THESIS

ONDISCUSS: VISUALIZING ASYNCHRONOUS ONLINE DISCUSSIONS THROUGH AN EPISTEMIC NETWORK ANALYSIS TOOL

Submitted by

Yanye Luther

Department of Computer Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2024

Master's Committee:

Advisor: Marcia Moraes

Sudipto Ghosh
James Folkestad

ABSTRACT

ONDISCUSS: VISUALIZING ASYNCHRONOUS ONLINE DISCUSSIONS THROUGH AN EPISTEMIC NETWORK ANALYSIS TOOL

Asynchronous online discussions are common assignments in both hybrid and online courses to promote critical thinking and collaboration among students. However, the evaluation of these assignments can require considerable time and effort from instructors. We created OnDiscuss, a learning analytics visualization tool for instructors that utilizes text mining algorithms and Epistemic Network Analysis (ENA) to generate visualizations of student discussion data. Natural language processing and text mining techniques are used to generate an initial codebook for the instructor as well as automatically code the data. This tool allows instructors to edit their codebook and then view the resulting ENA networks for the entire class and individual students. Through empirical investigation, we assess this tool's effectiveness to help instructors in analyzing asynchronous online discussion assignments.

Our findings highlight several key insights regarding the implications of this tool for enhancing the accessibility and usability of ENA as a learning analytics visualization tool. OnDiscuss is helpful to those unfamiliar with ENA since it abstracts many of the intricacies of ENA by providing an easy interface to manipulate a codebook and thus the resulting ENA networks. Future refinements, such as the addition of a baseline ENA model, can make it more helpful to those familiar with ENA. Despite the tool's automated keyword generation capabilities, it is clear that instructor intervention remains crucial for refining the codebook. Therefore, while automated techniques like Latent Dirichlet Allocation (LDA) provide valuable insights given a large amount of data, these processes must be complemented by expert guidance.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Dr. Marcia Moraes, for her invaluable guidance, unwavering support, and insightful feedback while working on this thesis and during my entire research experience. I would have never thought I would be in grad school completing my Master's when I started as an undergraduate research assistant under Marcia my sophomore year. Marcia's expertise and patience have been instrumental in shaping this research and in my development in academia. Thank you for spending all that time in your office and over Teams listening to the flood of things I'd have to say about anything and everything.

I would also like to extend my sincere thanks to the members of my thesis committee, Dr. Sudipto Ghosh and Dr. James Folkestad, for their time, guidance, and support throughout my research.

To my family and friends, thank you for keeping me sane during my studies. While I'm sure it was hard to pull me away from my desk sometimes, thank you for sticking to it and not letting me work too hard. I'm grateful for the many laughs and late night homework sessions in the 120 lab. Your camaraderie was everything to me during my time at CSU. And of course, thank you to my parents for supporting my sudden spur of the moment, "I want to go to grad school" decision which I'm sure you both were expecting to some degree.

Extra huge shout out to my partner, Emma, for listening to all my research related rants on top of all my other computer science rants. Also for editing all my papers littered with spelling and grammar mistakes. You always came to the rescue making sure my posters and slide shows looked presentable. Since I couldn't quite put your name on my other published works, here is your much deserved recognition.

DEDICATION

I would like to dedicate this thesis to all the students writing their discussion post in Canvas for class - don't forget to reply to two other classmates.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

Introduction

The widespread integration of digital technology in higher education impacts both teaching and learning practices, enabling access to data primarily derived from online learning environments that can be leveraged to enhance student learning. Online learning has become an integral part of higher education facilitating both synchronous and asynchronous interaction in virtual environments [2]. Thus higher education institutions are implementing Learning Analytics (LA) systems to better understand and support student learning [3]. According to the 1st International Conference on Learning Analytics and Knowledge, learning analytics is *"the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs"* [4]. Learning analytics has significant potential to drive personalized and adaptive educational experiences.

Given the importance of understanding student interactions, it is crucial to recognize that social interactions play a significant role in the learning process. Vygotsky's [5] social learning theory states that learning occurs primarily within social and cultural settings where students interact with their peers, teachers, and parents as active participants to construct their own knowledge. In this framework interpersonal interactions and discussions are a vital part of teaching and learning [6,7]. Therefore, asynchronous online discussions are a common tool to facilitate social interactions in both hybrid and online courses [8, 9]. Asynchronous online discussions offer many benefits to students such as a deeper understanding of course material [10], more effective communication with group members [10], improvements in critical thinking and writing skills [8], and increased student performance in meeting learning outcomes [11].

However, despite these benefits, instructors face challenges in assessing students' contributions [12]. According to de Lima et al. [12] this is 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.

1.1 Epistemic Network Analysis

In order to address some of those difficulties, Epistemic Network Analysis (ENA) has been presented as learning analytics visualization tool to show the relationships among the different concepts students discuss in an asynchronous online discussion [1].

ENA was developed to model cognitive networks based on the fundamental theories of learning analytics that *the structure of connections among cognitive elements is more important than the mere presence or absence of those elements in isolation* [13]. Shaffer [13] characterized learning as the development of an epistemic frame, a pattern of associations among knowledge, skills, and habits of mind. Based on these theoretical foundations, ENA models the structure and weight of these connections among cognitive elements as a dynamic network.

Now in order to tag and discover these elements found in qualitative data such as student discussion posts, researchers must code the data. Coding is the process of labeling segments of data with tags or codes that represent specific themes, concepts, or categories [14]. This helps researchers organize and make sense of qualitative data, such as interviews, focus group discussions, or open-ended survey responses and is predominantly used in the social sciences. There are many types of coding and ways to arrive at a codebook. Researchers may generate the codes from the data, theory, or literature from the field and then perform a long process of identifying, organizing, aggregating, refining, and applying the codes. It is a rigorous procedure to not only define a codebook but then also code the data. Despite the rich data analysis that can emerge from coding, it is fundamentally a very labor intensive and challenging task [15].

The researchers in Moraes et al. [1] created an "a prior" codes that represented concepts that students were learning and thus should connect in online discussion. By manually coding the student discussion posts they were able to create ENA networks such as Figure 1.1.

A key assumption of using ENA to model any kind of network is that the structure of the connections in the data is the most important in the analysis [13]. So much like the fundamental theories of learning it's the structure of those connections rather than the presence or absence of elements in isolation. The connections among elements are derived from each unit of analysis (e.g.,

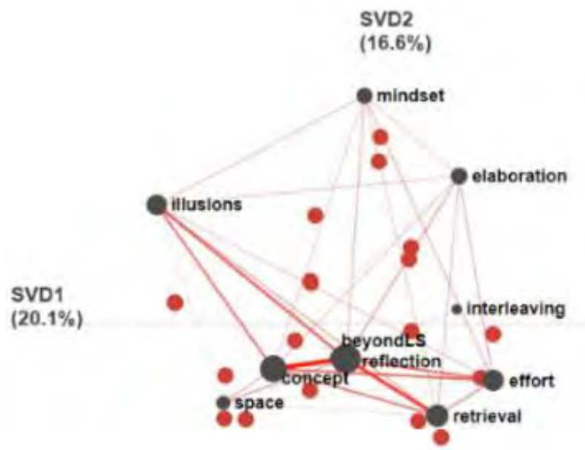


Figure 1.1: Group Average Network from Moraes et al. [1]

study subject) based on the code co-occurrences in the data subsets, called stanzas (e.g., sentence, paragraph, or document). Codes that occur within the same stanza receive a 1, and codes that do not co-occur receive a 0. The following are important aspects of an ENA network [16]:

- **Nodes** represent individual codes.
- **Links or connections** between nodes represent the strength of association of those codes. This is proportional to the relative frequency of their co-occurrence. The thickness of the lines between the codes indicates the strength of connections, thicker lines indicate stronger connections, whereas thinner lines indicate weaker connections.
- **Co-occurrence** of two codes means that they are both interpretations of the data in the same temporal context.
- **Centroid** summarizes the network as a single point in space. This allows comparison among different networks because centroids located close together represent networks with similar patterns of connections, while centroids located far apart represent networks with different patterns of connections.

1.2 Motivation and Contributions

The motivation behind this thesis was based on the initial work done in Moraes et al. [1]. Based on Moraes et al. [1], we submitted a project proposal to our university's teaching innovation grant to examine the use of ENA as a visualization tool. One piece of feedback provided by the reviewers was that instructors would not have time to be involved in the coding process, even if that process used nCoder [17]. The reviewers pointed out that they would like to have access to the visualization but not have the "burden" to build the codes. However, they would be willing to provide keywords that should be present in the codes.

Considering that feedback, as a part of my undergraduate work, we decided to apply text mining and Natural Language Processing (NLP) algorithms to automate the extraction of codes from data and use the keywords that were provided by the instructors as guiding input to the algorithms [18]. We found that Latent Dirichlet Allocation (LDA) with 5 topics produced the most codes provided by humans manually coding the data. Then we wanted to compare ENA networks constructed by manual human coding versus automated coding with the addition of provided keywords (Chapter 3). The networks showed no statistical difference from one another. We then developed a complete ENA learning analytics visualization tool, OnDiscuss, utilizing the same LDA algorithm and allowed instructors to modify the generated codebook. A case study on two instructors revealed the OnDiscuss's ability to support instructors when evaluating their course's asynchronous online discussions (Chapter 4).

This work explores the use of ENA as a learning analytics visualization tool to support instructors in evaluating asynchronous online discussions. This thesis is a culmination of several works completed and published during my studies.

Chapter 2

Literature Review

2.1 Topic Modeling

When dealing with very large text datasets, manually coding all of the text becomes impractical. Recent advancements in machine learning and natural language processing (NLP) enable the automatic identification of patterns in data that may elude human detection, as well as specific instances of existing codes that might be overlooked. One of the most widely used NLP techniques for automated coding is topic modeling. This method analyzes a large dataset to create clusters of words, known as topics, each characterized by a probability distribution that assigns likelihoods to the words within the data. The words with the highest probabilities help convey the "meaning" of the identified topics. Therefore, statistical and machine learning tools such as nCoder [17] and LightSIDE [19] have emerged.

