, exists between pre-service teachers' STEM learning design competence and their lesson plans?
S65	Schnitzler et al. (2020) [96]	Connecting judgment process and accuracy of student teachers: differences in observation and student engagement cues to assess student characteristics	5	1. concerning judgment accuracy: How accurately can student teachers judge student profiles? How does student teachers' judgment accuracy differ across student profiles? Which student profiles do student teachers interchange predominantly? 2. concerning observation of behavioral activity: indicate high and low judgment accuracy in student teachers' eye movements with regard to the number of fixations and the average fixation duration? 3. concerning utilization of student cues as a cognitive activity: Which student cues do student teachers utilize to assess student cognitive and motivational-affective characteristics? What combinations of student cues do student teachers with high and low judgment accuracy use to assess student cognitive and motivational-affective characteristics?
S66	Uz-Bilgin et al. (2020) [75]	Exploring how role and background influence through analysis of spatial dialogue in collaborative problem-solving games	3	1. How do individuals communicate spatial information during a role playing cross-platform collaborative game? 2. How does an individuals' role and prior knowledge in biology impact their dialogue during gameplay?

TABLE 1. *(Continued.)* **Primary studies references.**

S67	Wang et al. (2020) [18]	How multimedia influence group interaction in stem education an epistemic network analysis for online synchronous collaborative learning	1	How Multimedia Influence Group Interaction in STEM Education?
S68	Barany et al. (2021) [84]	Connecting curricular design and student identity change: an epistemic network analysis	1	How did the design of Virtual City Planning Session 3 support student exploration of science career identities?
S69	Espino et al. (2021) [101]	Student emotions in the shift to online learning during the covid-19 pandemic	3	Examine the emotions exhibited by students around factors of online learning as they transitioned abruptly from in-person instruction
S70	Misiejuk et al. (2021) [44]	Using learning analytics to understand student perceptions of peer feedback	7	1. What is the relationship between student's perception of the usefulness of feedback, improvement suggestions, and comments on the feedback? 2. What is the relationship between rubric characteristics and student's perception of the usefulness of feedback?
S71	Mulvey et al. (2021) [85]	Making connections using individual epistemic network analysis to extend the value of nature of science assessment	0	1. How can we interpret changes in individual participants' understanding of NOS through traditional VNOS analysis? 2. How can we interpret changes in individual participants' understanding of NOS through traditional VNOS analysis? How might iENA extend those interpretations?
S72	Nguyen et al. (2021) [19]	"We're looking good": social exchange and regulation temporality in collaborative design	4	1. How do teams engage in project-based learning (PBL) in a collaborative engineering design course? 2. How do team regulation patterns relate to individuals' learning outcomes, namely final course grade and perceived collective efficacy?
S73	Pantić et al. (2021) [94]	Making sense of teacher agency for change with social and epistemic network analysis	5	1. What is the nature of change that teachers try to achieve or orient themselves toward? 2. How is teacher agency for change associated with inclusive pedagogy? 3. What kind of interactions characterize teachers who exercise agency for change?
S74	Paquette et al. (2021) [86]	Using epistemic networks to analyze self-regulated learning in an open-ended problem-solving environment	0	Investigate how students self-regulate their learning in the Betty's Brain environment, where they engage in three categories of open-ended problem-solving actions
S75	Scianna et al. (2021) [45]	Counting the game: visualizing changes in play by incorporating game events	1	How do player responses to in-game feedback change between first and second games played?
S76	Vega et al. (2021) [62]	Negotiating tensions: a study of pre-service English as foreign language teachers' sense of identity within their community of practice	2	How do EFL pre-service teachers make sense of dominant NS-NNS discourses and perceived legitimate practices when negotiating their emerging teacher identities?

VI. CONCLUSION

In summary, a large volume of research has explored the potential of ENA across a diverse range of problems. The current analysis supports the conclusion that ENA has helped visualize connections between coded elements, enabled comparisons across groups, and helped augment the insights of other methods, for example, process mining. However, ENA has not been able to fulfill many of is aspirations. The current implementations of ENA have neither been scalable nor automated. Our analysis has also shown that there is insufficient evidence that ENA has helped quantitatively investigate qualitative data nor has it helped assess learners' performance nor chart the temporal trajectory of interactions and therefore, it is fair to conclude that the expectations for ENA were set too high, but many fell short of promise. The small group of developers behind the development of the theory, software, and conceptualization of the field limits the role of the broader community to mere consumers of locked tools rather than partners who can contribute and drive the field forward. Whereas we expect more growth in applications in both volume and breadth, we hope for more involvement of the wider community in shaping the perspectives of the methods. Only then can we expect more diversity in implementations, richer quantitative capabilities, and novel perspectives.

