

Scoping the Emerging Field of Quantitative Ethnography: Opportunities, Challenges & Future Directions

Rogers Kaliisa¹, Kamila Misiejuk², Golnaz Arastoopour Irgens^{and} Morten Misfeldt⁴

¹Department of Education, University of Oslo, Oslo, Norway
{rogers.kaliisa@iped.uio.no}

²Centre for the Science of Learning & Technology, University of Bergen, Bergen, Norway
{Kamila.Misiejuk@uib.no}

³Department of Education & Human Development, Clemson University, South Carolina, USA
{garasto@clemson.edu}

⁴Department of Computer Science, University of Copenhagen, Copenhagen, Denmark
{misfeldt@ind.ku.dk}

Cite as: Kaliisa, R., Misiejuk, K., Irgens, G. A., & Misfeldt, M. (2021, February). Scoping the Emerging Field of Quantitative Ethnography: Opportunities, Challenges and Future Directions. In *International Conference on Quantitative Ethnography* (pp. 3-17). Springer, Cham. https://doi.org/10.1007/978-3-030-67788-6_

Abstract. Quantitative Ethnography (QE) is an emerging methodological approach that combines ethnographic and statistical tools to analyze both Big Data and smaller data to study human behavior and interactions. This paper presents a methodological scoping review of 60 studies employing QE approaches with an intention to characterize and establish where the boundaries of QE might and should be in order to establish the identity of the field. The key finding is that QE researchers have enough commonality in their approach to the analysis of human behavior with a strong focus on grounded analysis, the validity of codes and consistency between quantitative models and qualitative analysis. Nonetheless, in order to reach a larger audience, the QE community should attend to a number of conceptual and methodological issues (e.g. interpretability). We believe that the strength of work from individual researchers reported in this review and initiatives such as the recently established International Society for Quantitative Ethnography (ISQE) can present a powerful force to shape the identity of the QE community.

Keywords: Quantitative Ethnography, Epistemic Network Analysis

1 Introduction

The increasing use of technology in many areas of society and life has led to an increasing amount of Big Data about human behavior and interaction [1]. However, this volume of data is usually too large and strains the capabilities of human interpretation and the traditional social science research approaches. In response to this challenge, a number of theoretical and methodological approaches have been suggested. Quant ethnography (QE) [2], is one of these approaches, and is the focus of this paper. QE links the power of statistics with the power of in-depth, ethnographic approaches, and examines both “Big data” and smaller data sets to understand the breadth of human behavior. QE varies from other approaches to big data analysis with its focus on validity and the linkages and consistency between quantitative models and qualitative analysis, both of which are sound on their own, and both of which are attending to the same mechanisms at work in the same set of data. Simply put, unlike the traditional mixed-methods approach, QE brings two broad approaches to fair sampling together into a single solution of research techniques other than two separate but related analyses. Thus, providing researchers with a thicker and richer description of the data as it yields quantifiable information about the network and visualization of discourse over time for individuals and groups [4]. The result is a more unified

mixed-methods approach that uniquely links qualitative and quantitative approaches to data by treating them as two sides of the same coin and results in a valid and robust understanding of human behavior [5].

This paper reports the first scoping review of research that employs QE approaches and techniques to study human behavior and interaction. A scoping review is a process of summarizing a range of evidence in order to convey the breadth and depth of a field [6]. This approach has been reported as appropriate for reviewing educational research across a range of domains, particularly those ‘breaking new ground’ [7] as is the case with QE. In the following sections, we provide a brief description of the central conceptual and theoretical underpinnings of QE, research questions, review methodology, presentation and discussion of findings, and future directions of the QE approach.

2 Theoretical Foundations of QE

The QE methodology is grounded in several theoretical and epistemological assumptions. For example, QE is based on the premise that any culture of learning is characterized by a Discourse (with a big D) which [8] defines as a particular way of “talking, listening, writing, reading, acting, interacting, believing, valuing, and feeling” (p.25) within some community. In this regard, the notion of “Big ‘D’ Discourse” sets a larger context for the analysis of “discourse” (with a little “d”), which describes how people interact (e.g. the flow of language-in-use across time and the patterns and connections across this flow of language) [8]. Moreover, for researchers to characterize learning and behavior and make sense of Discourse, they derive codes (with a small c) and Codes (with a big C). Using this parallel terminology, a Code is the culturally relevant meaning of some action or an interpretation of something that happens, while a code is a warrant to justify the claims made based on a discourse (e.g. the things a group of people could say as evidence for that interpretation) [2]. However, in order to create Thick Descriptions, QE seeks to find connections between the different Codes into what [9] calls ‘Cultures’ which consist of symbols that interact to form a web of meanings. It is from this background that QE researchers are interested in modelling Codes in their data to see how they are systematically related to one another in a broader Discourse. In addition, QE researchers tend to make use of existing theories to explain the observed phenomena and to close the interpretive loop by making it possible to search for understanding in a corpus of data, selection of appropriate analytic methods and the interpretation of a model against the original data [10].

3 Tools & Techniques of QE

QE uses a range of conceptual and practical tools for sensibly analyzing big data and human behavior. One common tool used for the model specification is Epistemic Network Analysis (ENA), a network analysis technique for analyzing the structure of connections among coded data by quantifying and modeling the co-occurrence of codes. The co-occurrences are modeled as dynamic, weighted node-link networks. Such networks can be compared visually and statistically, allowing for comparisons between groups or samples. The tool also allows the researcher to view the original qualitative data to close the interpretative loop. ENA stems from the operationalization of epistemic frame theory, a learning theory that models learning as ways of thinking, acting, and being in the world of some community of practice [11]. This theory suggests that any community of practice has a culture and that culture has a grammar, a structure composed of skills, knowledge, identity, values and epistemology, which forms the epistemic frame of the community [12]. Using ENA, researchers analyze the development of learners’ epistemic frames by modeling connections in student discourse [13] and measuring the co-occurrence of concepts within the conversations, topics, or activities that take place during learning [11,12].