Previous works such as Cai et al. [20] utilized nCoder for topic modeling for two large datasets, one from a medical education course and another from engineering design teams participating in an educational simulation, investigating how close the human codes were to the codes identified by nCoder. 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 [21]. In their comparison, they indicated nCoder had a higher accuracy. However, nCoder is not fully automated. It requires human in the loop to read through the text and validate if the coding is conceptually valid. It also suffers from low recall. nCoder+ [22] aimed to improve low recall issue in nCoder through semantic component addition.

Our work in Saravani et al. [18] has several distinctive features from prior works. First, the initial codebook is completely automated and created through Latent Dirichlet Allocation (LDA). Second, we utilize coherence analysis [23] to identify the optimal number of topics in the discussion data, thus avoiding arbitrary selection of the number. Finally, while human intervention is not needed for the initial codebook generation, instructors are then able to manually edit the initial codebook by removing, adding, and renaming keywords.

2.2 ENA Applications and Uses

Epistemic Network Analysis (ENA) has been used as an analysis technique in several different domains such as health care [24], politics [25], history [26], and psychology [27]. ENA has been used to support many facets of education and learning in several areas [28–30]. A growing area of interest for ENA is its application as a visualization tool [31] to help instructors evaluate clinical team simulations [32], support teachers' interventions in students' virtual collaboration [33], evaluate teamwork [34]. Vega and Irgens [35] introduced participatory quantitative ethnography (QE) which includes participants in co-construction and co-interpretation of ENA models. This work demonstrates the deeper analysis that arises from interpretations of data by modifying codes, adding/removing connections, and reacting to codes. All of these studies utilized different coding strategies to code the data used in the ENA visualizations.

2.2.1 ENA in Collaborative Learning

ENA has been used in several works related with collaborative learning in asynchronous online discussions [36–39]. Rolim et al. [36] used ENA to provide insights on the relationship between social and cognitive presence in asynchronous online discussions. The work was based on the Community of Inquiry (CoI) model defines three dimensions that mold the learning experience (social presence, cognitive presence, and teaching presence) and assumes an overlapping relationship among the three presences that enhance the students online learning capability. Scianna and Kaliisa [39] proposed an analysis workflow and visualization method called Social Sentiment Embedded Epistemic Networks (SSEEN) which combines ENA and sentiment analysis (SA) to consider how sentiment manifests and explores its usefulness in understanding student interactions in asynchronous online discussions.

Gašević et al. [37] proposed the use of social epistemic network signature (SENS), which combines ENA and Social Network Analysis (SNA) to analyze collaborative learning. ENA was related to the content of the student discourse and SNA was used to link that to the roles the students played in their communities. Each analysis technique was independent of one another and thus resulted in two separate networks. Swiecki and Shaffer [38] extended SENS and proposed the integrated social-epistemic network signature (iSENS), an approach that provides the simultaneous investigation of cognitive and social connections in collaborative learning from a single network combined from both ENA and SNA.

Overall ENA has been used by itself and in conjunction with other analysis techniques to fully understand the many aspects that go into collaborative problem solving. While the social presence, role, and sentiment of a student does impact their collaborative learning, in this thesis we are entirely focused on the content the student is discussing.

2.2.2 Need for an Instructor Tool for Visualizing Learning with ENA

ENA has been used to visualize many aspects of learning and education [32]. Fougt et al. [40] analyzed instructors' ability to assess student papers and considered a different range of number of topics in order to capture different levels of complexity of the ENA models. Visually, ENA had the potential to indicate the quality of the student assignment. The instructors reported it can be difficult to choose the correct number of codes and keywords and that it should be left to the instructor to make that informed decision. Thus demonstrating the need for instructors to be able to edit their codebook for multiple iterations. Herder et al. [33] created an ENA tool integrated into a virtual internship for teachers to support students. While the teachers all valued the possibilities this tool provided in supporting students' work, two of the three teachers interviewed reported either having difficulties interpreting the networks or it using a lot of cognitive energy to understand. While the teachers' experience with ENA was not mentioned, for the study presented in this thesis instructors received a short presentation introducing them to ENA.

Unlike previous works that used ENA in collaborative learning analysis, this work combines LDA and ENA in a tool that: builds an initial codebook for instructor, automatically codes the data of asynchronous online discussions using that initial codebook, and generates individual ENA visualizations of student's content connections. In addition, past works do not have the ability to modify the codebook and thus instantly update the ENA model on their own without any researcher intervention.

Chapter 3

Combining Automated Coding and Instructor Input[1]

3.1 Introduction

In this work, we aim to use the automatically generated codes provided in Saravani et al. [18] to code the same dataset used in Moraes et al. [1]. Then evaluate the quality of the automated coding with the human coding presented in Moraes et al. [1] by interrater reliability. If necessary we then improve the automated code generation in order to reach a satisfactory interrater reliability between algorithm and human. Applying the codes that were automatically generated in the creation of ENA visualizations. Analyzing and evaluating those ENA visualizations with human coded ENA visualizations. 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 [1]?
- **RQ2.** How does an instructor evaluate the ENA visualizations generated?

3.2 Methodology

The main goal of our approach was to extract topic keywords from a relatively small online discussion dataset using Latent Dirichlet Allocation (LDA) [42], 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.

[1]The content of this chapter was adapted from the published work, Marcia Moraes, Sadaf Ghaffari, Yanye Luther, and James Folkesdtad. Combining automatic coding and instructor input to generate ena visualizations for asynchronous online discussion. In Golnaz Arastoopour Irgens and Simon Knight, editors, *Advances in Quantitative Ethnography*, pages 381–394, Cham, 2023. Springer Nature Switzerland

3.2.1 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 3.1 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.

Table 3.1: Priori codes

Code Name	**Definition**	**Kappa (κ)**
Retrieval practice, 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 effect. Spacing out practice allows people to a little forgetting that helps their process of consolidation. 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 mastery	Researches have pointed out that students usually have a misunderstanding about how learning occurs and engage with learning strategies that are not beneficial for their long-term retention, 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 impediments (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

3.2.2 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 [42], 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 [**?**,23]. Table 3.2 shows the extracted words for each topic. In Table 3.2, **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.

Table 3.2: Five topics extracted by Latent Dirichlet Allocation

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	learn_differ	confidence_memory	space_retrieval
acquire	parachute	disable	accuracy	tend
instruct	fall	intelligent	peer	day
learn_learn	land	prefer	answer	long_term
impact	jump	support	event	week
demand	parachute_land	dyslex	state	myth
lecture_classroom	land_fall	focus	calibration_learn	practice_space

*Stands for Parachute Landing Fall.

3.3 Experiments

With the LDA extracted keywords for 4 topics, we conducted experiments with those keywords alone as described in 3.3.1, and along with the keywords identified by the instructor as described in 3.3.2 and 3.3.3.

Table 3.3: Interrater reliability between automated coding process and human coding process.

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 3.4: 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

3.3.1 Experiment 1

Following the results obtained from 3.2.2, we automatically generated a well-formatted table [13] 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 [43]. We then ran interrater reliability between the automated coding and the human coding provided by [1]. Table 3.3 presents the Cohen's kappa results for each code.

Table 3.3 shows the only code that had a Cohen's kappa moderate level of agreement [13] 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.

Table 3.5: 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

3.3.2 Experiment 2

After receiving the keywords from the instructor (Table 3.6), 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 3.6 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.

Table 3.4 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.

3.3.3 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 [1] to create a joint well-formatted table that

Table 3.6: Keywords provided by the instructor.

Code Name	Definition
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, misunderstanding, 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 practice, spacing practice, spaced 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, interleaving, interleaved practice, interleave, interleaved

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 3.2.2. 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 3.4, 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.

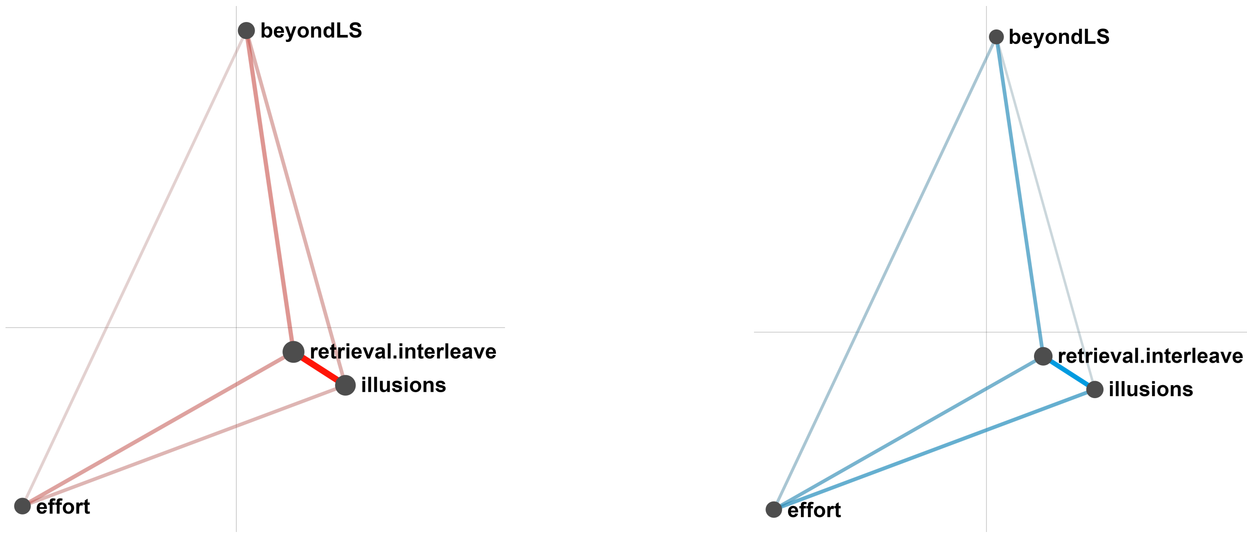
3.4 Results

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.