APPENDIX

See Table 1.

REFERENCES

[1] D. W. Shaffer, "Models of situated action," in *Games, Learning, and Society*, C. Steinkuehler, K. Squire, S. Barab, Eds. Cambridge, U.K.: Cambridge Univ. Press, 2012, pp. 403–432.

[2] D. Bowman, Z. Swiecki, Z. Cai, Y. Wang, and B. Eagan, "The mathematical foundations of epistemic network analysis," in *Advances in Quantitative Ethnography*, vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 91–105.

[3] D. W. Shaffer, D. Hatfield, G. N. Svarovsky, P. Nash, A. Nulty, E. Bagley, K. Frank, A. A. Rupp, and R. Mislevy, "Epistemic network analysis: A prototype for 21st-century assessment of learning," *Int. J. Learn. Media*, vol. 1, no. 2, pp. 33–53, May 2009, doi: 10.1162/ijlm.2009.0013.

[4] R. Kaliisa, K. Misiejuk, G. A. Irgens, and M. Misfeldt, "Scoping the emerging field of quantitative ethnography: Opportunities, challenges and future directions," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 3–17.

[5] D. W. Shaffer, *Quantitative Ethnography*. Madison, WI, USA: Cathcart Press, 2017.

[6] B. Eagan, B. Rogers, P. Rebecca, and D. W. Shaffer, "Rho for inter rater reliability," GNU R-Package version (1.3.0.3), Free Softw. Found., Boston, MA, USA, 2021. [Online]. Available: https://cran.r-project.org/web/packages/rhoR

[7] D. Shaffer, F. Borden, A. Srinivasan, and J. Saucerman, "The nCoder: A technique for improving the utility of inter-rater reliability statistics," Univ. Wisconsin-Madison, Madison, WI, USA, Epistemic Games Group Work. Paper 2015-01.

[8] Z. Cai, A. Siebert-Evenstone, B. Eagan, D. W. Shaffer, X. Hu, and A. C. Graesser, "nCoder+: A semantic tool for improving recall of nCoder coding," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 41–54, doi: 10.1007/978-3-030-33232-7_4.

[9] S. Zörgö and G.-J. Y. Peters, "Epistemic network analysis for semi-structured interviews and other continuous narratives: Challenges and insights," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 267–277, doi: 10.1007/978-3-030-33232-7_23.

[10] 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, doi: 10.18608/jla.2016.33.3.

[11] I. Herman, G. Melançon, and M. S. Marshall, "Graph visualization and navigation in information visualization: A survey," *IEEE Trans. Vis. Comput. Graphics*, vol. 6, no. 1, pp. 24–43, Jan. 2000, doi: 10.1109/2945.841119.

[12] E. Bagley, "The epistemography of urban and regional planning 912: Appropriation in the face of resistance," in *Proc. 9th Int. Conf. Learn. Sci.*, 2010, pp. 81–88.

[13] D. Hatfield and D. W. Shaffer, "The epistemography of journalism 335: Complexity in developing journalistic expertise," in *Proc. 9th Int. Conf. Learn. Sci. (ICLS)*, 2010, pp. 628–635.

[14] M. Bodin, "Mapping university students' epistemic framing of computational physics using network analysis," *Phys. Rev. Special Topics-Phys. Educ. Res.*, vol. 8, no. 1, p. 10115, Apr. 2012, doi: 10.1103/PhysRevST-PER.8.010115.

[15] Z. Cai, B. Eagan, N. M. Dowell, J. W. Pennebaker, D. W. Shaffer, and A. C. Graesser, "Epistemic network analysis and topic modeling for chat data from collaborative learning environment," in *Proc. 10th Int. Conf. Educ. Data Mining*, Jan. 2017, pp. 104–111.

[16] A. Csanadi, B. Eagan, D. W. Shaffer, I. Kollar, and F. Fischer, "Collaborative and individual scientific reasoning of pre-service teachers: New insights through epistemic network analysis (ENA)," in *Proc. Comput. Collab. Learn. Conf.*, 2017, pp. 215–222.

[17] A. Csanadi, B. Eagan, I. Kollar, D. W. Shaffer, and F. Fischer, "When coding-and-counting is not enough: Using epistemic network analysis (ENA) to analyze verbal data in CSCL research," *Int. J. Computer-Supported Collaborative Learn.*, vol. 13, no. 4, pp. 419–438, Dec. 2018, doi: 10.1007/s11412-018-9292-z.

