Combining Automatic Coding and Instructor Input to Generate ENA Visualizations for Asynchronous Online Discussion

Marcia Moraes^{*1[0000-0002-9652-3011]}, Sadaf Ghaffari^{*1[0000-0002-2413-5829]}, Yanye Luther^{1[0000-0002-0995-3356]}, and James Folkestad^{2[0000-0003-0301-8364]}

Department of Computer Science¹, School of Education², Colorado State University, Fort Collins CO 80523, USA

Abstract. Asynchronous online discussions are a common fundamental tools to facilitate social interaction in hybrid and online courses. However, instructors lack the tools to accomplish the overwhelming task of evaluating asynchronous online discussion activities. In this paper we present an approach that uses Latent Dirichlet Analysis (LDA) and the instructor's keywords to automatically extract codes from a relatively small dataset. We use the generated codes to build an Epistemic Network Analysis (ENA) model and compare this model with a previous ENA model built by human coders. The results show that there is no statistical difference between the two models. We present an analysis of these models and discuss the potential use of ENA as a visualization to help instructors evaluating asynchronous online discussions.

Keywords: ENA Visualization · Automated Coding · Unsupervised Learning · Instructor's Keywords

1 Introduction

Asynchronous online discussions are a common fundamental tool to facilitate social interaction in hybrid and online courses. They have been shown to improve students' critical thinking [10], knowledge construction [13], writing skills [2], and learning outcomes [24]. However, instructors lack the tools to accomplish the overwhelming task of evaluating asynchronous online discussion activities [14]. According to de Lima et al. [14], instructors reported struggling to assess students' contributions in forum activities due to difficulties in following the discussions, the lack of specific reports related to the subjects discussed, the students' contributions to those subjects, and the lack of visualizations to convey messages in a graphical format.

In order to address some of those difficulties, Epistemic Network Analysis (ENA) has been presented as learning analytics visualization tool to show the

^{*} These authors contributed equally to this work.

relationships among the different concepts students discuss in an asynchronous online discussion [16]. In this particular work [16], two human annotators manually coded the text data. In a more recent work [19], three different text mining approaches, namely Latent Semantic Analysis (LSA), Latent Dirichlet Analysis (LDA) and Clustering Word Vectors, were applied to automate code extraction from a relatively small discussion board dataset obtained from [16]. Based on the study presented in [16], we submitted a project proposal to our university's teaching innovation grant to investigate the application of ENA as a visualization tool to support instructors. Reviewers mentioned that instructors would like to have the visualizations but would not have time to be involved in the coding process, even if that process used nCoder [15]. However, the instructors would be willing to provide keywords that should be present in the codes. Based on the provided feedback, and unlike previous works, we not only looked into extracting codes from relatively small data using LDA but also illustrated how this can lead to generating ENA visualization without using nCoder [15] throughout the process.

This paper builds on those previous works by:

- 1. Using the automatically generated codes provided in [19] to code the same dataset used in [16].
- 2. Evaluating the quality of the automated coding with the human coding presented on [16] by interrater reliability.
- 3. Improving the automated code generation in order to reach a satisfactory interrater reliability between algorithm and human if necessary.
- 4. Applying the codes that were automatically generated in the creation of ENA visualizations.
- 5. Analyzing and evaluating those ENA visualizations with human coded ENA visualizations.
- 6. Validating both ENA visualizations with instructors.

In this paper we aim to answer the following research questions:

- RQ1. What is/are the difference(s) between the ENA visualizations generated using an automated coding process and a human coding process for the data presented in [16]?
- RQ2. How does an instructor evaluate the ENA visualizations generated?

The remainder of this paper is structured as follows. Section 2 summarizes prior research efforts on using ENA as a visualizations tool and on automated coding processes. Sections 3 and 4 describe the approach used in the study. Section 5 is dedicated to presenting the results obtained from our experiments. We discuss the results and outline directions for further work in Section 6.