Beyond ENA, the QE community has continued to develop new tools that align with the epistemological assumptions of QE. For example, to support inter-rater reliability (IRR), [16] developed a statistic called Shaffer's rho to improve on the measurement of percent agreement and to reach valid theoretical saturation. Moreover, to aid QE researchers with the coding of large datasets, [17] developed nCoder, a tool that helps researchers to discover and code key concepts in text data with minimum human judgements and as a theoretical warrant that their codes were theoretically saturated. Moreover, to deal with the shortcomings of existing tools, QE researchers have developed additional tools. For example, to address the presence of false negatives due to codes that occur infrequently, [18] recently created nCoder+, which is an add-on to nCoder equipped with a semantic component that helps to solve low recall. [19], developed the Reproducible Open Coding Kit (ROCK) a tool that eases manual coding of QE data sets, while [20] suggested a multimodal matrix approach to support the modeling of QE data.

4 Purpose of this Scoping Review and Research Questions

To the best of our knowledge, there are no studies that have mapped the landscape of QE to characterize this emerging field and how it differs from traditional qualitative, quantitative, and mixed methods approaches. Thus, in this methodological review, we identified and mapped the available evidence and the current progress and trends in the emerging approach of QE. We argue that a scoping review of QE studies is needed (i) to characterize and conceptualize the existing body of work, and establish where the boundaries of QE might and should be in order to develop the identity of the field; and to (ii) identify the opportunities, challenges and future directions of the QE approach. Against this background, the research questions in this study are:

1. Which research areas (e.g. learning sciences) and settings (e.g. educational situations) have been investigated using QE approaches.
2. What are the methodological characteristics (e.g. data, tools, visualizations, coding, and analysis) of studies employing QE approaches?
3. How are studies employing QE theoretically grounded (e.g. what theoretical approaches are commonly used) and what role does theory play in QE studies?

5 Methodology

The methodological framework of undertaking a scoping review by [6] and used in previous scoping reviews [21] underpinned this review. The review followed the following five stages: (i) identifying the research questions (ii) identifying relevant studies; (iii) study selection; (iv) charting data; (v) collating, summarizing and reporting results. To identify relevant studies, we used the following digital sources: ProQuest; ERIC; Web of Science; Scopus; Google Scholar; Science Direct; and PubMed. In addition, we performed manual searching through reference lists of some QE papers, as well as conference proceedings (e.g. the International Conference on Learning Analytics & Knowledge [LAK], and the first International Conference on Quantitative Ethnography [ICQE]). The search string used was "Quantitative Ethnography" OR "Epistemic Network Analysis" OR "ncoder" OR "ncoder+". The initial search resulted in 1,231 studies that were filtered based on our inclusion criteria (e.g. study applies QE as defined by Shaffer, presents empirical findings and published in English between 2009 and April 2020). The screening was based on titles, abstracts and full-text skimming, and began on 1 February 2020 until 30 April 2020. The final dataset included 60 papers: 34 conference papers and 26 journal articles

(see appendix I, and others indicated by * in the reference list). Editorials, theoretical papers, and posters were excluded from the analysis but where applicable, these have been used in the background and discussion of the empirical studies.

The coding was performed through several stages. Initially, two coders took a grounded approach and reviewed 10 studies for training purposes and to gain familiarity with the literature. In this case, initial codes were formed based on the descriptions and contextual information provided in the papers. Next, each of the coders independently coded a further 10 papers and then discussed coding challenges to refine the coding scheme. As other QE researchers [22], we used social moderation where two raters coded all the papers and then discussed all the areas where ratings differed until an agreement was reached. Finally, the coders split the papers and proceeded with coding the full sample following the revised codes. We undertook a narrative analysis of the identified studies using individual papers as the unit of analysis. We tabulated the included studies to provide an overview of the different codes. The full coding details can be found at [<https://onedrive.live.com/view.aspx?resid=E3B64EB830502E3E!119&ihint=file%2cxlsx&authkey=!Ao6IIVyrsFLIcgI>].

6 Results

6.1 Descriptive Information of Included Studies

The 60 studies included in this review consist of conference papers (n=34) and journal articles (n= 26) published by 9 different conferences and 18 journals. The analysis also showed that the number of studies employing QE approaches has been growing steadily between 2009 and 2020 with a significant spike of studies in 2019 (n=28) (Fig 1). This is partly explained by the fact that the first international conference on QE was held in October 2019 [23], which resulted in an increase, in the number of QE studies. The steady increase in both conference and journal publications imply that the field is maturing and reaching a larger audience. The findings further revealed that QE studies are generally applying small sample sizes with most papers (n=44) having a sample size of less than 100. Only six studies had a sample size between 100-1000 participants, yet only one study had a sample size higher than 1000 participants did. Nine papers were not explicit with the actual sample size