[18] H. Wang, C. Wang, and F. Wu, "How multimedia influence group interaction in stem education an epistemic network analysis for online synchronous collaborative learning," in *Proc. Int. Symp. Educ. Technol.*, 2020, pp. 303–306, doi: 10.1109/ISET49818.2020.00072.

[19] H. Nguyen, K. Y. Lim, L. L. Wu, C. Fischer, and M. Warschauer, "'Were looking good': Social exchange and regulation temporality in collaborative design," *Learn. Instruct.*, vol. 74, Aug. 2021, Art. no. 101443, doi: 10.1016/j.learninstruc.2021.101443.

[20] P. Nash and D. W. Shaffer, "Mentor modeling: The internalization of modeled professional thinking in an epistemic game," *J. Comput. Assist. Learn.*, vol. 27, no. 2, pp. 173–189, Apr. 2011, doi: 10.1111/j.1365-2729.2010.00385.x.

[21] G. N. Svarovsky, "Exploring complex engineering learning over time with epistemic network analysis," *J. Pre-College Eng. Educ. Res.*, vol. 1, no. 2, pp. 19–30, 2011, doi: 10.5703/1288284314638.

[22] G. Arastoopour, N. Chesler, and D. W. Shaffer, "A simulation-based approach for increasing women in engineering," in *Proc. Comput.-Supported Collaborative Learn. Conf. (CSCL)*, 2013, pp. 217–218.

[23] 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, Mar. 2019, doi: 10.1016/j.chb.2018.07.003.

[24] N. A. Uzir, J. Jovanovic, W. Matcha, L.-A. Lim, and A. Fudge, "Analytics of time management and learning strategies for effective online learning in blended environments," in *Proc. 10th Int. Conf. Learn. Analytics Knowl.*, Mar. 2020, pp. 392–401, doi: 10.1145/3375462.3375493.

[25] S. S. Fougt, A. Siebert-Evenstone, B. Eagan, S. Tabatabai, and M. Misfeldt, "Epistemic network analysis of students' longer written assignments as formative/summative evaluation," in *Proc. 8th Int. Conf. Learn. Anal. Knowl.*, Mar. 2018, pp. 126–130, doi: 10.1145/3170358.3170414.

[26] D. Oner, "A virtual internship for developing technological pedagogical content knowledge," *Australas. J. Educ. Technol.*, pp. 27–42, Jan. 2020, doi: 10.14742/ajet.5192.

[27] D. W. Shaffer, "Epistemic frames and islands of expertise: Learning from infusion experiences," in *Proc. 6th Int. Conf. Learn. Sci. (ICLS)*, 2004, pp. 473–480.

[28] J. Lave and E. Wenger, *Situated Learning*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

[29] D. W. Shaffer, "Epistemic frames for epistemic games," *Comput. Educ.*, vol. 46, no. 3, pp. 223–234, Apr. 2006, doi: 10.1016/j.compedu.2005.11.003.

[30] D. W. Shaffer, "Epistemic network analysis: Understanding learning by using big data for thick description," *Int. Handb. Learn. Sci.*, vol. 4, pp. 520–531, Oct. 2018, doi: 10.4324/9781315617572.

[31] D. W. Shaffer and A. R. Ruis, "How We Code," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 62–77.

[32] N. C. Chesler, A. R. Ruis, W. Collier, Z. Swiecki, G. Arastoopour, and D. W. Shaffer, "A novel paradigm for engineering education: Virtual internships with individualized mentoring and assessment of engineering thinking," *J. Biomech. Eng.*, vol. 137, no. 2, Feb. 2015, doi: 10.1115/1.4029235.

[33] J. L. Green and S. M. Bridges, "Interactional ethnography," *Int. Handb. Learn. Sci.*, vol. 4, pp. 475–488, Apr. 2018, doi: 10.4324/9781315617572.

[34] C. L. Marquart, Z. Swiecki, W. Collier, B. Eagan, R. Woodward, and D. W. Shaffer, "rENA: Epistemic network analysis," GNU R-Package Version (1.3.0.3), Free Softw. Found., Boston, MA, USA, 2021. [Online]. Available: https://cran.r-project.org/web/packages/rENA/

[35] M. Page, J. McKenzie, P. Bossuyt, and I. Boutron, "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," *Systematic Rev.*, vol. 10, no. 1, p. 89, Dec. 2021, doi: 10.1186/s13643-021-01626-4.

[36] C. Okoli, "A guide to conducting a standalone systematic literature review," *Commun. Assoc. Inf. Syst.*, vol. 37, no. 1, pp. 879–910, 2015, doi: 10.17705/1cais.03743.

[37] B. Kitchenham and S. Charters, *Guidelines for Performing Systematic Literature Reviews in Software Engineering*. Keele, U.K.: Keele Univ., 2007.