2 Related Works

ENA has been used to support many facets of education and learning in several areas [4, 18, 21]. One use of ENA that is gaining more attention is its use as a

visualization tool [22] to help instructors evaluate clinical team simulations [9], support teachers' interventions in students' virtual collaboration [12], evaluate teamwork [23], include participants in co-construction and co-interpretation of ENA representations [25], and unveil the conceptual connections that students are making in asynchronous online discussions [16]. All of these studies utilized different coding strategies to code the data used in the ENA visualizations.

A recent work by Cai et al. 6 centered around nCoder where they investigated how close human created code words were to codes identified by topic models in two large datsets. Another work compared the performance of neural networks in a supervised learning manner with nCoder to assess which approach required the least human coding effort while achieving a sufficient and accurate classification [8]. In their comparison, they indicated nCoder had a higher accuracy. Although nCoder is a popular learning analytics platform used to develop coding schemes, it is not fully automated. In other words, it requires active human efforts to read through every item in the data and validate if the choice of coding is conceptually valid. It also suffers the low recall problem. Previous literature studies have identified ways to improve the problem of low recall or high false negative rate in nCoder. nCoder+ [7] aims to improve low recall issue in nCoder through semantic component addition. However, as mentioned by the authors of nCoder+, the idea is still a prototype and is not a public tool yet. In another research effort, the use of Negative Reversion Set (NRS) sampling has been shown to improve the low recall for Regular Expressions based classifiers such as nCoder [5].

Our work has several distinctive features compared to prior works. First, we only take advantage of a Natural Language Processing (NLP) unsupervised learning technique to automate the code extraction, despite having a relatively small dataset without the use of nCoder. Second, we utilize coherence analysis [17] to identify the optimal number of topics in the discussion data, thus avoiding arbitrary selection of the number. Third, we use instructor keywords in addition to the LDA extracted keywords to generate the visualization.

3 Method

The main goal of our approach was to extract topic keywords from a relatively small online discussion dataset using Latent Dirichlet Allocation (LDA) [3], use those keywords to automatically code asynchronous online discussion data, and generate ENA visualizations based on that data. In this section, we describe how we automated this process.

3.1 Dataset Preprocessing

The data utilized to investigate the research questions comprised of online discussion posts from seven semesters: Fall 2017, Fall 2018, Fall 2019, Spring 2020, Fall 2020, Spring 2021, and Fall 2021. The data consisted of 2,648 postings collected from an online class for organizational leaders as part of a Masters of

Education program at a Research 1 land-grant university. Table 1 represents prior codes in our dataset.

The problem of interest was based on code retrieval. This highlighted the importance of the preprocessing step in our setup. The preprocessing steps in our automatic extraction task consisted of tokenization, lowercasing, named-entity removal, stop words removal, in-document frequency filtering, and generating bigrams and trigrams since our interest was to retrieve the code containing two or three words.

Code Name	Definition	$\begin{array}{c} \mathbf{Kappa} \\ (\kappa) \end{array}$
Retrieval prac- tice, Spacing out practice, Interleaving	Retrieval practice is the act of recalling facts or concepts or events from memory and are also known as testing effect or retrieval-practice ef- fect. Spacing out practice allows people to a lit- tle forgetting that helps their process of con- solidation. Interleaving the practice of two or more concepts or skills help develop the ability to discriminate later between different kinds of problems and select the better solution.	0.85
Illusion of mas- tery	Researches have pointed out that students usu- ally have a misunderstanding about how learn- ing occurs and engage with learning strategies that are not beneficial for their long-term re- tention, such as rereading the material several times and cramming before exams.	0.89
Effortful learning	Learning is deeper and more durable when it is effortful, meaning that efforts, short-terms im- pediments (desirable difficulties), learning from mistakes, and trying to solve some problem before knowing the correct answer makes for stronger learning.	0.85
Get beyond learning styles	Researchers found that when instructional style matches the nature of the content, all learners learn better, regardless of their learning styles.	0.86

Table 1: Priori codes

3.2 Latent Dirichlet Analysis

We aimed to determine which codes are associated with each discussion post, i.e. in each document, and extract them. To accomplish this we used LDA [3], a generative probabilistic model, to extract the codes from the online discussion data to help understand what topics were discussed in the course. In order to find high probability words within each topic, the number of topics was set to 5 to get the high topic coherence score [17,19]. Table 2 shows the extracted words for

each topic. In Table 2, **Topic 1** code words are associated with *Effortful learning* code, **Topic 2** code words are associated with *Get beyond learning styles*, **Topic 3** code words are associated with *Illusion of mastery*, and **Topic 4** code words are associated with *Retrieval practice*, *Spacing out practice*, and *Interleaving*. Only **Topic 0** did not represent any codes.