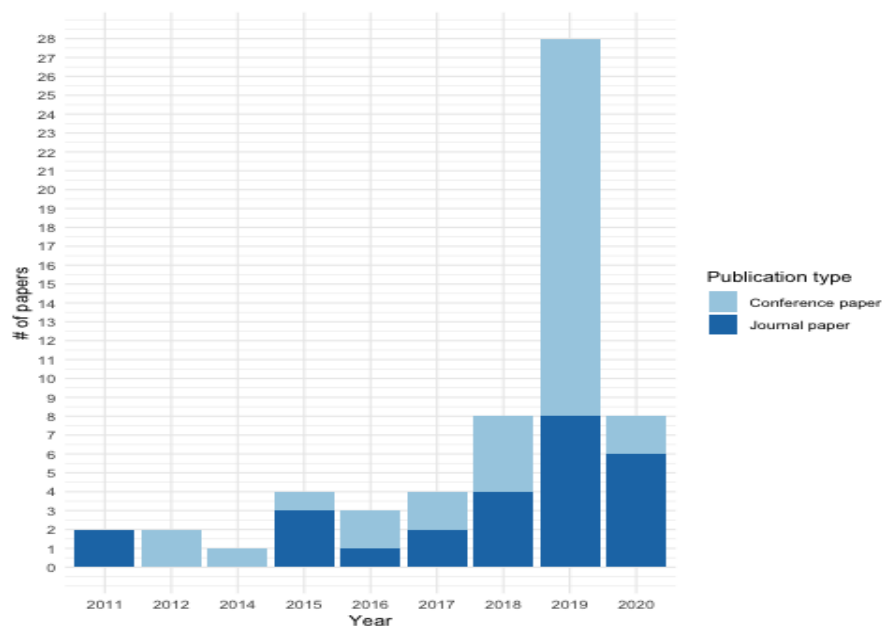


Fig. 1. The trend and type of QE publications

6.2 Discipline and Contextual Diversity in QE

The coding showed that the bulk of current QE studies are overwhelmingly within the learning sciences ($n=53$) and specifically in domains such as engineering education; science education; teacher education; collaborative learning; educational and epistemic games; language learning; and learning design. Beyond the learning sciences, six studies were from the health sciences in domains such as surgery [15], medicine [24] and care transitions [25]. The review found one study within the behavioral sciences [14], which examined adolescents' self-evaluative emotions and judgements towards aggression. Overall, studies employing QE draw on a diverse range of research fields, which could imply that QE caters to the sensitivity of different ontologies and worldviews.

The findings showed that formal learning settings ($n=31$) such as university ($n=21$) and K-12 ($n=10$), are a major site for QE research. These are closely followed by studies conducted in informal settings ($n=27$) such as work-place learning environments, online/virtual learning environments including simulations and games, and social media sites. Non-formal settings were only represented in one study [40]. These findings suggest that QE is broadening its reach across different settings beyond formal educational organizations and gaining attention in other contexts, such as informal learning settings.

6.3 Text and Non-Text Data Sources in QE Studies

The coding revealed that the most frequently utilized types of data were online discussion forums (n=19) mainly used to gain insights into students' cognitive presence; interviews (n=10) used to get a sense of the emic codes that people in a culture use to make sense of the world; and Third-party data (n=10) such as gaming environments/simulations used to evaluate students' learning products. The other forms of data used but less prominent include audio (n=6) and video (n=3); assessment (n=3) such as grades, and pre-post test scores; institutional data (n=2) such as lesson plans from LMSs; observation (n=2); as well as reflection journals, literature review, social media interaction data, eye tracking, students essays and survey each used in one study respectively. Overall, QE studies utilize a number of sources of data to study human behavior. Nonetheless, it is surprising that observation, which is one of the key ethnographic data collection approaches, was used by only two QE studies. This area might require further consideration to strengthen the study of human behavior using QE. Moreover, it was noted that most of the data used in QE studies are originally text-based (discussion forums), yet other non-text data such as videos are always converted to text forms for QE analysis.

6.4 Common Coding Approaches

Coding of the data is an essential part of the QE process since it "should accurately reflect something that is meaningful in the discourse" [26]. The articles in this review were coded by the following aspects of the coding process: 1) Coding type: approaches to data coding (top-down, bottom-up, hybrid); 2) Data coding: techniques used for data coding (automated, manual, mixed); 3) Raters: number of human and non-human raters coding and/or validating the data; 4) Coding validation: methods used to validate the coding scheme.

The most popular coding type method is top-down coding, that is, using predefined codes to code the data, used by 25 articles. Eighteen articles used bottom-up coding, where codes emerge from the data. Eleven articles used hybrid coding that combines both top-down and bottom-up coding methods. Six articles did not specify their coding type.

In 30 articles, a human or humans coded the data manually. Three articles used qualitative data analysis software, MAXQDA, for manual coding. Eighteen articles coded their data automatically: using Latent Dirichlet Allocation (LDA) (n=2), agent identification system (n=1), BeGaze software that annotates gaze targets (n=2), nCoder (n=5), or an automated coding algorithm similar to nCoder (n=7). [27] used topic modeling, word frequencies and n-grams to develop the codes used for the automated data coding with nCoder, while [28] used a mixed-method (manual+automated) approach. First, two human coders manually coded for cognitive presence, and then, LDA was applied to detect topics in the discussion messages. Eleven articles did not specify the data coding method.

Twenty-six articles coded and/or validated their codes with the help of two human raters. Another popular method was two human raters + machine adopted by 10 articles, while five articles relied only on automated methods. Two papers coded and/or validated their data by four raters. [29] used nine raters, whereas [30] used only one rater.

The most popular statistics used to validate the codes was Cohen's kappa reported in 30 articles. Shaffer's rho was reported in nine articles. [31] used Krippendorff's alpha. [32] and [33] mentioned using interrater reliability without specifying the exact statistics applied. Two papers used social moderation to validate their codes, while one paper used verbal agreement. [34] reported accuracy as their validation measure. Twenty-three papers did not specify their coding validation method. Overall, 30

articles specified all the aspects (coding type, data coding, raters, and validation methods), which constitutes only 50% of all papers included in this literature review. Two articles did not report any aspects of their coding process.