[38] A. Fink, *Conducting Research Literature Reviews: From the Internet to Paper*. Thousand Oaks, CA, USA: SAGE, 2010.

[39] J. Chen, M. Wang, P. A. Kirschner, and C.-C. Tsai, "The role of collaboration, computer use, learning environments, and supporting strategies in CSCL: A meta-analysis," *Rev. Educ. Res.*, vol. 88, no. 6, pp. 799–843, Dec. 2018, doi: 10.3102/0034654318791584.

[40] P. Shea and T. Bidjerano, "Learning presence: Towards a theory of self-efficacy, self-regulation, and the development of a communities of inquiry in online and blended learning environments," *Comput. Educ.*, vol. 55, no. 4, pp. 1721–1731, Dec. 2010, doi: 10.1016/j.compedu.2010.07.017.

[41] M. Skjott Linneberg and S. Korsgaard, "Coding qualitative data: A synthesis guiding the novice," *Qualitative Res. J.*, vol. 19, no. 3, pp. 259–270, Jul. 2019, doi: 10.1108/QRJ-12-2018-0012.

[42] M. G. Moore, "Editorial: Three types of interaction," *Amer. J. Distance Educ.*, vol. 3, no. 2, pp. 1–7, Jan. 1989, doi: 10.1080/08923648909526659.

[43] J. Webster and R. T. Watson, "Analyzing the past to prepare for the future: Writing a literature review," *MIS Quart.*, vol. 26, no. 2, pp. 23–30, 2002, doi: 10.1.1.104.6570.

[44] K. Misiejuk, B. Wasson, and K. Egelandsdal, "Using learning analytics to understand Student perceptions of peer feedback," *Comput. Hum. Behav.*, vol. 117, Apr. 2021, Art. no. 106658, doi: 10.1016/j.chb.2020.106658.

[45] J. Scianna, D. Gagnon, and B. Knowles, "Counting the game: Visualizing changes in play by incorporating game events," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 218–231.

[46] G. Arastoopour, D. W. Shaffer, Z. Swiecki, A. R. Ruis, and N. C. Chesler, "Teaching and assessing engineering design thinking with virtual internships and epistemic network analysis," *Int. J. Eng. Educ.*, vol. 32, no. 3, pp. 1492–1501, 2016.

[47] A. L. Siebert-Evenstone, G. Arastoopour Irgens, W. Collier, Z. Swiecki, A. R. Ruis, and D. Williamson Shaffer, "In search of conversational grain size: Modeling semantic structure using moving stanza windows," *J. Learn. Anal.*, vol. 4, no. 3, pp. 631–638, Dec. 2017, doi: 10.18608/jla.2017.43.7.

[48] A. R. Ruis, A. L. Siebert-Evenstone, R. Pozen, B. R. Eagan, and D. W. Shaffer, "Finding common ground: A method for measuring recent temporal context in analyses of complex, collaborative thinking," in *Proc. Comput.-Supported Collaborative Learn. Conf. (CSCL)*, 2019, pp. 136–143.

[49] H. Sung, C. Shengyang, A. R. Ruis, and D. W. Shaffer, "Reading for breadth, reading for depth: Understanding the relationship between reading and complex thinking using epistemic network analysis," in *Proc. 13th Comput.-Supported Collaborative Learn. Conf. (CSCL)*, 2019, pp. 376–383.

[50] X. Peng, B. Wu, and Y. Hu, "Modeling STEM learning design competence through discourse analysis," in *Proc. 20th Int. Conf. Adv. Learn. Technol.*, 2020, pp. 201–205, doi: 10.1109/ICALT49669.2020.00065.

[51] B. Wu, Y. Hu, and M. Wang, "Scaffolding design thinking in online STEM preservice teacher training," *Brit. J. Educ. Technol.*, vol. 50, no. 5, pp. 2271–2287, Sep. 2019, doi: 10.1111/bjet.12873.

[52] Y. Yi, X. Lu, and J. Leng, "Exploring the development of reflection among pre-service teachers in online collaborative writing: An epistemic network analysis," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 257–266.

[53] S. Zhang, Q. Liu, and Z. Cai, "Exploring primary school teachers' technological pedagogical content knowledge (TPACK) in online collaborative discourse: An epistemic network analysis," *Brit. J. Educ. Technol.*, vol. 50, no. 6, pp. 3437–3455, Nov. 2019, doi: 10.1111/bjet.12751.

[54] Y. Hod, S. Katz, and B. Eagan, "Refining qualitative ethnographies using epistemic network analysis: A study of socioemotional learning dimensions in a humanistic knowledge building community," *Comput. Educ.*, vol. 156, Oct. 2020, Art. no. 103943, doi: 10.1016/j.compedu.2020.103943.