3.4.1 ENA Models

Figure 3.1a 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 3.5.

Figure 3.1b 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 retrieval-interleave 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 3.5.



(a) ENA Network from the A+HK coding process.

(b) ENA Network from the H coding process.

Figure 3.1: A+HK and H Resulting ENA Networks

As shown by Figures 3.1a, 3.1b, and Table 3.5, both coding processes generated the same relationships between all the codes, with some differences between the strengths in the relationships. Figure 3.2 shows the difference between the A+HK 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 ($Mdn = -0.01, N = 25, U = 318.00, p = 0.03, r = -0.02$). Therefore, there is no statistical difference between the model generated by the A+HK process and the human manual coding process.

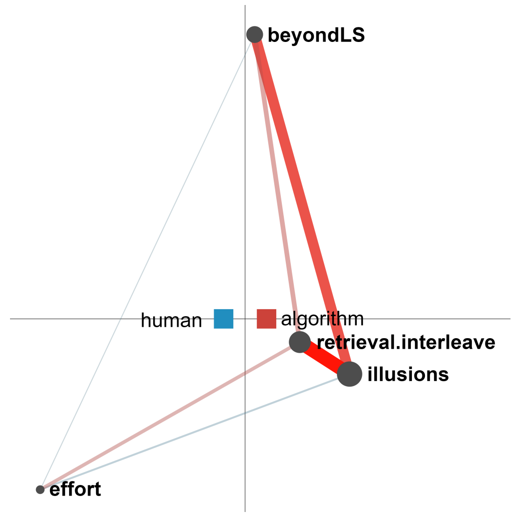


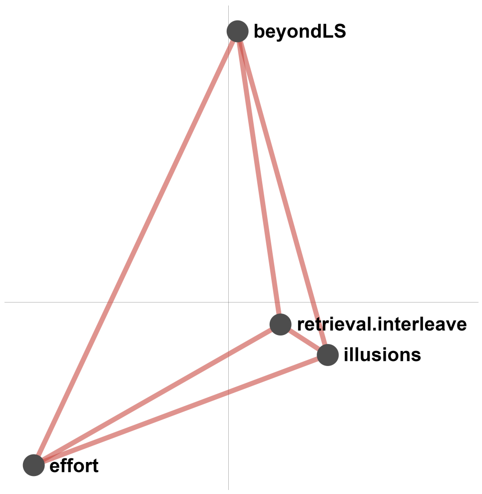
Figure 3.2: Difference between ENA generated by the A+HK coding process and the H coding process.

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 3.3a and 3.3b. 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 A.1a and A.1b).

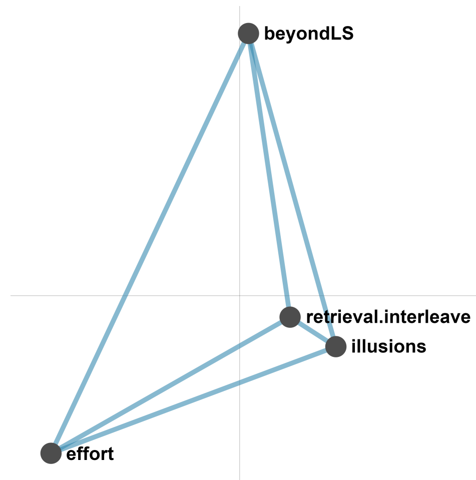
The remaining 13 visualizations (52%) had a different structure. In 10 of those, the A+HK process found more connections than the human process. In the remaining three, the human found more connections that the A+HK process.

3.4.2 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 3.4.1 to the



(a) User 142854 using A+HK coding process.



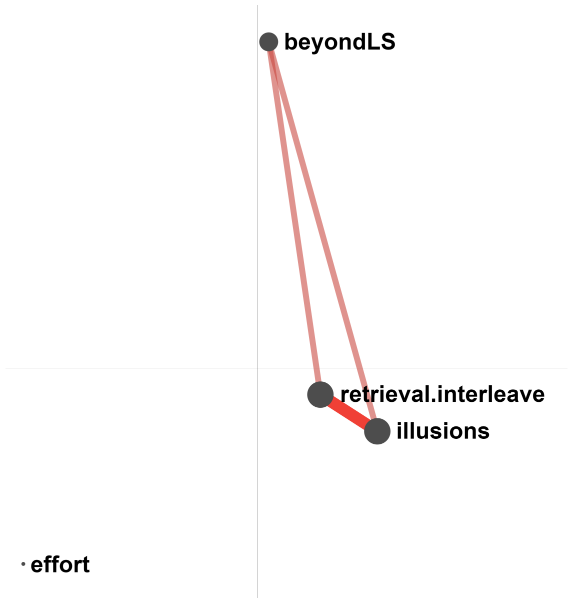
(b) User 142854 using H coding process.

Figure 3.3: ENA visualizations with the same structure and strengths.

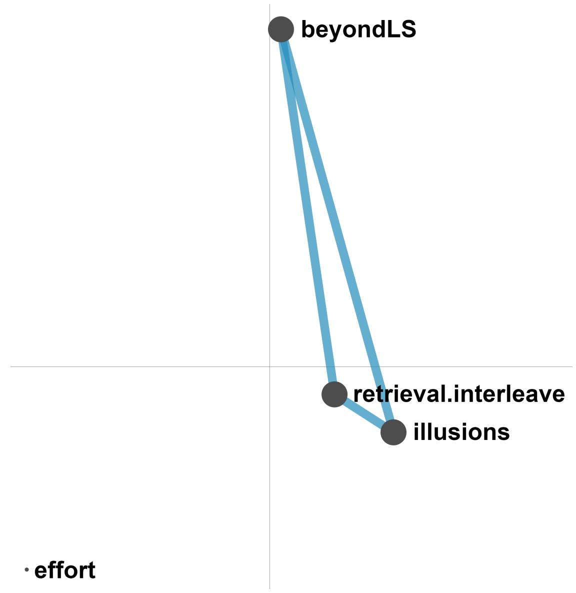
instructor who's familiar with Quantitative Ethnography and the ENA Web tool. Each one of the 25 visualizations were walked through with the instructor using the ENA Web tool.

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 A.1c) the A+HK process found a relation that the human had missed, for the other (135030 user as presented in A.1a), the A+HK process presented a false positive between retrieval-interleave 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.

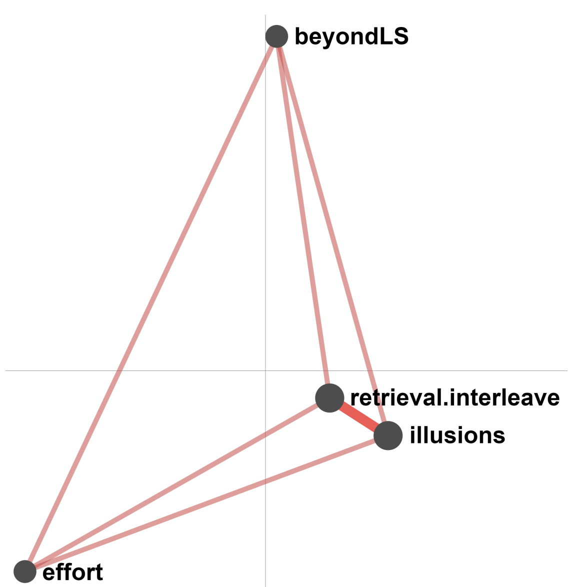
The next step was to analyze those 13 visualizations that had different structures between the A+HK and human processes. We started by analyzing the visualizations where the A+HK process found more connections between the codes than the human process. Out of those 10 cases, in seven cases the A+HK process found connections between codes that the human had missed. Only



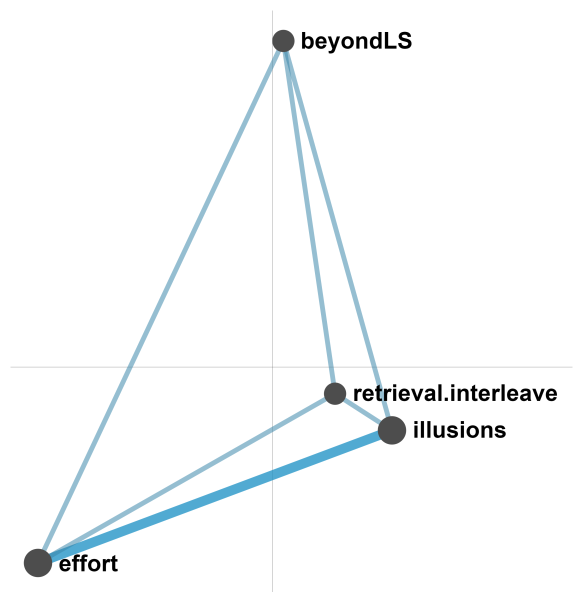
(a) ENA for user 135030 using A+HK coding process.



(b) ENA for user 135030 using H coding process.



(c) ENA for user 125919 using A+HK coding process.



(d) ENA for user 125919 using H coding process.

Figure 3.4: ENA visualizations with the same structure and different strengths.

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.

3.5 Discussion

3.5.1 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 [1]?

As we can observe from Section 3.4.1 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 3.2). 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.

3.5.2 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 conceptual connections in the written passages and identified more correct connections in 11 out of the 25 written passages (44%). The level of accuracy 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 [44] for a stronger connection between codes discussed in a course.

3.6 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) [42]. 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.

Chapter 4

OnDiscuss: An ENA Tool[2]

4.1 Introduction

In this paper, we present OnDiscuss, a learning analytics (LA) visualization tool to support instructors in evaluating asynchronous online discussions. The tool utilizes both text mining and Epistemic Network Analysis (ENA). Latent Dirichlet Allocation (LDA) [42] is used to perform automatic topic extraction from the discussion data to create an initial codebook based on the findings of [18]. Bigrams and trigrams are incorporated in LDA and topics are extracted over the entire discussion thus using an infinite stanza window [18]. The use of LDA with the intervention and addition of the instructor's own keywords has demonstrated significant potential to assist instructors in evaluating discussion based assignments [41]. However, in [41] the ENA visualizations were only demonstrated to a single instructor, requiring manual presentation. This current study investigates the potential for broader applications by exploring how different instructors utilize a tool that enables them to modify their own codebook and observe immediate changes in ENA models.