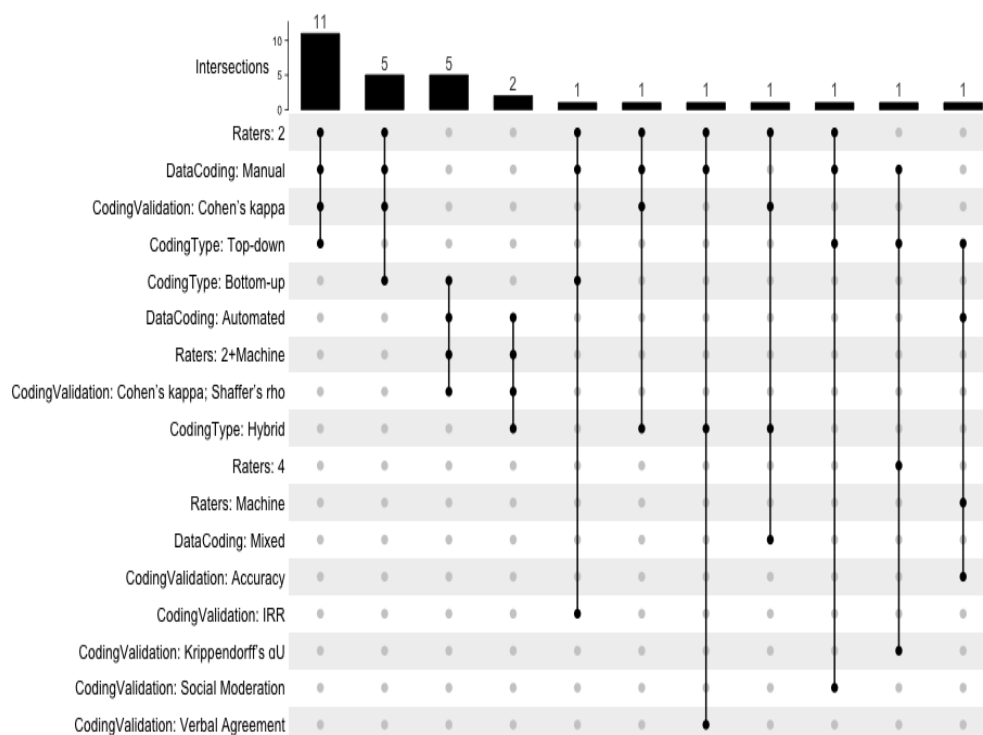


Fig. 2. Coding patterns.

The most popular coding pattern is manual top-down coding validated by two raters using Cohen's kappa ($n=11$), and a variation of this pattern, where the coding is bottom-up ($n=5$). Another popular pattern is automated bottom-up coding, where the codes are validated by two raters and a machine, and reporting Cohen's kappa and Shaffer's rho as validation statistics ($n=5$). Two papers used hybrid automated coding validated by two raters and a machine using Cohen's kappa and Shaffer's rho. The main finding is that only half of the papers reported on all steps of the coding and the coding validation process. Moreover, we discovered two main patterns of data coding: 1) two raters would manually code the data using either bottom-up or top-down method and validate the coding with Cohen's Kappa; 2) data would be automatically coded using either bottom-up or top-down method, and validated by two raters and the machine by applying Shaffer's rho and Cohen's Kappa.

6.5 ENA as a Dominant Analysis Method in QE Studies

All articles included in this literature review used ENA as one of their data analysis methods, and 31 articles used ENA as their only data analysis method. To complement the ENA findings, 12 articles used descriptive statistics, whereas eight articles reported both descriptive and inferential statistics. [35]

used ENA in combination with cluster analysis and descriptive statistics. [36] applied Principal Component Analysis and descriptive statistics in addition to ENA. ENA and time series were used by both [37] in combination with descriptive statistics, and [38], who visualized the time series with a CORDTRA diagram. Some articles compared the insights gained by using ENA to other data analysis methods. [39] compared ENA and process mining, and complemented the analysis with descriptive statistics. [29] used both ENA and time series (FRIEZE representation) to analyze temporal patterns.

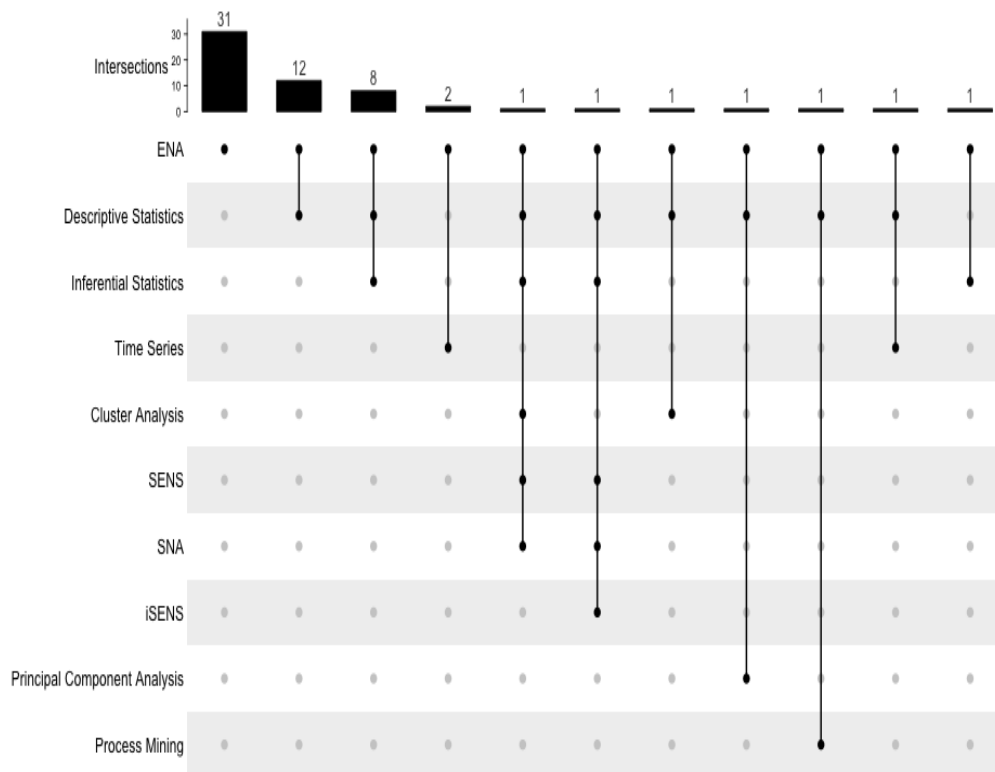


Fig. 3. Data analysis methods.