[55] R. Ferreira and V. Rolim, "Towards combined network and text analytics of student discourse in online discussions," in *Artificial Intelligence in Education* (Lecture Notes in Computer Science), vol. 10947, H. U. Hoppe, R. Luckin, M. Mavrikis, K. Porayska-Pomsta, B. McLaren, and B. Boulay, Eds. Cham, Switzerland: Springer, 2018, pp. 111–126.

[56] R. Ferreira, "What is the effect of a dominant code in an epistemic network analysis," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 66–76.

[57] V. Rolim, R. Ferreira, V. Kovanovi, and D. Gasevic, "Analysing social presence in online discussions through network and text analytics," in *Proc. 19th Int. Conf. Adv. Learn. Technol. (ICALT 2019)*, Jul. 2019, pp. 163–167, doi: 10.1109/ICALT.2019.00058.

[58] V. Rolim, R. Ferreira, and R. D. Lins, "A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry," *Internet Higher Educ.*, vol. 42, pp. 53–65, Jul. 2019, doi: 10.1016/j.iheduc.2019.05.001.

[59] M. Ferreira, V. Rolim, R. Ferreira, R. D. Lins, G. Chen, and D. Gaáevic, "Towards automatic content analysis of social presence in transcripts of online discussions," in *Proc. PervasiveHealth Pervasive Comput. Technol. Healthc.*, Mar. 2020, pp. 141–150, doi: 10.1145/3375462.3375495.

[60] P. Nash and D. W. Shaffer, "Epistemic trajectories: Mentoring in a game design practicum," *Instructional Sci.*, vol. 41, no. 4, pp. 745–771, Jul. 2013, doi: 10.1007/s11251-012-9255-0.

[61] G. Arastoopour and D. W. Shaffer, "Epistemography and professional CSCL environment design," in *Proc. Comput.-Supported Collaborative Learn. Conf.*, 2015, vol. 1, pp. 204–211.

[62] H. Vega, G. A. Irgens, and C. Bailey, "Negotiating tensions: A study of pre-service English as foreign language teachers' sense of identity within their community of practice," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 277–291.

[63] E. A. Bagley and D. W. Shaffer, "Stop talking and type: Comparing virtual and face-to-face mentoring in an epistemic game," *J. Comput. Assist. Learn.*, vol. 31, no. 6, pp. 606–622, Dec. 2015, doi: 10.1111/jcal.12092.

[64] D. Hatfield, "The right kind of telling: An analysis of feedback and learning in a journalism epistemic game," *Int. J. Gaming Computer-Mediated Simulations*, vol. 7, no. 2, pp. 1–23, Apr. 2015, doi: 10.4018/IJGCMS.2015040101.

[65] L.-A. Lim, "Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference?" in *Proc. Pervasive-Health Pervasive Comput. Technol. Healthc.*, Mar. 2019, pp. 250–259, doi: 10.1145/3303772.3303804.

[66] L.-A. Lim, S. Dawson, and D. Gaševic, "Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: An exploratory study of four courses," *Assess. Eval. High. Educ.*, vol. 46, no. 3, pp. 1–21, Apr. 2020, doi: 10.1080/02602938.2020.1782831.

[67] L.-A. Lim, S. Dawson, A. Fudge, A. Pardo, and S. Gentili, "Students' sense-making of personalised feedback based on learning analytics," *Australas. J. Educ. Technol.*, vol. 36, no. 6, pp. 15–33, Dec. 2020, doi: 10.14742/ajet.6370.

[68] S. M. Pratt, "A mixed methods approach to exploring the relationship between beginning Readers' dialog about their thinking and ability to self-correct oral reading," *Reading Psychol.*, vol. 41, no. 1, pp. 1–43, Jan. 2020, doi: 10.1080/02702711.2019.1674434.

[69] S. M. Pratt and J. M. Coleman, "Using epistemic network analysis to visually map a metacognitive continuum of urban fourth Graders' strategies for navigating multimodal science texts," *Elementary School J.*, vol. 121, no. 2, pp. 224–255, Dec. 2020, doi: 10.1086/711296.

[70] E. Bagley and D. W. Shaffer, "Learning in an urban and regional planning practicum: The view from educational ethnography," *J. Interact. Learn. Res.*, vol. 26, no. 4, pp. 369–393, 2015.

[71] D. M. Bressler, A. M. Bodzin, B. Eagan, and S. Tabatabai, "Using epistemic network analysis to examine discourse and scientific practice during a collaborative game," *J. Sci. Educ. Technol.*, vol. 28, no. 5, pp. 553–566, Oct. 2019, doi: 10.1007/s10956-019-09786-8.