This paper aims to answer the following research questions:

- **RQ1.** Are ENA visualizations of asynchronous online discussion data helpful for instructors who are novices in ENA?

- **RQ2.** Do the ENA visualizations have the potential to reduce the time and effort spent assessing asynchronous online discussions for instructors of any experience level with ENA?

[2]The content of this chapter was adapted from a paper accepted to the International Conference of Quantitative Ethnography (ICQE) 2024

4.2 Methodology

4.2.1 Participant Selection

In order to answer the research questions, we selected two instructors; one instructor that did not have previous experience with ENA and one instructor with previous experience with ENA to evaluate their respective asynchronous online discussion with the aid of OnDiscuss. Both instructors taught at the same Research 1 land-grant university. The instructor without experience in ENA is referred to as Instructor A, the novice, and the instructor with experience is referred to as Instructor B, the expert. Instructor A taught a graduate level Computer Science course with discussion discourse amongst small groups of 5-6 students with about 2-3 groups in the entire class. Instructor B taught a graduate level Education course with discourse amongst all 20 students in the class.

The courses that were chosen for the experiment had to meet the following criteria: the course must have asynchronous online discussion assignments; those assignments must require students to assimilate knowledge and synthesize concepts into a coherent discussion post rather; and the discussions must be entirely raw text since OnDiscuss can't parse potential keywords in images, videos, links to websites, etc.

4.2.2 OnDiscuss Description and Functionality

OnDiscuss is a learning analytics visualization tool for instructors that utilizes text mining algorithms and Epistemic Network Analysis (ENA) to generate visualizations of student discussion data. As highlighted by reviewers in [1], the process of establishing a codebook from scratch may present challenges for instructors new to ENA. Our tool uses the same process described in [41] to generate five topics with ten keywords, each derived from the discussion's data to be used as an initial codebook for the instructor. The rationale for the initial number of topics and keywords can be found in [18]. After the initial codebook is built, the tool uses text mining algorithms to automatically code the discussion for each post. That coded data is then represented in an ENA network. The tool uses rENA [45] package and the following settings to build the ENA model:

unit is the student ID, the conversation is the utterance from the student in their discussion post, and an infinite stanza window. That configuration enabled us to generate the connections between the codes that were to be discussed by each student.

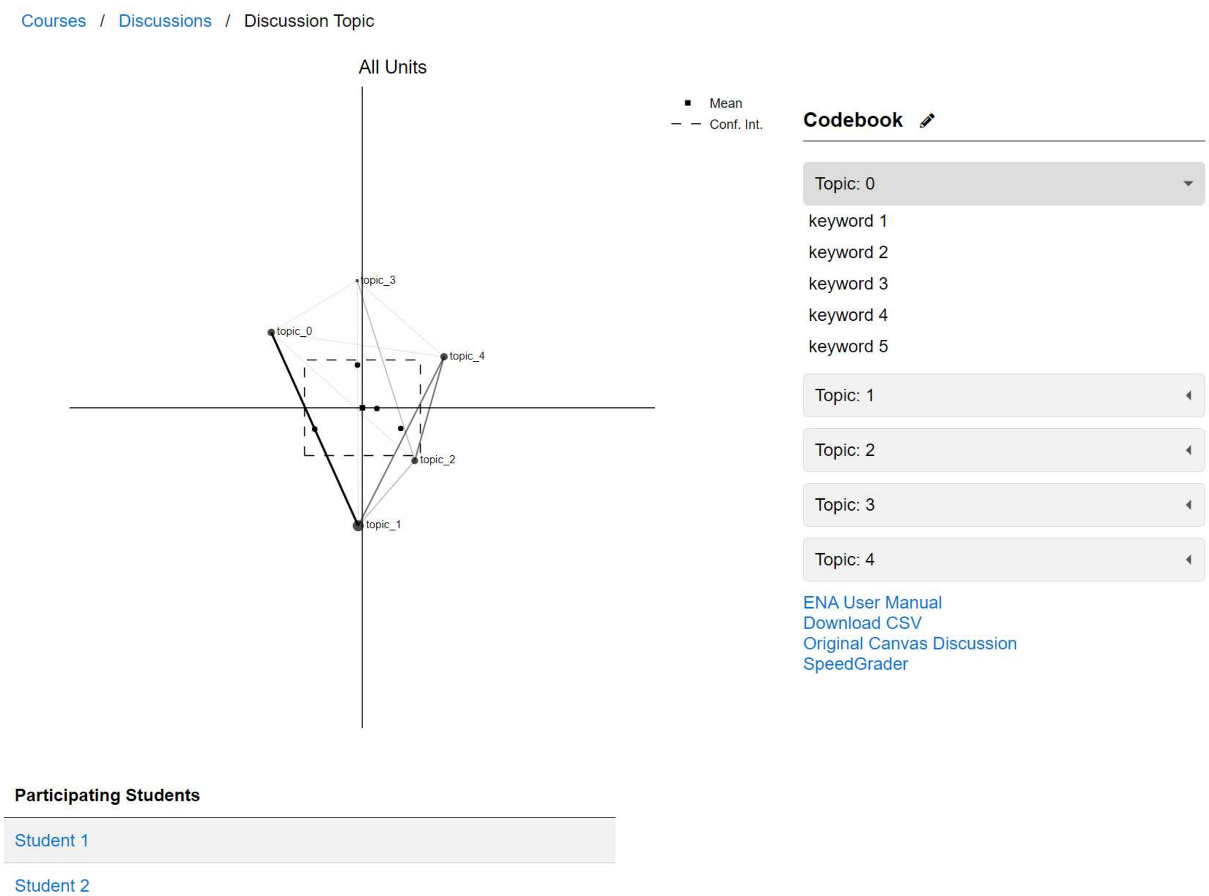


Figure 4.1: Example Class/Group ENA View for a Discussion Topic

OnDiscuss is integrated with Canvas Learning Management System (LMS) which allowed instructors to view a list of all discussions published to Canvas within a course. Clicking on a discussion displays the group ENA model and the associated codebook as shown in Figure 4.1. Instructors can edit their codebook by adding, removing, and editing keywords. They cannot edit the number of topics since [18] found 5 to be the optimal number of topics for grouping keywords. However, they can edit the names of the topics to be more descriptive since LDA will just assign the topics as numbers 0-4. Once the edits are completed, the ENA models automatically update to account for the codebook changes. Instructors are also able to view individual students' networks and the discussion posts that contributed to the network (Figure 4.2).

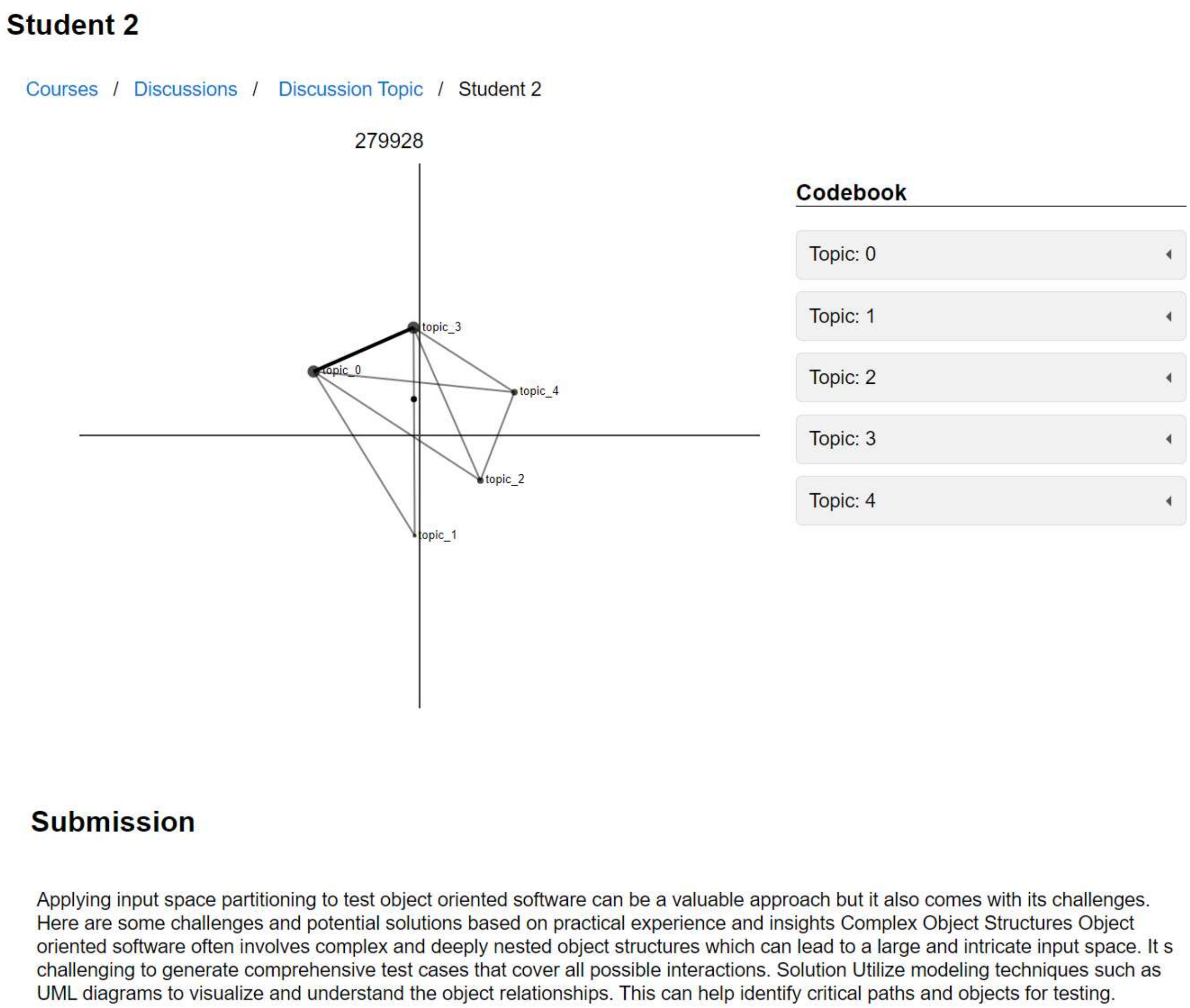


Figure 4.2: Example Individual ENA View for a Discussion Topic

Shown in Figure 4.1, OnDiscuss provides a link to the discussion in Canvas and a link to the SpeedGrader for easy access to grade and view the original discussion in Canvas. There was also a link to the "ENA User Manual," which provided reinforcement of what ENA is and how to interpret an example network. A link to download the comma separated value (CSV) file was provided in the correct format for the ENA Web Tool [46].