[40] did innovative work trying to expand ENA with social network analysis (SNA) to develop social and epistemic network signature (SENS). They also applied descriptive statistics, inferential statistics, and cluster analysis in their study. Building upon their findings, [41] compared ENA, SNA, SENS and social-epistemic network signature (iSENS) by using inferential statistics – hierarchical linear models – and descriptive statistics, such as effect size (Cohen’s d). In summary, most papers use only ENA as their data analysis method. Other papers complemented ENA with other methods, compared ENA to other methods, or tried to expand ENA by using other methods.

6.6 The Dynamics of Theory Integration in QE Studies

Theory plays a crucial role in QE by supporting the selection of relevant variables, development and interpretation of models, and converting those interpretations to meaningful and scientifically justified actions. The coding in this study revealed 44 studies that either used a theoretical framework ($n=36$) or/and some disciplinary concepts ($n=11$) to frame the research and in some cases support the coding, and interpretation of the results. In addition, the analysis revealed 16 studies, which were atheoretical.

That is, the studies were not explicitly linked to any specific theoretical framework meaning that they took a data-driven approach. Of those studies that used theoretical frameworks, they mainly grounded their research in the epistemic frame theory (n=14) which was used as an approach to conceptualize and characterize the structure of connections that students make among elements of authentic practice [42]. The next most used theoretical frameworks were the collaborative learning framework (n=10); Engineering design thinking (n=4); the community of inquiry framework (n=3) and projective reflection (n=4).

The analysis further revealed that the role played by the theoretical frameworks varied. Some studies used the frameworks as a basis to position their work and as a lens to interpret their findings but did not use them for coding. For example, [33] used epistemic frames to explore the relationship between beginning readers' Dialog about their thinking and ability to self-correct oral reading. However, the author took a more grounded approach to the segmentation of the data without necessarily using epistemic frame elements as [43] did. On the other hand, some QE studies used theoretical frameworks in an inductive way by using theoretical insights as a basis to support the filtering of data to develop relevant codes and later for closing the interpretative loop. For example, [44] used the community of inquiry framework as an analytical lens to understand the association between the individual phases of cognitive and social presences in communities of inquiry and at the same time, used the same framework to code the presence or absence of each dimension and to make sense out of the resulting models. Moreover, it was noticed that in some QE studies the theoretical frameworks are neither strictly deductive nor inductive, but represent a combination of both. In other words, theory was used as an analytical lens and for the development of codes but in combination with data-driven approaches. For example, [27] used a computational thinking-STEM taxonomy as a guiding framework but at the same time employed thematic analysis to identify student-constructed computational thinking practices that fit under the broader taxonomy categories. In sum, the findings suggest that QE is not aligned to a single theoretical perspective, but rather underpinned by different theoretical outlooks. In addition to employing a diverse catalog of perspectives, QE researchers also use theoretical frameworks to inform various stages of the analytical process.

7 Discussion

This methodological review was set out to map the available evidence and trends in the emerging approach of QE with an intention to characterize and establish the boundaries of QE in order to shape the identity of the field. First, in relation to RQ1 on research fields and settings of QE implementation, the findings indicated that QE researchers are overwhelmingly associated with the learning sciences. The findings also noted that QE is slowly being adopted in other fields such as health and behavioral sciences. Consequently, the majority of QE studies are implemented within formal educational settings but with a steady increase in other contexts such as informal and non-formal settings. The diversity of research fields and settings imply that QE is broadening its reach and caters to the sensitivity of different ontologies, worldviews, and contexts, something that provides hope for the future of this emerging field.

Second, in response to RQ2, which identified the methodological characteristics of QE studies, findings revealed that QE studies utilize a number of data sources to study human behavior with discussion forums, interviews and Third-party data from virtual learning environments being the most common. However, only two QE studies used observation as a source of data for QE analysis, despite this being one of the key ethnographic data collection approaches and a key premise of QE. More importantly, even though QE provides tools for analysis of Big data, currently few QE studies have analyzed Big data from platforms such as Twitter with most studies relying on data from very small sample sizes which could lead to challenges of drawing conclusions and integrating statistical approaches on too little ethnographic data as also noted by [2]. This implies that any conclusions made from QE studies

employing small samples are limited to that of the sample population, which could affect the application of QE at scale, yet an indication of maturity of research and applied methods relates to the scale of the studies in research and practice [45]. Moreover, it was noted that most of the data used in QE studies are originally text-based (discussion forums), yet other non-text data such as videos are always converted to text forms for QE analysis. This raises a question as to whether QE data always need to be in a text form for analysis to take place. Therefore, given the increasing use of non-text data in QE studies, the QE community should develop tools specific to non-text data to avoid the possibility of describing such data as *what it is not* rather than *what it is*.

In addition, only half of the papers in this review reported on all steps of the coding and the coding validation process. This is despite the emphasis that QE places on data coding and validation [46]. Of those papers that did report on the coding process, two main patterns of data coding emerged: 1) a traditional approach in which two raters manually coding the data and validating the coding with Cohen's Kappa; 2) as recommended in the QE foundational book [2], using automated classifiers to code the data that would be validated against two raters and the machine by applying Shaffer's rho and Cohen's Kappa. Though the bottom-up approach to coding is central to grounded analysis that plays a significant role in QE, articles in our literature review apply different coding approaches depending on the context.