[72] T. Mochizuki, "Effects of perspective-taking through tangible puppetry in microteaching and reflection on the role-play with 3D animation," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 315–325.

[73] V. Nachtigall and H. Sung, "Students' collaboration patterns in a productive failure setting: An epistemic network analysis of contrasting cases," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 165–176.

[74] B. Wu, Y. Hu, A. R. Ruis, and M. Wang, "Analysing computational thinking in collaborative programming: A quantitative ethnography approach," *J. Comput. Assist. Learn.*, vol. 35, no. 3, pp. 421–434, Jun. 2019, doi: 10.1111/jcal.12348.

[75] C. Uz-Bilgin, M. Thompson, and M. Anteneh, "Exploring how role and background influence through analysis of spatial dialogue in collaborative problem-solving games," *J. Sci. Educ. Technol.*, vol. 29, no. 6, pp. 813–826, Dec. 2020, doi: 10.1007/s10956-020-09861-5.

[76] S. Sullivan, C. Warner-Hillard, B. Eagan, R. J. Thompson, A. R. Ruis, K. Haines, C. M. Pugh, D. W. Shaffer, and H. S. Jung, "Using epistemic network analysis to identify targets for educational interventions in trauma team communication," *Surgery*, vol. 163, no. 4, pp. 938–943, Apr. 2018, doi: 10.1016/j.surg.2017.11.009.

[77] A. R. Ruis, A. A. Rosser, J. N. Nathwani, M. V. Beems, S. A. Jung, and C. M. Pugh, "Multiple uses for procedural simulators in continuing medical education contexts," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 211–222.

[78] A.-L.-D. D'Angelo, A. R. Ruis, W. Collier, D. W. Shaffer, and C. M. Pugh, "Evaluating how residents talk and what it means for surgical performance in the simulation lab," *Amer. J. Surg.*, vol. 220, no. 1, pp. 37–43, Jul. 2020, doi: 10.1016/j.amjsurg.2020.02.016.

[79] A. Foster, M. Shah, A. Barany, and H. Talafian, "High school students' role-playing for identity exploration: Findings from virtual city planning," *Inf. Learn. Sci.*, vol. 120, nos. 9–10, pp. 640–662, Oct. 2019, doi: 10.1108/ILS-03-2019-0026.

[80] J. Oshima, R. Oshima, A. Ohsaki, and J. Michael Splichal, "Collective knowledge advancement through shared epistemic agency: Socio-semantic network analyses," in *Proc. 13th Comput.-Supported Collaborative Learn. Conf. (CSCL)*, 2019, pp. 57–64.

[81] E. E. Peters-Burton, J. C. Parrish, and B. K. Mulvey, "Extending the utility of the views of nature of science assessment through epistemic network analysis," *Sci. Educ.*, vol. 28, nos. 9–10, pp. 1027–1053, Dec. 2019, doi: 10.1007/s11191-019-00081-x.

[82] G. Arastoopour Irgens, S. Dabholkar, C. Bain, P. Woods, K. Hall, H. Swanson, M. Horn, and U. Wilensky, "Modeling and measuring high school Students' computational thinking practices in science," *J. Sci. Educ. Technol.*, vol. 29, no. 1, pp. 137–161, Feb. 2020, doi: 10.1007/s10956-020-09811-1.

[83] J. Oshima, R. Oshima, and S. Saruwatari, "Analysis of students' ideas and conceptual artifacts in knowledge-building discourse," *Brit. J. Educ. Technol.*, vol. 51, no. 4, pp. 1308–1321, Jul. 2020, doi: 10.1111/bjet.12961.

[84] A. Barany, M. Shah, and A. Foster, "Connecting curricular design and student identity change: An epistemic network analysis," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 155–169.

[85] B. K. Mulvey, J. C. Parrish, J. W. Reid, J. Papa, and E. E. Peters-Burton, "Making connections: Using individual epistemic network analysis to extend the value of nature of science assessment," *Sci. Educ.*, vol. 30, no. 3, pp. 527–555, Feb. 2021, doi: 10.1007/s11191-020-00189-5.

[86] L. Paquette, T. Grant, Y. Zhang, G. Biswas, and R. Baker, "Using epistemic networks to analyze self-regulated learning in an open-ended problem-solving environment," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 185–201.

[87] E. E. Peters-Burton, "Outcomes of a self-regulated learning curriculum model: Network analysis of middle school students' views of nature of science," *Sci. Educ.*, vol. 24, nos. 7–8, pp. 855–885, Oct. 2015, doi: 10.1007/s11191-015-9769-3.