Topic 0	Topic 1	Topic 2	Topic 3	Topic 4
lecture	desire	dylexia	confidence	mass
solution	$desire_difficulty$	$learn_style$	feedback	$mass_practice$
classroom	plf^*	individual	calibration	$interleaving_practice$
surgeon	resonate	${\rm learn_differ}$	$confidence_memory$	$space_retrieval$
acquire	parachute	disable	accuracy	tend
instruct	fall	intelligent	peer	day
learn_learn	land	prefer	answer	long_term
impact	$_{ m jump}$	$\operatorname{support}$	event	week
demand	parachute_land	dyslex	state	myth
$lecture_classroom$	land_fall	focus	$calibration_learn$	practice_space

Table 2: Five topics extracted by Latent Dirichlet Allocation

*Stands for Parachute Landing Fall.

4 Experiments

With the LDA extracted keywords for 4 topics, we conducted experiments with those keywords alone as described in 4.1, and along with the keywords identified by the instructor as described in 4.2 and 4.3.

4.1 Experiment 1

Following the results obtained from 3.2, we automatically generated a wellformatted table [20] in which each row consisted of: post entry number, user id, date and time for that post entry, actual discussion post data, and list of codes with 1's or 0's corresponding to the existence or no existence of the specific code in each post. The table was entered into the ENA webtool in an Excel format [1]. We then ran interrater reliability between the automated coding and the human coding provided by [16]. Table 3 presents the Cohen's kappa results for each code.

Table 3 shows the only code that had a Cohen's kappa moderate level of agreement [20] was the Beyond Learning Styles code. Illusion of Mastery had a weak level of agreement and the remaining codes had minimal level of agreement. In order to improve those numbers, we had asked the instructor, who manually coded the data, to provide us with keywords that we could include in the automatic process.

-0_F	
Code	Cohen's (κ)
Effortful Learning	0.23
Beyond Learning Styles	0.77
Illusions of Mastery	0.52
Retrieval Practice, Spaced out Practice, Interleaving	0.36

Table 3: Interrater reliability between automated coding process and human coding process.

Table 4: Interrater reliability between Automated + Human Keywords (A+HK) coding process and Human (H) coding process.

Code	Cohen's (κ)
Effortful Learning	0.70
Beyond Learning Styles	0.81
Illusions of Mastery	0.79
Retrieval Practice, Spaced out Practice, Interleaving	0.79

Table 5: Comparison of strength of connections between Automated + Human Keywords (A+HK) coding process and Human (H) coding process.

Connection	Strength (A+HK)	Strength (H)
illusions and retrieval-interleave	0.40	0.36
beyondLS and retrieval-interleave	0.32	0.28
effort and retrieval-interleave	0.30	0.27
beyondLS and illusions	0.27	0.18
effort and illusions	0.27	0.28
effort and beyondLS	0.21	0.22

4.2 Experiment 2

After receiving the keywords from the instructor (Table 6), we combined extracted keywords from LDA and keywords from instructor and generated a new well-formatted data table containing the same elements present in the data table from Experiment 1. Table 6 demonstrates that some of the keywords provided by the instructor were very similar to each other. In order to preserve the instructor's process, those keywords were not changed since they were used in the instructor's process of manually coding the data.