4.2.3 Experimental Procedure

The experiment was conducted as an empirical observational study [47], where we observed the instructor's interaction with OnDiscuss during semi-structured interview [48] sessions that lasted around one hour. We recorded the interview sessions and took notes about instructor's interactions and questions. We provided assistance in response to any inquiries raised by the instructor during the session. Because discussion topics were pulled from past semesters of the course, the instructors were asked to review the selected discussion prior to the interview session. This was to allow the instructors to reacquaint themselves with the discussion topic and the students' postings and discourse.

To begin the session, we delivered a presentation that went over the basics of what ENA is, a case study illustrating ENA interpretation, and the impact of discussion data and codebook modifications on visualizations. The presentation began by explaining that ENA identifies the co-occurrences in segments of discourse data and modeling the weighted structure of co-occurrences as a dynamic network model [49]. We then presented a case study on one discussion topic from a single semester of a course. An initial codebook was generated using LDA and was used to create the group class and individual student models where instructors learned that the nodes represented topics, the thickness of the lines represented the strength of connections between topics, and the other points in the group network represented the centroids of the individual students' networks. The presentation gave examples of student posts and how keywords co-occurred in the discourse data to create the ENA model. Because this case study was performed on a single discussion topic from a single semester, additional networks of the same discussion topic from a different semester were shown to demonstrate the impact of the discussion data using the same codebook. Finally, we made our own edits to the initial codebook and demonstrated how such modifications could influence the resulting networks. This presentation was delivered to both instructors, irrespective of their familiarity with ENA. Sample slides from the presentation can be found in Appendix A.

After the presentation the instructors navigated to the chosen discussion assignment in OnDiscuss. The initial codebook was completely generated by LDA from previous semesters of the same discussion topic [18]. The instructors were allowed to make as many edits and iterations of the codebook while inspecting the models produced by each iteration. Once the instructor was satisfied with the codebook and/or the group and individual ENA models, a closing semi-structured interview was conducted. The complete list of questions can be found in Appendix B.

4.3 Results

4.3.1 Instructor A

The novice with ENA asked many questions about the basics of ENA during the presentation. For example, *What is a codebook?*, *How are the topics positioned in 2D space?*, and *What should I be expecting to understand from these visualizations?*. After responding to the instructor's questions and completing the presentation, the instructor navigated to their chosen discussion on OnDiscuss. At this point, they had access to the initial codebook created from running LDA on all the previous semesters of the same discussion topic (Table 4.1) the resulting group network (Figure 4.4a), and all the individual networks.

Table 4.1: Instructor A Initial Codebook

Topic	Keywords
0	devic, interfac, child, applic, potenti, post, input_paramet, parent, behavior, run
1	write, want, team, choic, custom, field, look, product, interfac, array
2	boundari, import, select, api, handl, rest_api, rest, encapsul, sure, partit_test
3	partit_method, categori_partit, leak, determin, memori_leak, languag, applic, system, databas, partit_test
4	subclass, tester, abstract, output, group, model, oop, overlap, detect_memori, disjoint

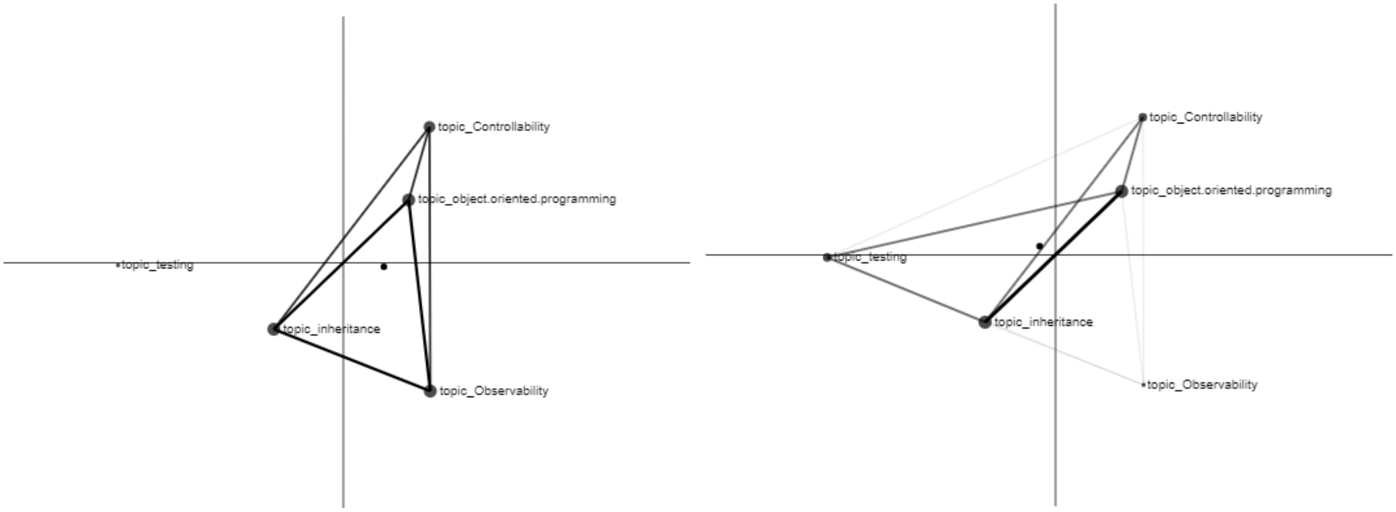
The instructor started their interaction with the tool by removing all the initial LDA generated keywords from each topic and added their own keywords. They said that half of all the automatically generated keywords were useful individually but the groupings of the keywords into topics was not helpful. Consequently, they removed all keywords and added those they thought would be present. The first iteration of their codebook did not have enough keywords that co-occurred in the discourse data so the network only consisted of a single line.

We intervened and suggested that this could've been due to multiple reasons such as not including the stems of keywords, not choosing enough keywords to represent a topic, and choosing keywords not present in the discussion. Taking this advice, the instructor included the proper word stems, added more keywords to existing topics, and completely reworked a topic that only had 2 keywords to a new topic with 12 keywords. The codebook that they were most pleased with is shown in Table 4.2 and the resulting group network is shown in Figure 4.4b.

Table 4.2: Instructor A Best Codebook

Code	**Keywords**
Observability	observability, visible, get, state, visibility, observable, getter, access, field, accessor
Controllability	configure, object, control, modify, state, mutate, mutator, update
inheritance	inheritance, child class, parent class, overriding, depth, subclass, superclass, sub class, super class, inherit, interface, abstract class
testing	black box, black-box, white-box, white box, automate, automation, industry, difficult, easy
object oriented programming	public, private, protected, package, simple, complex, abstraction, specialization, data, encapsulation, method, field

Analyzing the networks from the codebook shown in Table 4.2, Instructor A stated that the connection strength between topics in the group ENA shown in Figure 4.4b did indeed reinforce what they had gathered from rereading the discussion. The individual ENA graphs were also representative of both the "strong" and "weak" students' contributions. When analyzing a strong student's network (Figure 4.3a), the instructor was shocked to not see any connections to the topic "testing" but stated, *This student made good strong connections to controllability and observability*. When analyzing a weak student's network (Figure 4.3b), the instructor stated, *I was expecting weak*



(a) Strong Student Network **(b)** Weak Student Network

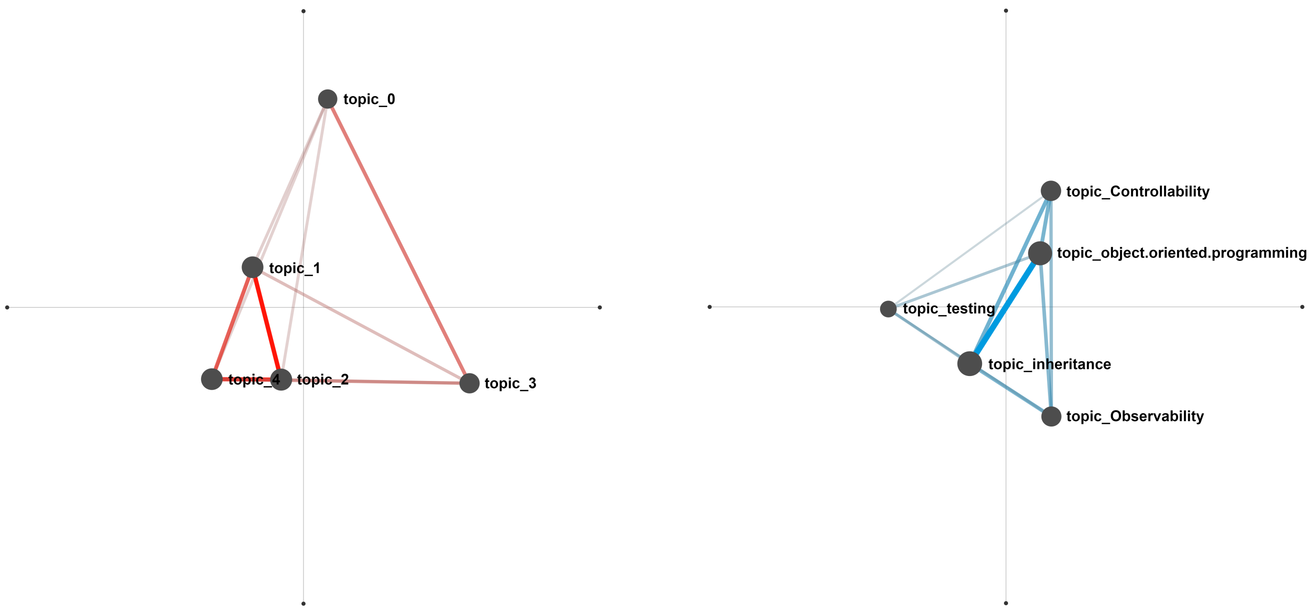
Figure 4.3: Instructor A Individual ENA Visualizations

connections to controllability and observability because they were initially answering a different question from the prompt.