[88] N. Melzner, M. Greisel, M. Dresel, and I. Kollar, "Using process mining (PM) and epistemic network analysis (ENA) for comparing processes of collaborative problem regulation," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 154–164.

[89] S. Karumbaiah, R. S. Baker, A. Barany, and V. Shute, "Using epistemic networks with automated codes to understand why players quit levels in a learning game," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 106–116.

[90] M. Shah, A. Foster, H. Talafian, and A. Barany, "Examining the impact of virtual city planning on high school students' identity exploration," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 199–210.

[91] H. Talafian, M. Shah, A. Barany, and A. Foster, "Promoting and tracing high school students' identity change in an augmented virtual learning environment," in *Proc. 13th Int. Conf. Comput. Supported Collaborative Learn. (CSCL)*, 2019, pp. 216–223.

[92] M. Shah, A. Foster, H. Talafian, A. Barany, and M. E. Petrovich, "Facilitating and interpreting high school students' identity exploration trajectories in STEM," *J. Experim. Educ.*, vol. 89, no. 3, pp. 541–559, Jul. 2021, doi: 10.1080/00220973.2020.1808941.

[93] J. Saint, W. Matcha, N. A. Uzir, and A. Pardo, "Combining analytic methods to unlock sequential and temporal patterns of self-regulated learning," in *Proc. PervasiveHealth Pervasive Comput. Technol. Healthc.*, Mar. 2020, pp. 402–411, doi: 10.1145/3375462.3375487.

[94] N. Pantic, S. Galey, L. Florian, and S. Joksimovic, "Making sense of teacher agency for change with social and epistemic network analysis," *J. Educ. Chang.*, vol. 17, pp. 1–33, Jan. 2021, doi: 10.1007/s10833-021-09413-7.

[95] J. Schneider, O. Zlatkin-Troitschanskaia, and H. Drachsler, "Epistemic network analyses of economics students' graph understanding: An eye-tracking study," *Sensors*, vol. 20, no. 23, pp. 1–32, Dec. 2020, doi: 10.3390/s20236908.

[96] K. Schnitzler, D. Holzberger, and T. Seidel, "Connecting judgment process and accuracy of student teachers: Differences in observation and student engagement cues to assess student characteristics," *Frontiers Educ.*, vol. 5, p. 259, Dec. 2020, doi: 10.3389/feduc.2020.602470.

[97] E. Bauer, M. Sailer, J. Kiesewetter, and C. Schulz, "Using ENA to analyze pre-service teachers' diagnostic argumentations: A conceptual framework and initial applications," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 14–25.

[98] E. Bauer, F. Fischer, J. Kiesewetter, and D. Shaffer, "Diagnostic activities and diagnostic practices in medical education and teacher education: An interdisciplinary comparison," *Frontiers Psychol.*, vol. 11, p. 2787, Oct. 2020, doi: 10.3389/fpsyg.2020.562665.

[99] L. Wu, Q. Liu, G. Mao, and S. Zhang, "Using epistemic network analysis and self-reported reflections to explore students' metacognition differences in collaborative learning," *Learn. Individual Differences*, vol. 82, Aug. 2020, Art. no. 101913, doi: 10.1016/j.lindif.2020.101913.

[100] S. Hu, K. T. Torphy, A. Opperman, K. Jansen, and Y.-J. Lo, "What do teachers share within socialized knowledge communities: A case of pinterest," *J. Prof. Capital Community*, vol. 3, no. 2, pp. 97–122, Apr. 2018, doi: 10.1108/JPCC-11-2017-0025.

[101] D. P. Espino, T. Wright, V. M. Brown, Z. Mbasu, M. Sweeney, and S. B. Lee, "Student emotions in the shift to online learning during the COVID-19 pandemic," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1312, A. R. Ruis and S. B. Lee, Eds. Cham, Switzerland: Springer, 2021, pp. 334–347.

[102] D. Gamage, I. Perera, and S. Fernando, "Exploring MOOC user behaviors beyond platforms," *Int. J. Emerg. Technol. Learn.*, vol. 15, no. 8, pp. 161–179, Apr. 2020, doi: 10.3991/IJET.V15I08.12493.

[103] T. Wakimoto, "Student teachers' discourse during puppetry-based microteaching," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 234–244.

[104] M. Phillips, V. Kovanovi, and I. Mitchell, "The influence of discipline on teachers' knowledge and decision making," in *Advances in Quantitative Ethnography* (Communications in Computer and Information Science), vol. 1112, B. Eagan, M. Misfeldt, and A. Siebert-Evenstone, Eds. Cham, Switzerland: Springer, 2019, pp. 177–188.