Title Suppressed Due to Excessive Length

Code	Keywords
Effortful Learning	difficult, difficulties, mistakes, failure, effortful learning, desirable difficulty, desirable, effortful
Beyond Learning Styles	instructional style, learning styles
Illusions of Mastery	illusion of mastery, illusions of mastery, misunderstand- ing, illusion of knowing, illusions of knowing, illusion of learning, illusions of learning, re read, cram
Retrieval Practice, Spaced out Practice, Interleaving	retrieval practice, retrieval process, testing effect, test effect, recall knowledge, retrieval, actively retrieving, periodically testing, retrieval activity, retrieval activities, low stakes, effective retrieval must be repeated, flash cards, quizzing, practice and retrieval, quiz over time, continually retrieve the information, frequently quizzing, retrieval practice activity, retrieval practice activities, testing efforts, active retrieval, practice, testing for its benefit in the learning process, short quiz, active recall, process of retrieval, practice sessions, self testing, recall the information, RPA, RPAs, spacing out, spacing out practice, spaced out, spaced retrieval, space retrieval, space practice, retrieval spaced, retrieve spaced, spaced application, spaced knowledge, space knowledge, spaced retrieval, retrieval practice is spaced, interleaved practice, interleaved practice, interleaved

Table 6: Keywords provided by the instructor.

Table 4 presents the Cohen's kappa for the interrater reliability between the automated process with instructor's provided keywords and the human coding. Compared to simply using the automated extracted codes, the level of agreement increased. Effortful learning and Retrieval Practice, Spaced out Practice, Interleaving codes, which previously had minimal level of agreement increased to a moderate level of agreement. Illusion of Mastery which had a weak level also increased to a moderate level, and Beyond Learning Styles, which had a moderate level increased to a strong level of agreement.

4.3 Experiment 3

The third experiment consisted of using the well-formatted data table produced from Experiment 2 and the well-formatted data table provided by [16] to create a joint well-formatted table that included an additional column, named source, to generate the ENA visualizations using the ENA webtool. All rows that contained data generated by the algorithm were labeled "algorithm" for the source column and all rows that contained human manual coding were labeled "human" for the source column.

We used the four codes produced by our approach described in Section 3.2. These codes were validated with the instructor to represent the concepts that the students were learning and therefore the concepts that the students should've connected in that online discussion. Those codes were Effortful Learning (represented simply as effort in ENA), Beyond Learning Styles (represented as beyondLS in ENA), Illusions of Mastery (represented as illusions in ENA), and Retrieval Practice, Spaced Out Practice, Interleaving (represented as retrieval-interleaving in ENA). As described in Experiment 2 and in Table 4, the interrater reliability using Cohen's Kappa reached at least $\kappa = 0.70$ for all the codes.

As we were interested in the individual student's network of concepts, both units of analysis and stanzas were students (i.e., all student messages) with an infinite stanza window. That configuration enabled us to visualize the connections between the codes for each student. To compare the model generated by the algorithm and the model generated by the human coder, the source column from our well-formatted data table was used.

5 Results

We analyzed the results produced by each data table to detect similarities and differences between the two models generated. After that, we had a meeting with the instructor to present the results to them and evaluate the two models produced. In this section we present the results from ENA generated using data from Experiment 3 and the evaluation process conducted by the instructor.

5.1 ENA Models

Figure 1a presents the group average network graph created using data from the automatic coding + human keywords (A+HK) process. The thickness of the lines between the codes indicates the strength of connections. Thicker lines indicate stronger connections, whereas thinner lines indicate weaker connections. The results indicated that for the A+HK process the strongest relationship was between the codes illusions and retrieval-interleave, followed by beyondLS and retrieval-interleave, and effort and retrieval-interleave. BeyondLS and illusions and effort and illusions had the same strength in relationship. The weakest relationship was between effort and beyondLS, as shown in Table 5.

Figure 1b presents the group average network graph created using data from the human coding process. Results show that the strongest relationships were between illusions and retrieval-interleave, followed by beyondLS and retrievalinterleave and effort and illusions. After that, the strongest relationships were between effort and retrieval-interleave and beyondLS and effort. BeyondLS and illusions connection had the weakest relation as we can observe from Table 5.

As shown by Figures 1a, 1b, and Table 5, both coding processes generated the same relationships between all the codes, with some differences between the strengths in the relationships. Figure 2 shows the difference between the A+HK



Fig. 1: ENA Visualizations.



Fig. 2: Difference between ENA generated by the A+HK coding process and the H coding process.