When asked if having these visualizations would impact how they reframe future discussion prompts, Instructor A stated they wouldn't change the prompt but rather better prepare the students in class for the discussion. Instructor A explained that they'd display the group network to the entire class to discuss the topics without connections. The instructor saw this as a helpful tool to then reinforce concepts that were missed without needing to read the entire discussion thus saving time and energy. They also noted that these ENA models would not only be helpful to the instructor but also to the students.

4.3.2 Instructor B

During the presentation, the expert with ENA asked the following questions: *What is LDA?*, *What stanza is being used?*, and *How are the student points being placed?*. When Instructor B began exploring OnDiscuss they were pleased with the initial codebook provided (Table 4.3). This initial codebook was created with more discussion data from various semesters so LDA was able to perform better topic modeling. Since this codebook was used in previous works, the topics had



(a) Initial Network **(b)** Best Network

Figure 4.4: Instructor A Group ENA Visualizations

descriptive names instead of the LDA assigned numbers and it had some added keywords from the same instructor [41].

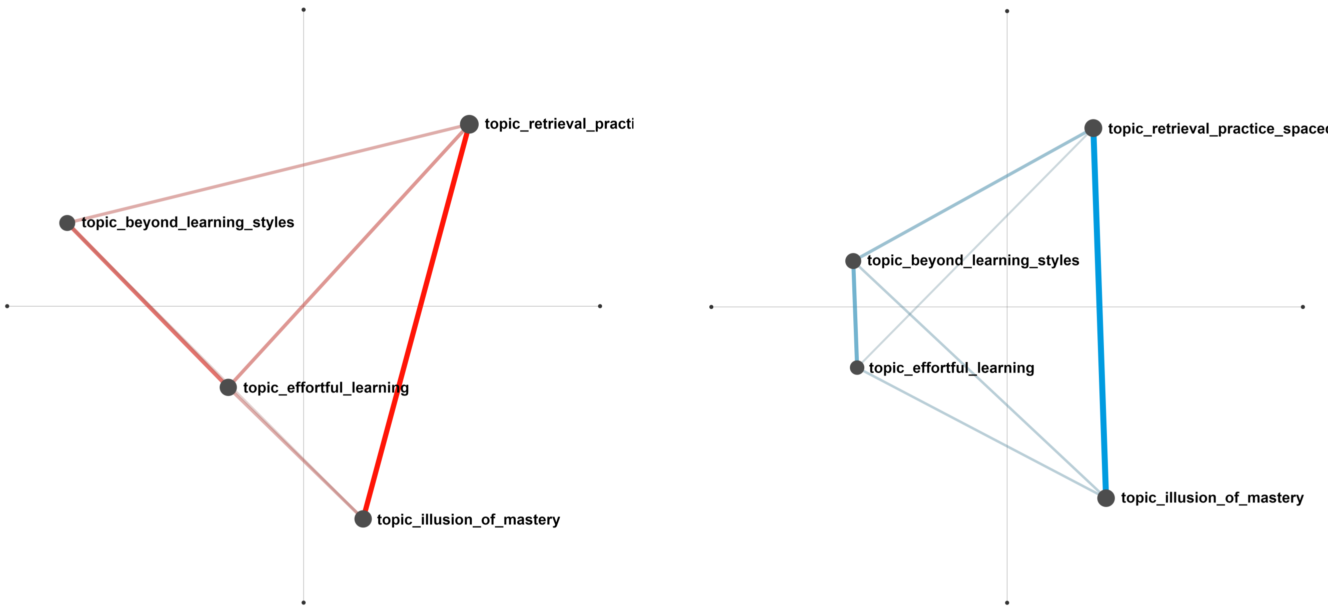
This instructor made very minimal edits to the initial codebook. They only removed "fall" and "desire difficulty" from the "effortful learning" topic. The network model changed as the instructor expected; however, they didn't think much could be concluded from the change between Figure 4.5a and Figure 4.5b. Figure 4.5b had a new connection between "beyond learning styles" and "illusion of mastery" that did not appear in Figure 4.5a. Comparing networks between codebook edits wasn't meaningful to Instructor B. Instead, they wanted a baseline model to compare the resulting network to in order to make more informed edits to the codebook. Instructor B suggested creating a baseline model from a textbook or an asynchronous online discussion amongst the instructors responding to the same discussion prompt as the students.

While examining individual student interactions, the instructor expressed the need for two separate networks: one focusing solely on each student's initial post and another incorporating

Table 4.3: Instructor B Initial Codebook

Code	Keywords
effortful learning	desire, plf, resonate, parachute, land, jump, commun, parachute land, land fall, difficult, difficulties, mistakes, failure, effortful learning, desirable difficulty, desirable, effortful
beyond learning styles	dylexia, learn style, individual, learn differ, disable, intelligent, prefer, support, dyslex, focus, instructional style, learning styles
illusion of mastery	confidence, feedback, calibration, confidence memory, accuracy, peer, answer, event, state, calibration learn, illusion of mastery, illusions of mastery, misunderstanding, illusion of knowing, illusions of knowing, illusion of learning, illusions of learning, re read, cram
retrieval practice spaced out practice interleaving	mass, mass practice, interleaving practice, space retrieval, tend, day, long term, week, myth, practice space, 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 practice, spacing practice, spaced 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, interleaving, interleaved practice, interleave, interleaved

both their original post and subsequent replies. This would provide a more isolated insight into the students' contribution and engagement with other students as well as their initial contribution. Additionally, the instructor suggested that a clearer indication of occurring keywords would be helpful when reading the student's discussion posts in the tool rather than just the plain text.



(a) Initial Network

(b) Best Network

Figure 4.5: Instructor A Group ENA Visualizations

4.4 Discussion

4.4.1 RQ1. Are ENA visualizations of asynchronous online discussion data helpful for instructors who are novices in ENA?

The novice instructor started with no understanding of ENA. Throughout the presentation and the time exploring the tool, this instructor was unsure what made a codebook better in terms of what made a topic a correct grouping of keywords. They were also unsure what makes an ENA network better in terms of the number of nodes, number and thickness of the edges, and placement of the centroid. Since this instructor is in the field of computer science, they seemed more inclined to understand the math behind the ENA visualizations. Future work should be done with novice instructors in other fields to determine if they also have the same inclination. This could inform future revisions to the presentation to emphasize and include more of the mathematical background behind ENA.

Once the instructor created a codebook with keywords they were pleased with, they were able to successfully draw conclusions from the networks. The placement of the points in the ENA network was still confusing to this instructor; however, they were able to extract lots of meaning from the thickness of the edges of both the individual and group networks. They were able to make interpretations such as, *The entire class isn't making many connections to these topics so I should emphasize them more in lectures* and *I thought this student made a strong contribution to the discussion and it's nice to see that visually they made lots of strong connections to all of the topics*. Instructor A stated that once they understood how to create a better codebook they'd feel even more familiar and inclined to use ENA.

The ENA presentation likely played a vital role in the novice instructor's basic understanding of ENA which manifested into their own ENA interpretations. We hypothesize that additional exposure to modifying codebooks and interpreting ENA networks from other discussions would be beneficial for the instructor's continued development. However, through just a single session this instructor demonstrated promising proficiency in ENA, indicating the potential for other instructors unfamiliar with ENA.

4.4.2 RQ2. Do the ENA visualizations have the potential to reduce the time and effort spent assessing asynchronous online discussions for instructors of any experience level with ENA?

Both Instructor A and B reported that having these ENA networks and a modifiable codebook was a helpful supplemental tool for evaluating their asynchronous online discussions. Both said in the semi-structured interview that these supplemental networks would make grading and analyzing the discussions faster. This tool is not meant to fully replace reading through the students' discussion posts but is instead a supplemental tool. The instructors stated it was faster for them to assess the discussions with the visual aid provided by the ENA networks. The instructors also stated by utilizing the individual networks they'd be able to provide more personalized feedback to individual students.

Instructor A found the group and individual networks helpful on their own, while Instructor B thought they were somewhat useful and needed more context, such as a baseline network, to be even more useful. Despite their differing perspectives and experiences with ENA, both instructors were able to draw their own conclusions from the networks based on their own codebooks. This diversity in interpretations highlights the power and flexibility of ENA and OnDiscuss, as instructors not only have the ability to handcraft their codebooks but also to derive nuanced insights tailored to their specific teaching contexts.

4.5 Conclusion and Future Works

In this study, we explored the potential of utilizing OnDiscuss, a tool that enables instructors to edit their own codebook and visualize the resulting ENA visualizations in real time. By providing instructors with the ability to directly interact with and customize, OnDiscuss opens up new opportunities to popularize ENA and make it more accessible to educators who may or may not be familiar with the intricacies of ENA.