Moreover, even though QE is positioned as a field encompassing ENA, we found few mentions of QE in the papers with more emphasis being on ENA. For example, only two papers described QE as a field [47, 34]. Five papers mentioned QE while describing the ENA method, whereas other 53 papers do not mention QE at all, though they usually cite [2]. This suggests a conflation between QE as a broader methodological approach and ENA as a tool used within QE. Perhaps, this is one area where the QE community should be more explicit as it strives to build an identity.

Lastly, in response to RQ3 on the theoretical grounding of QE studies, the findings revealed that the majority of QE studies integrated theory in their research practice. However, even though some QE papers highlight theoretical frameworks, they were not used to inform the analysis and interpretation of results. Moreover, 16 studies were entirely atheoretical which can lead to different, but equally troubling analytical weaknesses in QE since researchers may lack critical distance from the data and merely report findings based on preconceived notions and categories of discourse [10]. This may result in making macro-level claims that are not well supported by data and identifying emic concepts from the bottom-up without necessarily defining a set of categories in advance. In sum, since the definition of QE emphasizes the use of statistical techniques with the interpretive power of qualitative and grounded analysis to warrant claims about the quality of thick description, a hybrid approach as that taken by [27] who concurrently used a theoretical framework and thematic analysis to support the framing, coding and interpretation of data could be a better approach for QE researchers to promote coherence and ensure validity.

8 Implications and Conclusions

Overall, it is possible to conclude from this study that most of the QE researchers study human behavior in line with Shaffer's definition of QE. There is an observed effort among QE researchers to ensure the validity of their codes through IRR, model development and closing the interpretive loop. The review identified that most studies have a strong focus on validity and the linkages and consistency between quantitative models and qualitative analysis without necessarily treating them as separate. This is a promising trend towards the development of an identity for the QE community.

Nonetheless, our premise in this paper is that, as a new field trying to develop an identity, QE researchers need to attend to a number of conceptual and methodological issues. First, the over dominance of ENA in QE studies at the expense of QE could limit the clear understanding and development of the QE community if overshadowed by one of its tools of analysis. It is, therefore, necessary that QE

researchers put the principles of QE at the center of their research before introducing the specific tools of analysis. In addition, the papers reviewed for this paper suggests that the QE field has yet to define the standards, which would indicate the identity of the field. For example, even though the key principle of QE is to ensure ‘validity’, some QE studies still pay little attention to providing explicit details regarding the coding and validation processes. We argue that this is one of the areas where QE researchers should place more attention. Moreover, as a way to extend the QE methodology to a larger audience, it is important for the QE community to attend to technical aspects such as attendability and interpretability since some of the tools and outputs from QE analysis currently require sophisticated mathematical reasoning and prior experience and training in both qualitative and statistical analysis [48]. One possible approach is to develop simple tools with an interface as those suggested by [49] and [50] that support the sharing of QE outputs with the broader public to encourage uptake of QE in other fields of practice that are inherently less quantitative. We believe that the strength of work from individual researchers reported in this review and initiatives such as the recently established International Society for Quantitative Ethnography (ISQE) can present a powerful force to shape the identity of the QE community.

References

1. Brown, B., Chui, M., Manyika, J.: Are you ready for the era of ‘big data’. *McKinsey Quarterly*, **4**(1), 24-35 (2011)
2. Shaffer, D.W.: *Quantitative ethnography*. Cathcart Press, Madison, WI (2017)
3. *Wu, B., Hu, Y., Ruis, A., Wang, M.: Analysing computational thinking in collaborative programming: A quantitative ethnography approach. *J. Comput. Assist. Learn.* **35**(3), 421-434 (2019)
4. Shaffer, D.W., Collier, W., Ruis, A.: A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. *J. Learn. Anal.* **3**(3), 9-45 (2016)
5. Misfeldt, M., Spikol, D., Bruun, J., Saqr, M., Kaliisa, R., Ruis, A., Eagan, B.: Quantitative ethnography as a framework for network analysis—a discussion of the foundations for network approaches to learning analysis. In: *LAK’10 Companion Proceedings* (2020)
6. Levac, D., Colquhoun, H., O’Brien, K.K.: Scoping studies: advancing the methodology. *Implement. Sci.* **5**(1) (2010)
7. Munn, Z., Peters, M.D., Stern, C., Tufanaru, C., McArthur, A., Aromataris, E.: Systematic review or scoping review? Guidance for authors when choosing between a systematic or scoping review approach. *BMC Med. Res. Methodol.* **18**(1) (2018)
8. Gee, J.P.: Discourse, small d, big D. In: K. Tracy, T. Sandel, C. Ilie (eds.) *The International Encyclopedia of Language and Social Interaction*, pp.1-5 (2015)
9. Geertz, C.: Deep play: notes on the Balinese cockfight. In: Crothers L., Lockhart C. (eds.) *The Interpretation of Cultures. Selected Essays*, pp. 412-453. Palgrave Macmillan, New York (1973)
10. Wise, A.F., Shaffer, D.W.: Why theory matters more than ever in the age of big data. *J. Learn. Anal.* **2**(2), 5-13 (2015)
11. Shaffer, D.W., Ruis, A.R.: Epistemic network analysis: A worked example of theory-based learning analytics. In: C. Lang, G. Siemens, A.F. Wise, D. Gašević (eds.), *Handbook of learning analytics*, pp.175-187. Society for Learning Analytics Research (2017)
12. Rupp, A.A., Choi, Y., Gushta, M., Mislevy, R. J., Bagley, E., Nash, P., Hatfield, D., Svarowski, G., Shaffer, D.W.: Modeling learning progressions in epistemic games with epistemic network analysis: Principles for data analysis and generation. In: *LeaPS 2009 Proceedings* (2009)
13. Arastoopour Irgens, G., Shaffer, D.W.: Measuring social identity development in epistemic games. In: *CSCL 2013 Proceedings*, pp.42-48 (2013)