Fig. 3: ENA visualizations with the same structure and strengths.

model (named algorithm in the figure) and the human model, meaning that the A+HK made stronger connections between illusions and retrieval-interleave, beyondLS and retrieval-interleave, effort and retrieval-interleave, and beyondLS and illusions codes.

Using the ENA webtool, we performed a statistical analysis to verify that the difference between the two models was significant. Along the X axis (MR1), a Mann-Whitney test showed that Human (Mdn = -0.13, N = 25) was not statistically significantly different at the $\alpha = 0.05$ level from algorithm (Mdn =0.13, N = 25, U = 206.00, p = 0.04, r = 0.34). Along the Y axis (SVD2), a Mann-Whitney test showed that Human (Mdn = -0.01, N = 25) was not statistically significantly different at the $\alpha = 0.05$ level from algorithm (Mdn = -0.01, N =25, U = 318.00, p = 0.92, r = -0.02). Therefore, there is no statistical difference between the model generated by the A+HK process and the human manual coding process.

Out of the 25 ENA visualizations generated, 12 of those (48%) had the same structure for both the A+HK process and human model. From those 12, 10 had the exact same strength in connections. One example can be seen in Figures 3a and 3b. The remaining two had different strengths. In one of those two visualizations, the structure was the same but the A+HK process found stronger connections between retrieval-interleave and illusions code and the human found stronger connections between effort and illusions instead. In the other visualization, the human coding found stronger connections than the A+HK process for all codes (Figures 4a and 4b).

The remaining 13 visualizations (52%) had a different structure. In 10 of those, the A+HK process found more connections than the human process. In the remaining three, the human found more connections that the A+HK process.



Fig. 4: ENA visualizations with the same structure and different strengths.

5.2 Instructor's Evaluation

In order to evaluate the quality of the ENA visualizations generated using the A+HK process, we met with the course instructor. The intention was to gain feedback regarding the correctness of the models generated in cases where the automated process found connections that weren't supported by human analysis, as well as cases where the automated process found relationships that were missed by the human coder. We presented the results described on Section 5.1 to the instructor who's familiar with Quantitative Ethnography and the ENA Web tool. Each one of the 25 visualizations were walked through with the instructor using the ENA Web tool.

First, all ENAs that had the same structure were analyzed. All the qualitative data extracts used to find the connections between the codes for the A+HK process and the human process were looked through. As expected, for the visualizations that had the same structure and same strengths of connections, both processes used the exact same data. For those two that had a different strength between connections, in one of them (125919 user as presented in 4c) the A+HK process found a relation that the human had missed, for the other (135030 user as presented in 4a), the A+HK process presented a false positive between retrievalinterleave and illusions codes. The automated process produced a false positive in this case because it identified a keyword. However, simply having that keyword present was insufficient for the human coder to establish a relationship between the two codes.

The next step was to analyze those 13 visualizations that had different structures between the A+HK and human processes. We started by analyzing the visualizations where the A+HK process found more connections between the codes than the human process. Out of those 10 cases, in seven cases the A+HK process found connections between codes that the human had missed. Only in three cases the A+HK process generated false positives connections. To evaluate the three cases where the A+HK process found less connections, we analyzed all the qualitative data and confirmed that the A+HK process missed those connections.

6 Discussion

- RQ1. What is/are the difference(s) between the ENA visualizations generated using an automated coding process and a human coding process for the data presented in [16]?

As we can observe from Section 5.1 both the A+HK and the human processes generated the same structure for their network, with a small difference in the strength of some connections. The A+HK process found stronger connections between illusions and retrieval-interleave, beyondLS and retrieval-interleave, effort and retrieval-interleave, and beyondLS and illusions codes (Figure 2). After running a statistical analysis, we observed that there was no statistical difference between the model generated by the A+HK and the human. This could potentially be a good indicator that an automated process that used combined LDA keywords and human keywords can contribute to generating ENA visualizations to help instructors in evaluating asynchronous online discussion data. Further tests need to be done in other sections of the course to confirm that similar structures will be found between the A+HK and human process for those new datasets.

- RQ2. How does an instructor evaluate the ENA visualizations generated?