Our findings highlight several key insights regarding the implications of this tool for enhancing the accessibility and usability of ENA as a learning analytics visualization tool. OnDiscuss is helpful to those unfamiliar with ENA since it abstracts many of the intricacies of ENA by providing an easy interface to manipulate a codebook and thus the resulting ENA networks. Future refinements, such as the addition of a baseline ENA model, can make it more helpful to those familiar with ENA. Despite the tool's automated keyword generation capabilities, it is clear that instructor intervention remains crucial for refining the codebook. Therefore, while automated techniques like Latent Dirichlet Allocation (LDA) provide valuable insights given a large amount of data, they must be complemented by expert guidance.

Chapter 5

Conclusion

This thesis builds on the foundational work of Moraes et al. [1]. While ENA proved to have potential as a learning analytics visualization, creating a codebook and then coding data is very time consuming. This was significantly streamlined leveraging NLP algorithms like Latent Dirichlet Allocation (LDA). Additionally, we were able to automatically code all of the discussion data based on the codebook.

Furthermore, we improved the accessibility and usability of ENA as a visualization tool for learning analytics. OnDiscuss was designed to allow users of all expertise levels of ENA to visualize and understand discourse in text discussion. By abstracting many of the intricacies of ENA and coding, it offers a more intuitive interface for manipulating the codebook and viewing the resulting ENA networks. While the automated codebook generation is a significant feature, it is essential to complement it with the insights of educators for further refinements. The key takeaway from this is the critical role of instructors' interpretations in making sense of the ENA networks.

The future of learning analytics holds immense potential for driving personalized and adaptive educational experiences. As tools like OnDiscuss continue to evolve, they will empower educators to utilize data more effectively, fostering deeper insights into student learning and engagement. By integrating these analytics with pedagogical expertise, we can create more responsive learning environments that support all learners.

5.1 Limitations

One limitation of our study is that we relied solely on the a priori codebook from EDOD 651 for our initial analyses. While this provided an excellent source of data, it limited our ability to explore a diverse range of perspectives and contexts that could emerge from other courses with different codebooks. Ideally, incorporating additional courses and their corresponding codebooks would enrich the data validation process and enhance the robustness of our findings.

By having only one codebook, we constrained our validation to a single perspective, which may not fully capture the complexity of discourse across various learning environments. This lack of diversity in coding frameworks may affect the generalizability of our results, as different courses often emphasize unique learning outcomes and instructional approaches. Future research would benefit from using multiple codebooks from a variety of courses, allowing for a more comprehensive validation process and broader insights into the application of ENA across different educational contexts.

One current limitation of our coding scheme and ENA networks are that they only represent the raw text data within the discussions. Consequently, this may overlook other forms of media, such as images, videos, and links to external websites, which could provide valuable insights or perspectives that are not captured through text alone.

Another potential limitation concerns the study population of Chapter 4. The instructors were purposefully chosen by us to participate in the research project. While this approach was necessary to ensure access to relevant discussion data and expertise with ENA, it introduced a selection bias. Also although the participants in our study are drawn from the same university, it is important to note that they exclusively represent graduate level courses with relatively small class sizes in two different disciplines: computer science and education.

Also in Chapter 4 Instructor A had considerably less available discussion data than Instructor B to create the codebook. Instructor A had 363 posts with 37,254 total words, while Instructor B had 2,648 posts with 444,364 total words. Because Instructor B had more data, LDA was able to provide more accurate topic modeling [42]. Moreover, given that Instructor B's codebook has been utilized in previous studies, it is plausible that the codebook may have undergone refinements over time, potentially rendering it more polished compared to Instructor A's.

5.2 Future Works

5.2.1 Enhancements to OnDiscuss

In Chapter 4 the expert in ENA (Instructor B) provided lots of future improvements to OnDiscuss during the semi structured interview. Adding a baseline ENA model from either the textbook or discussion posts of the instructor could have potential to assist the instructor even further by allowing them to make comparisons between the class and base model. Another addition to the tool could be two separate networks for the individual students: one based on the student's initial post responding to the discussion prompt and then another based on both their original post and their subsequent replies to other classmates. Because the codebook is comprised of many keywords contributing to a code, the tool could also highlight the occurring keywords in the student's discussion post.

5.2.2 Study Population Diversity

Further exploration is needed across different academic disciplines from Computer Science and Education. While the two disciplines examined in Chapter 4 are vastly different there are plenty of other disciplines to investigate. All of the data was from graduate level courses so future studies should explore undergraduate courses. In many undergraduate courses the class sizes are much larger which would be even more time consuming to assess those discussion assignments. Exploring larger class sizes would also be important.

5.2.3 Different Study Methodologies

The study in Chapter 4 was carried out on historical discussion data, underscoring the necessity for future studies to be done on an in progress semester of discussions to uncover further insights into the implications of utilizing OnDiscuss. An impactful approach would be to design an empirical study involving multiple instructors from various disciplines who utilize OnDiscuss to evaluate their online discussion assignments, incorporating qualitative data from surveys and interviews. Additionally, conducting an experimental study that measures quantitative variables, such as the time taken to evaluate student contributions, would provide valuable insights. Another promising direction is a longitudinal study that tracks instructors who do not use OnDiscuss in one semester, followed by multiple subsequent semesters using the tool. This would allow for a comparative analysis of the differences in their pedagogy and class structures, as well as any changes in their interpretations and uses of the ENA networks as they become more familiar with the tool and its capabilities for enhancing student engagement and learning outcomes.

5.2.4 Other Discourse Types

OnDiscuss was primarily intended to be used on asynchronous online discussion data from students responding to prompts that required them to assimilate their knowledge. Expanding to other types of discourse in classes such as group discussions about group projects in messaging platforms like Microsoft Teams or Slack could also enrich an instructor's understanding of the collaborative learning taking place in group projects. This would allow for a broader application of OnDiscuss and further enhance its capabilities as a learning analytics tool.

Bibliography

[1] Marcia Moraes, James Folkestad, and Kelly McKenna. Using epistemic network analysis to help instructors evaluate asynchronous online discussions. In *Second International Conference on Quantitative Ethnography: Conference Proceedings Supplement*, pages 19–22, 2021.

[2] J. Broadbent and W.L. Poon. Self-regulated learning strategies academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education*, 27:1–13, 2015.

[3] Olga Viberg, Mathias Hatakka, Olof Bälter, and Anna Mavroudi. The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89:98–110, 2018.

[4] 1st international conference on learning analytics and knowledge. In *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*.

[5] L. S. Vygotsky. *Mind in society*. Harvard University Press, Cambridge, MA, 1978.

[6] Richard S Prawat. Teachers' beliefs about teaching and learning: A constructivist perspective. *American journal of education*, 100(3):354–395, 1992.

[7] Michelle L Davis, Sara M Witcraft, Scarlett O Baird, and Jasper AJ Smits. Learning principles in cbt. In *The science of cognitive behavioral therapy*, pages 51–76. Elsevier, 2017.

[8] Maya Aloni and Christine Harrington. Research based practices for improving the effectiveness of asynchronous online discussion boards. *Scholarship of Teaching and Learning in Psychology*, 4(4):271, 2018.

[9] Sarah Fehrman and Sunnie Lee Watson. A systematic review of asynchronous online discussions in online higher education. *American Journal of Distance Education*, 35(3):200–213, 2021.

[10] Jessica Decker and Valerie Beltran. Graduate students' perceptions of the benefits and drawbacks of online discussion tools. *International Journal of Online Pedagogy and Course Design (IJOPCD)*, 6(1):1–12, 2016.

[11] Majed Gharmallah Alzahrani. The effect of using online discussion forums on students' learning. *Turkish Online Journal of Educational Technology-TOJET*, 16(1):164–176, 2017.

[12] Dhanielly PR de Lima, Marco A Gerosa, Tayana U Conte, and José Francisco de M. Netto. What to expect, and how to improve online discussion forums: the instructors' perspective. *Journal of Internet Services and Applications*, 10:1–15, 2019.

[13] David Williamson Shaffer, Wesley Collier, and Andrew R Ruis. 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.

[14] Barney Glaser and Anselm Strauss. *Discovery of Grounded Theory: Strategies for Qualitative Research*. Routledge, 1st edition, 1999.

[15] Johnny M. Saldana. *The Coding Manual for Qualitative Researchers*. SAGE Publications, London, England, 2 edition, 2013.

[16] David Shaffer. *Epistemic network analysis: Understanding learning by using big data for thick description*, pages 520–531. 2018.

[17] Cody L Marquart, Zachari Swiecki, Brendan Eagan, and David Williamson Shaffer. Package 'ncodeR', 2019. Online. [Accessed: May 18, 2022].

[18] Sina Mahdipour Saravani, Sadaf Ghaffari, Yanye Luther, James Folkestad, and Marcia Moraes. 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*, pages 227–238. Springer, 2023.

[19] E. Mayfield, D. Adamson, and C. P. Rosé. Lightside researcher's workbench, 2013.

[20] Zhiqiang Cai, Amanda Siebert-Evenstone, Brendan Eagan, and David Williamson Shaffer. Using topic modeling for code discovery in large scale text data. In *International Conference on Quantitative Ethnography*, pages 18–31. Springer, 2021.

[21] J. Choi, A. R. Ruis, Z. Cai, B. R. Eagan, and D. W. Shaffer. 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*.

[22] Zhiqiang Cai, Amanda Siebert-Evenstone, Brendan Eagan, David Williamson Shaffer, Xiangen Hu, and Arthur C Graesser. ncoder+: a semantic tool for improving recall of ncoder coding. In *International Conference on Quantitative Ethnography*, pages 41–54. Springer, 2019.

[23] David Newman, Jey Han Lau, Karl Grieser, and Timothy Baldwin. Automatic evaluation of topic coherence. In *Human language technologies: The 2010 annual conference of the North American chapter of the association for computational linguistics*, pages 100–108, 2010.

[24] Szilvia Zörgő, Anna Jeney, Krisztina Csajbók-Veres, Samvel Mkhitaryan, and Anna Susánszky. Mapping the content structure of online diabetes support group activity on facebook. In *International Conference on Quantitative Ethnography*, pages 221–236. Springer, 2021.