14. *Frey, K.S., Kwak-Tanquay, S., Nguyen, H.A., Onyewuenyi, A.C., Strong, Z.H., Waller, I.A.: Adolescents' Views of Third-Party Vengeful and Reparative Actions. In: ICQE 2019 Proceedings, pp.89-105 (2019)
15. *D'Angelo, A.L.D., Ruis, A.R., Collier, W., Shaffer, D.W., Pugh, C.M.: Evaluating how residents talk and what it means for surgical performance in the simulation lab. *Am. J. Surg.* **220**(1), 37-43 (2020)
16. Eagan, B.R., Rogers, B., Pozen, R., Marquart, C., Shaffer, D.W.: rhoR: Rho for inter-rater reliability (Version 1.1.0) (2016)
17. Shaffer, D.W., Borden, F., Srinivasan, A., Saucerman, J., Arastoopour, G., Collier, W., Ruis, A.R., Frank, K.A.: The nCoder: A Technique for Improving the Utility of Inter-Rater Reliability Statistics. Epistemic Games Group Working Paper 2015-01. (2015)
18. *Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D.W., Hu, X., Graesser, A.C.: nCoder+: A semantic tool for improving recall of nCoder coding. In: ICQE 2019 Proceedings, pp.41-54 (2019)
19. Zörgő, S., Peters, G.J.Y.: Epistemic network analysis for semi-structured interviews and other continuous narratives: Challenges and insights. In: ICQE 2019 Proceedings, pp.267-277 (2019)
20. Buckingham Shum, S., Echeverria, V., Martinez-Maldonado, R.: The Multimodal Matrix as a quantitative ethnography methodology. In: ICQE 2019 Proceedings, pp.26-40 (2019)
21. Major, L., Warwick, P., Rasmussen, I., Ludvigsen, S., Cook, V.: Classroom dialogue and digital technologies: A scoping review. *Educ. Inf. Technol.* **23**(5), 1995-2028 (2018)
22. *Espino, D., Lee, S., Eagan, B., Hamilton, E.: An initial look at the developing culture of online global meet-ups in establishing a collaborative, STEM media-making community. In: CSCL 2019 Proceedings, pp.608-611 (2019)
23. Eagan, B., Misfeldt, M., Siebert-Evenstone, A. (eds.): *Advances in Quantitative Ethnography: First International Conference, ICQE 2019, Madison, WI, USA, October 20–22, 2019, Proceedings* (Vol. 1112). Springer Nature (2019)
24. *Sullivan, S., Warner-Hillard, C., Eagan, B., Thompson, R.J., Ruis, A.R., Haines, K., Pugh, C.M., Shaffer, D.W., Jung, H.S.: Using epistemic network analysis to identify targets for educational interventions in trauma team communication. *Surgery*, **163**(4), 938-943 (2018)
25. *Wooldridge, A.R., Haefli, R.: Using epistemic network analysis to explore outcomes of care transitions. In: ICQE 2019 Proceedings, pp.245-256 (2019)
26. Shaffer, D.W.: Big data for thick description of deep learning. In: Millis, K., Long, D., Magliano, J., Wiemer, K. (eds.) *Deep Comprehension*, pp.265-277. Routledge, New York (2018)
27. *Arastoopour Irgens, G., Dabholkar, S., Bain, C., Woods, P., Hall, K., Swanson, H., Horn, M., Wilensky, U.: Modeling and measuring high school students' computational thinking practices in science. *J. Sci. Educ. Tech.* **29**(1), 137-161 (2020)
28. *Ferreira, R., Kovanović, V., Gašević, D., Rolim, V.: Towards combined network and text analytics of student discourse in online discussions. In: AIED 2018 Proceedings, pp.111-126 (2018)
29. *Lund, K., Quignard, M., Shaffer, D.W.: Gaining insight by transforming between temporal representations of human interaction. *J. Learn. Anal.* **4**(3), 102-122 (2017)
30. *Nash, P., Shaffer, D.W.: Mentor modeling: The internalization of modeled professional thinking in an epistemic game. *J. Comput. Assist. Learn.* **27**(2), 173-189 (2011)
31. *Bauer, E., Sailer, M., Kiesewetter, J., Schulz, C., Pfeiffer, J., Gurevych, I., Fischer, M.R., Fischer, F.: Using ENA to analyze pre-service teachers' diagnostic argumentations: A conceptual framework and initial applications. In: ICQE 2019 Proceedings, pp.14-25 (2019)
32. *Hu, S., Torphy, K.T., Chen, Z., Eagan, B.: How do US teachers align instructional resources to the common core state standards: A case of Pinterest. In: SMSociety 2018 Proceedings, pp.315-319 (2018)