During the evaluation process, the instructor pointed out that the course used a series of assignments that required students to synthesize concepts into coherent discussion posts. Consequently, grading those posts demanded frequent reading to identify concepts and to evaluate how well each student integrated them. Recognizing the grading challenges, the instructor considered the potential usefulness of an algorithm generated ENA for potentially improving the efficiency and accuracy of grading. Comparing the A+HK coding to the human coded equivalent, the instructor mentioned that it was impressive that the A+HK generated identical coding for 11 of the 25 students. Additionally, they noted it was encouraging to see that the algorithm found more connections in 11 additional cases. In other words, 22 out of the 25 cases (88%) the A+HK identified the correct connections in 11 out of the 25 written passages and identified more correct was promising, suggesting that it may be possible to use the A+HK to highlight the majority of the connections automatically.

However, the A+HK method found fewer relationships in three cases (12%). Examining those cases, the human coder identified accurate relationships, but those relationships were extrapolated from the subtle meaning and content in the post. Future work needs to be done on how to include those aspects in the automated process. For example, improving the keywords offered by the instructor and using Large Language Models to consider the context of words [11] for a stronger connection between codes discussed in a course.

7 Conclusion

In this paper we presented an approach that clustered topic keywords into meaningful categories from a relatively small online course discussion dataset using Latent Dirichlet Allocation (LDA) [3]. Those keywords and the instructor's keywords were then used to automatically code asynchronous online discussion data. Finally, ENA visualizations were generated based on the data. The visualizations were compared with the corresponding visualizations generated by human coding process, and both visualizations were evaluated by an instructor. Results indicated that there is no statistical difference between the model generated by the A+HK process and the human.

Overall, the result of the A+HK demonstrates significant potential to assist instructors in evaluating discussion based assignments that demand the students'

synthesis and integration of concepts, especially in larger classrooms. An automated method allows instructors teaching classes with hundreds of students to use discussion posts to promote these higher order learning outcomes. It is important to acknowledge that our approach is considered as a tool for instructors to enhance their evaluation process of asynchronous online discussions. Additional efforts need to be made in order to verify its applicability to other class settings such as different student populations and different course materials.

References

- 1. Ena web tool (2023), https://app.epistemicnetwork.org/login.html
- Aloni, M., Harrington, C.: Research based practices for improving the effectiveness of asynchronous online discussion boards. Scholarship of Teaching and Learning in Psychology 4(4), 271 (2018)
- Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. Journal of machine Learning research 3(Jan), 993–1022 (2003)
- Bressler, D.M., Annetta, L.A., Dunekack, A., Lamb, R.L., Vallett, D.B.: How stem game design participants discuss their project goals and their success differently. In: Wasson, B., Zörgő, S. (eds.) Advances in Quantitative Ethnography. pp. 176–190. Springer International Publishing, Cham (2022)
- Cai, Z., Eagan, B., C., M., Shaffer, D.: Lstm neural network assisted regex development for qualitative coding. Advances in Quantitative Ethnography: Fourth International Conference, International Conference on Quantitative Ethnography 2022 https://par.nsf.gov/biblio/10354430
- Cai, Z., Siebert-Evenstone, A., Eagan, B., Shaffer, D.W.: Using topic modeling for code discovery in large scale text data. In: International Conference on Quantitative Ethnography. pp. 18–31. Springer (2021)
- 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: International Conference on Quantitative Ethnography. pp. 41–54. Springer (2019)
- Choi, J., Ruis, A.R., Cai, Z., Eagan, B.R., Shaffer, D.W.: Does active learning reduce human coding?: A systematic comparison of a neural network with ncoder. Advances in Quantitative Ethnography: Fourth International Conference, International Conference on Quantitative Ethnography 2022 https://par.nsf.gov/biblio/ 10354410
- Fernandez-Nieto, G.M., Martinez-Maldonado, R., Kitto, K., Buckingham Shum, S.: Modelling spatial behaviours in clinical team simulations using epistemic network analysis: Methodology and teacher evaluation. In: LAK21: 11th International Learning Analytics and Knowledge Conference. p. 386–396. LAK21, Association for Computing Machinery, New York, NY, USA (2021). https://doi.org/10.1145/3448139.3448176, https://doi.org/10.1145/3448139.3448176
- Garrison, D.R., Anderson, T., Archer, W.: Critical thinking, cognitive presence, and computer conferencing in distance education. American Journal of distance education 15(1), 7–23 (2001)
- Ghaffari, S., Krishnaswamy, N.: Grounding and distinguishing conceptual vocabulary through similarity learning in embodied simulations. arXiv preprint arXiv:2305.13668 (2023)