[25] Yuanru Tan, Binrui Yang, Brendan Eagan, and Peter Levine. Using ordered network analysis to visualize ideologies in political survey data. In *Proceedings of the International Conference on Quantitative Ethnography*, October 2022. Poster.

[26] Xiner Liu Juhan Kim, Amanda Barany and Andres Felipe Zambrano. The stories we tell: Uncovering hidden narratives in history textbooks through epistemic network analysis. In G. Arastoopour Irgens and S. Knight, editors, *Advances in Quantitative Ethnography: Fifth International Conference, ICQE 2023*, pages 261–274, Melbourne, Australia, 2023. Springer / International Society for Quantitative Ethnography (ISQE).

[27] Szilárd Dávid Kovács, Katherine Mulholland, Lara Condon, Zsuzsa Koncz, and Szilvia Zörgö. Interaction of diagnostic criteria in the narratives of patients with borderline personality disorder. In G. Arastoopour Irgens and S. Knight, editors, *Advances in Quantitative Ethnography: Fifth International Conference, ICQE 2023*, pages 316–329. Springer / International Society for Quantitative Ethnography (ISQE), Melbourne, Australia, 2023.

[28] Denise M. Bressler, Leonard A. Annetta, Alexis Dunekack, Richard L. Lamb, and David B. Vallett. How stem game design participants discuss their project goals and their success differently. In Barbara Wasson and Szilvia Zörgő, editors, *Advances in Quantitative Ethnography*, pages 176–190, Cham, 2022. Springer International Publishing.

[29] Ha Nguyen. Exploring group discussion with conversational agents using epistemic network analysis. In Barbara Wasson and Szilvia Zörgő, editors, *Advances in Quantitative Ethnography*, pages 378–394, Cham, 2022. Springer International Publishing.

[30] Mamta Shah, Amanda Siebert-Evenstone, Hollie Moots, and Brendan Eagan. Quality and safety education for nursing (qsen) in virtual reality simulations: A quantitative ethnographic examination. In Barbara Wasson and Szilvia Zörgő, editors, *Advances in Quantitative Ethnography*, pages 237–252, Cham, 2022. Springer International Publishing.

[31] S. B. Shum. Qe visualizations as tools for thinking [powerpoint slides], international conference on quantitative ethnography 2021 keynote speaker, 2021.

[32] Gloria Milena Fernandez-Nieto, Roberto Martinez-Maldonado, Kirsty Kitto, and Simon Buckingham Shum. Modelling spatial behaviours in clinical team simulations using epistemic network analysis: Methodology and teacher evaluation. In *LAK21: 11th International Learning Analytics and Knowledge Conference*, LAK21, page 386–396, New York, NY, USA, 2021. Association for Computing Machinery.

[33] Tiffany Herder, Zachari Swiecki, Simon Skov Fougt, Andreas Lindenskov Tamborg, Benjamin Brink Allsopp, David Williamson Shaffer, and Morten Misfeldt. Supporting teachers'

intervention in students' virtual collaboration using a network based model. In *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, LAK '18, page 21–25, New York, NY, USA, 2018. Association for Computing Machinery.

[34] Zachari Swiecki, AR Ruis, and David Williamson Shaffer. Modeling and visualizing team performance using epistemic net-work analysis. In *Proceedings of the 7th annual GIFT users symposium*, pages 148–156, 2019.

[35] Hazel Vega and Golnaz Arastoopour Irgens. Constructing interpretations with participants through epistemic network analysis: Towards participatory approaches in quantitative ethnography. In Barbara Wasson and Szilvia Zörgő, editors, *Advances in Quantitative Ethnography*, pages 3–16, Cham, 2022. Springer International Publishing.

[36] Vitor Rolim, Rafael Ferreira, Rafael Dueire Lins, and Dragan Găsević. A network-based analytic approach to uncovering the relationship between social and cognitive presences in communities of inquiry. *The Internet and Higher Education*, 42:53–65, 2019.

[37] Dragan Gašević, Srećko Joksimović, Brendan R Eagan, and David Williamson Shaffer. Sens: Network analytics to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, 92:562–577, 2019.

[38] Zachari Swiecki and David Williamson Shaffer. isens: an integrated approach to combining epistemic and social network analyses. In *Proceedings of the tenth international conference on learning analytics & knowledge*, pages 305–313, 2020.

[39] Jennifer Scianna and Rogers Kaliisa. 'sseen': a networked approach to uncover connections between sentiment, social, and epistemic elements of student online forum discourse. *Educational technology research and development*, pages 1–23, 2023.

[40] Simon Skov Fougt, Amanda Siebert-Evenstone, Brendan Eagan, Sara Tabatabai, and Morten Misfeldt. Epistemic network analysis of students' longer written assignments as forma-

tive/summative evaluation. LAK '18, page 126–130, New York, NY, USA, 2018. Association for Computing Machinery.

[41] Marcia Moraes, Sadaf Ghaffari, Yanye Luther, and James Folkesdtad. Combining automatic coding and instructor input to generate ena visualizations for asynchronous online discussion. In Golnaz Arastoopour Irgens and Simon Knight, editors, *Advances in Quantitative Ethnography*, pages 381–394, Cham, 2023. Springer Nature Switzerland.

[42] David M Blei, Andrew Y Ng, and Michael I Jordan. Latent dirichlet allocation. *Journal of machine Learning research*, 3(Jan):993–1022, 2003.

[43] Ena web tool, 2023.

[44] Sadaf Ghaffari and Nikhil Krishnaswamy. Grounding and distinguishing conceptual vocabulary through similarity learning in embodied simulations. *arXiv preprint arXiv:2305.13668*, 2023.

[45] rena package, 2024.

[46] Ena web tool, 2024.

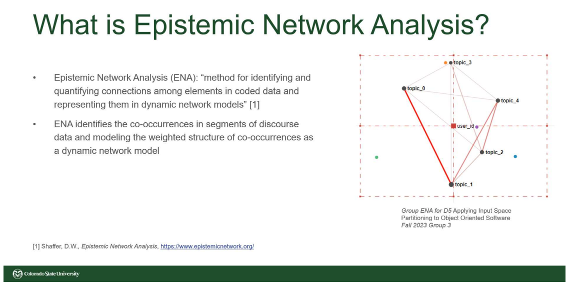
[47] Fred Pyrczak and Deborah M. The empirical approach to knowledge. In *Making Sense of Statistics*, pages 1–5. Routledge, 2018.

[48] Danielle Magaldi and Matthew Berler. Semi-structured interviews. *Encyclopedia of personality and individual differences*, pages 4825–4830, 2020.

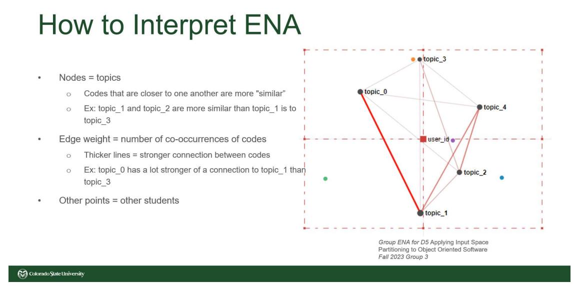
[49] D.W. Shaffer. *Quantitative Ethnography*. Cathcart Press, Madison, WI, 2017.

Appendix A

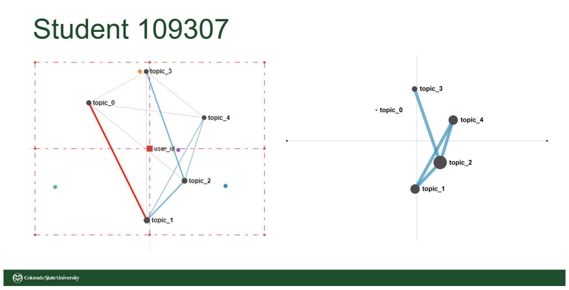
ENA Presentation



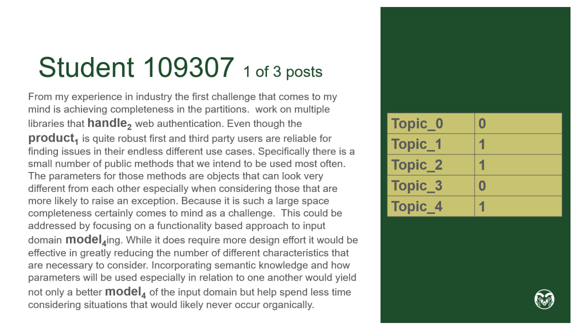
(a) What is ENA



(b) How to Interpret ENA



(c) Example Student Network



(d) Example Student Post

Figure A.1: Sample Slides from ENA Presentation

Appendix B

Interview Questions

- Do these automatically generated topics reflect your learning objectives for this discussion?
  - After revising the codebook do you think these new topics/keywords better reflect your learning objectives?
  - Did the autogenerated topics reflect the topics/concepts students should be discussing?
  - Did the autogenerated topics help provide a foundation for a codebook?
  - How many iterations of the codebook did you make?
    - How many iterations of the codebook did you make?
    - Why?
  - If you couldn't edit the codebook, would you still find it useful?
- When investigating the group ENA (entire class)
  - Do you have a sense of the class discussion?
  - What topics are students discussing?
  - Are these topics the ones you are expecting students to be discussing?
- When investigating the individual ENA (single student)
  - Do these ENA models provide useful context?
  - Can you easily identify what topics the student discussed? What topics are they missing?
  - Can you easily identify what connections among the topics the student did?
  - Does this influence how you'd grade this student?

- Do you feel grading and analyzing these discussions is faster with the supplemental ENA models?
- Does this allow you to provide more personalized feedback to the student?

- Would this impact how you frame future discussion assignments?
  - If students are not discussing all the topics, would you reframe/revise the way you describe the prompt?
  - Based on the ENA visualizations topics and connections, would you consider revising your mediation process?
  - Before the discussion assignment closed, would you intervene in the discussion if students were not making certain connections?