33. *Pratt, S.M.: A mixed methods approach to exploring the relationship between beginning readers' dialog about their thinking and ability to self-correct oral reading. *Read. Psychol.* **41**(1), 1-43 (2020)
34. *Karumbaiah, S., Baker, R. S., Barany, A., Shute, V.: Using epistemic networks with automated codes to understand why Players quit levels in a learning game. In: *ICQE 2019 Proceedings*, pp. 106-116 (2019)
35. Peters-Burton, E.E.: Outcomes of a self-regulated learning curriculum model. *Sci. Educ.* **24**(7-8), 855-885 (2015)
36. *Swiecki, Z., Ruis, A.R., Farrell, C., Shaffer, D.W.: Assessing individual contributions to collaborative problem solving: a network analysis approach. *Comput. Hum. Behav.* **104** (2020)
37. *Zhang, S., Liu, Q., Cai, Z.: Exploring primary school teachers' technological pedagogical content knowledge (TPACK) in online collaborative discourse: An epistemic network analysis. *Br. J. Educ. Technol.* **50**(6), 3437-3455 (2019)
38. *Siebert-Evenstone, A., Arastoopour Irgens, G., Collier, W., Swiecki, Z., Ruis, A.R., Shaffer, D.W.: In search of conversational grain size: Modelling semantic structure using moving stanza windows. *J. Learn. Anal.* **4**(3), 123-139 (2017)
39. *Melzner, N., Greisel, M., Dresel, M., Kollar, I.: Using Process Mining (PM) and Epistemic Network Analysis (ENA) for comparing processes of collaborative problem regulation. In: *ICQE 2019 Proceedings*, pp.154-164 (2019)
40. *Gašević, D., Joksimović, S., Eagan, B.R., Shaffer, D.W.: SENS: Network analytics to combine social and cognitive perspectives of collaborative learning. *Comput. Hum. Behav.* **92**, 562-577 (2019)
41. *Swiecki, Z., Shaffer, D.W.: iSENS: an integrated approach to combining epistemic and social network analyses. In: *LAK'10 Proceedings*, pp.305-313 (2020)
42. *Bagley, E., Shaffer, D.W.: Epistemic mentoring in virtual and face-to-face environments. In: *ICLS 2012 Proceedings*, pp.256-260 (2012)
43. *Svarovsky, G.N.: Exploring complex engineering learning over time with epistemic network analysis. *J-PEER* **1**(2) (2011)
44. *Rolim, V., Ferreira, R., Lins, R.D., Gašević, D.: A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry. *Internet High. Educ.* **42**, 53-65 (2019)
45. Ognjanović, I., Gašević, D., Dawson, S.: Using institutional data to predict student course selections in higher education. *Internet High. Educ.* **29**, 49-62 (2016)
46. Shaffer, D.W.: QE-COVID data challenge. Why QE? [White paper]. (2020) <https://sites.google.com/wisc.edu/qe-covid-data-challenge/why-qe>
47. *Wu, B., Hu, Y., Ruis, A., Wang, M.: Analysing computational thinking in collaborative programming: A quantitative ethnography approach. *J. Comput. Assist. Learn.* **35**(3), 421-434 (2019)
48. *Swiecki, Z., Shaffer, D.W.: Toward a taxonomy of team performance visualization tools. In: *ICLS 2018 Proceedings*, pp.144-151 (2018)
49. Swiecki, Z., Marquart, C., Sachar, A., Hinojosa, C., Ruis, A.R., Shaffer, D.W.: Designing an Interface for sharing quantitative ethnographic research data. In: *ICQE 2019 Proceedings*, pp.334-341 (2019)
50. Herder, T., Swiecki, Z., Fougat, S. S., Tamborg, A.L., Allsopp, B.B., Shaffer, D.W., Misfeldt, M.: Supporting teachers' intervention in students' virtual collaboration using a network based model. In *LAK'8 Proceedings*, pp.21-25 (2018)

Appendix i

List of studies included in the scoping review but not cited in the main text (n=36)

Authors	Authors
Andrist, Collier, Gleicher, Mutlu, Shaffer (2015)	Orrill, Shaffer (2012)
Andrist, Ruis, Shaffer (2018)	Peters-Burton (2015)
Arastoopour Irgens, Shaffer, Swiecki, Ruis, Chesler (2015)	Peters-Burton, Parrish, Mulvey (2019)
Arastoopour Irgens, Shaffer (2015)	Phillips, Kovanović, Mitchell, Gašević (2019)
Barany, Foster (2019)	Ruis, Rosser, Nathwani, Beams, Jung, Pugh (2019)
Bressler, Bodzin, Eagan, Tabatabai (2019)	Ruis, Rosser, Quandt-Walle, Nathwani, Shaffer, Pugh (2018)
Brown, Nagar, Orrill, Weiland, Burke (2016)	Ruis, Siebert-Evenstone, Pozen, Eagan, Shaffer (2019)
Cai, Eagan, Dowell, Pennebaker, Shaffer, Graesser (2017)	Shah, Foster, Talafian, Barany (2019)
Chesler, Ruis, Collier, Swiecki, Arastoopour Irgens, Shaffer (2015)	Siebert-Evenstone, Shaffer (2019)
Csanadi, Eagan, Shaffer, Kollar, Fischer (2017)	Sinclair, Ferreira, Gašević, Lucas, Lopez (2019)
Espino, Lee, Van Tress, Baker, Hamilton (2020)	Sung, Cao, Ruis, Shaffer (2019)
Fisher, Hirshfield, Siebert-Evenstone, Arastoopour Irgens, Koretsky (2016)	Talafian, Shah, Barany, Foster (2019)
Fougt, Siebert-Evenstone, Eagan, Tabatabai, Misfeldt (2018)	Wakimoto, Sasaki, Hirayama, Mochizuki, Eagan, Yuki, Kato (2019)
Foster, Shah, Barany, Talafian (2019)	Whitelock-Wainwright, Tsai, Lyons, Kaliff, Bryant, Ryan, Gašević (2020)
Knight, Arastoopour Irgens, Shaffer, Buckingham Shum, Littleton (2014)	Wooldridge, Carayon, Shaffer, Eagan (2018)
Lim, Dawson, Joksimović, Gašević (2019)	Yi, Lu, Leng (2019)
Nachtigall, Sung (2019)	Yue, Hu, Xiao (2019)
Oner (2020)	