- 12. Herder, T., Swiecki, Z., Fougt, 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: Proceedings of the 8th International Conference on Learning Analytics and Knowledge. p. 21–25. LAK '18, Association for Computing Machinery, New York, NY, USA (2018). https://doi.org/10.1145/3170358.3170394, https://doi.org/10.1145/3170358.3170394
- Koh, J.H.L., Herring, S.C., Hew, K.F.: Project-based learning and student knowledge construction during asynchronous online discussion. The Internet and Higher Education 13(4), 284–291 (2010)
- de Lima, D.P., Gerosa, M.A., Conte, T.U., de M. Netto, J.F.: What to expect, and how to improve online discussion forums: the instructors' perspective. Journal of Internet Services and Applications 10, 1–15 (2019)
- Marquart, C.L., Swiecki, Z., Eagan, B., Shaffer, D.W.: Package 'ncodeR' (2019), https://cran.r-project.org/web/packages/ncodeR/ncodeR.pdf, Online. [Accessed: May 18, 2022]
- Moraes, M., Folkestad, J., McKenna, K.: Using epistemic network analysis to help instructors evaluate asynchronous online discussions. In: Second International Conference on Quantitative Ethnography: Conference Proceedings Supplement. pp. 19–22 (2021)
- Newman, D., Lau, J.H., Grieser, K., Baldwin, T.: Automatic evaluation of topic coherence. In: Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics. pp. 100– 108 (2010)
- Nguyen, H.: Exploring group discussion with conversational agents using epistemic network analysis. In: Wasson, B., Zörgő, S. (eds.) Advances in Quantitative Ethnography. pp. 378–394. Springer International Publishing, Cham (2022)
- Saravani, S.M., Ghaffari, S., Luther, Y., Folkestad, J., Moraes, M.: Automated code extraction from discussion board text dataset. In: Advances in Quantitative Ethnography: 4th International Conference, ICQE 2022, Copenhagen, Denmark, October 15–19, 2022, Proceedings. pp. 227–238. Springer (2023)
- Shaffer, D.W., Collier, W., Ruis, A.R.: A tutorial on epistemic network analysis: Analyzing the structure of connections in cognitive, social, and interaction data. Journal of Learning Analytics 3(3), 9–45 (2016)
- Shah, M., Siebert-Evenstone, A., Moots, H., Eagan, B.: Quality and safety education for nursing (qsen) in virtual reality simulations: A quantitative ethnographic examination. In: Wasson, B., Zörgő, S. (eds.) Advances in Quantitative Ethnography. pp. 237–252. Springer International Publishing, Cham (2022)
- 22. Shum, S.B.: Qe visualizations as tools for thinking [powerpoint slides], international conference on quantitative ethnography 2021 keynote speaker (2021), https://simon.bucking-hamshum.net/wp-content/uploads/2021/02/SBS_ ICQE2020_Keynote.pdf
- Swiecki, Z., Ruis, A., Shaffer, D.W.: Modeling and visualizing team performance using epistemic net-work analysis. In: Proceedings of the 7th annual GIFT users symposium. pp. 148–156 (2019)
- 24. Thomas, M.J.: Learning within incoherent structures: The space of online discussion forums. Journal of Computer Assisted Learning 18(3), 351–366 (2002)
- Vega, H., Irgens, G.A.: Constructing interpretations with participants through epistemic network analysis: Towards participatory approaches in quantitative ethnography. In: Wasson, B., Zörgő, S. (eds.) Advances in Quantitative Ethnography. pp. 3–16. Springer International Publishing, Cham (2022)