[105] D. W. Johnson and R. T. Johnson, "Making cooperative learning work," *Theory Into Pract.*, vol. 38, no. 2, pp. 67–73, Mar. 1999, doi: 10.1080/00405849909543834.

[106] E. Panadero, "A review of self-regulated learning: Six models and four directions for research," *Frontiers Psychol.*, vol. 8, p. 422, Apr. 2017, doi: 10.3389/fpsyg.2017.00422.

[107] D. R. Garrison, T. Anderson, and W. Archer, "The first decade of the community of inquiry framework: A retrospective," *Internet Higher Educ.*, vol. 13, nos. 1–2, pp. 5–9, Jan. 2010, doi: 10.1016/j.iheduc.2009.10.003.

[108] M. J. Koehler and P. Mishra, "What happens when teachers design educational technology? The development of technological pedagogical content knowledge," *J. Educ. Comput. Res.*, vol. 32, no. 2, pp. 131–152, Mar. 2005, doi: 10.2190/0EW7-01WB-BKHL-QDYV.

[109] N. G. Lederman, "Nature of science: Past, present, and future," in *Handbook of Research on Science Education*, S. K. Abell and N. G. Lederman, Eds. Mahwah, NJ, USA: Lawrence Erlbaum Associates, 2007, pp. 831–879.
[110] J. H. Flavell, "Metacognition and cognitive monitoring: A new area of cognitive-developmental inquiry," *Amer. Psychologist*, vol. 34, no. 10, pp. 906–911, 1979, doi: 10.1037/0003-066X.34.10.906.
[111] J. Wing, "Computational thinking," *Commun. ACM*, vol. 49, pp. 33–35, Mar. 2006, doi: 10.1145/1118178.1118215.
[112] D. W. Shaffer and A. R. Ruis, "Epistemic network analysis: A worked example of theory-based learning analytics," in *Handbook of Learning Analytics*, L. Charles, S. George, W. Alyssa, and D. Gašević, Eds., 1st ed. Society for Learning Analytics Research (SoLAR), 2017, pp. 175–187, doi: 10.18608/hla17.015.
[113] 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, Nov. 2017, doi: 10.1016/j.edurev.2017.08.005.
[114] 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, doi: 10.1007/s11412-020-09322-6.
[115] 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, doi: 10.1126/science.1165821.
[116] S. J. Cranmer, P. Leifeld, S. D. McClurg, and M. Rolfe, "Navigating the range of statistical tools for inferential network analysis," *Amer. J. Political Sci.*, vol. 61, no. 1, pp. 237–251, Jan. 2017, doi: 10.1111/ajps.12263.
[117] D. Borsboom, M. Deserno, and M. Rhemtulla, "Network analysis of multivariate data in psychological science," *Nat. Rev. Methods Prim.*, vol. 1, no. 1, pp. 1–18, 2021, doi: 10.1038/s43586-021-00055-w.

RAMY ELMOAZEN is currently pursuing the Ph.D. degree in learning analytics with the University of Eastern Finland. He is also a Researcher at the School of Computing, University of Eastern Finland. His research interests include developing computer-supported collaborative learning and learning analytics, including social networks and epistemic network analyses.

MOHAMMED SAQR received the Ph.D. degree in learning analytics from Stockholm University, Sweden. He had a postdoctoral training at the Université de Paris, France. He currently works as a Senior Researcher at the University of Eastern Finland, Finland. He is working on artificial intelligence, big data, network science, and scientometrics. His research interests include advancing novel methods, such as temporal networks, graphical Gaussian models, multichannel sequence analysis, and temporal processes in general. He is also an active member of several scientific organizations and acts as an academic editor in leading academic publications.

MATTI TEDRE is currently a Professor of computer science, especially computing education, and the philosophy of computer science, with the University of Eastern Finland. His recent book *Computational Thinking* (The MIT Press) with Peter J. Denning presented a rich picture of computing's disciplinary ways of thinking and practicing, and his 2014 book *Science of Computing* (Taylor & Francis/CRC Press) portrayed the conceptual and technical development of computing as a discipline. He has given more than a dozen conference keynotes on computing education, including ACM ITiCSE, ISSEP, and ECSS.

LAURA HIRSTO is currently a Professor of educational science with the Department of Applied Educational Science and Teacher Education, University of Eastern Finland. She is also supervising various research and development projects related to academic and educational development, student learning and teacher learning, and pedagogical perspectives on learning analytics in various contexts. Her research interests include higher education students' and teacher students' as well as primary level pupils' learning and motivational processes, and in variations of effective teaching and learning environments.