




Article

Evaluating the Coverage and Depth of Latent Dirichlet Allocation Topic Model in Comparison with Human Coding of Qualitative Data: The Case of Education Research

Gaurav Nanda ¹, Aparajita Jaiswal ², Hugo Castellanos ³, Yuzhe Zhou ³, Alex Choi ³ and Alejandra J. Magana ^{3,*} 

¹ School of Engineering Technology Purdue University, West Lafayette, IN 47906, USA

² Center for Intercultural Learning, Mentorship, Assessment and Research (CILMAR), Purdue University, West Lafayette, IN 47906, USA

³ Department of Computer and Information Technology, Purdue University, West Lafayette, IN 47906, USA

* Correspondence: admagana@purdue.edu

Abstract: Fields in the social sciences, such as education research, have started to expand the use of computer-based research methods to supplement traditional research approaches. Natural language processing techniques, such as topic modeling, may support qualitative data analysis by providing early categories that researchers may interpret and refine. This study contributes to this body of research and answers the following research questions: (RQ1) What is the relative coverage of the latent Dirichlet allocation (LDA) topic model and human coding in terms of the breadth of the topics/themes extracted from the text collection? (RQ2) What is the relative depth or level of detail among identified topics using LDA topic models and human coding approaches? A dataset of student reflections was qualitatively analyzed using LDA topic modeling and human coding approaches, and the results were compared. The findings suggest that topic models can provide reliable coverage and depth of themes present in a textual collection comparable to human coding but require manual interpretation of topics. The breadth and depth of human coding output is heavily dependent on the expertise of coders and the size of the collection; these factors are better handled in the topic modeling approach.

Keywords: topic modeling; latent Dirichlet allocation; qualitative analysis; human coding; natural language processing; unsupervised machine learning



Citation: Nanda, G.; Jaiswal, A.; Castellanos, H.; Zhou, Y.; Choi, A.; Magana, A.J. Evaluating the Coverage and Depth of Latent Dirichlet Allocation Topic Model in Comparison with Human Coding of Qualitative Data: The Case of Education Research. *Mach. Learn. Knowl. Extr.* **2023**, *5*, 473–490.

<https://doi.org/10.3390/make5020029>

Received: 13 March 2023

Revised: 22 April 2023

Accepted: 10 May 2023

Published: 14 May 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Social science fields, such as education research, have started to expand the use of computer-based research methods to supplement traditional approaches. This shift is motivated by the availability of large datasets and the emergence of multimodal data. Indeed, new subfields have emerged, such as educational data mining, learning analytics, and computational ethnography. Specifically for the case of qualitative methods, researchers have argued that computer-based analytical methods can support the process and analysis of qualitative data. Specifically, computer-based analytical methods can support the exploration (visualization), classification, grouping, and validation of data and patterns derived from data [1]. Other methods involving natural language processing techniques may support qualitative data analysis by providing early categories that researchers may interpret and refine. Likewise, initial qualitative data analysis may improve topic modeling outcomes when used as training data [2].

Our study focuses on a specific natural language processing approach for supplementing qualitative data analysis called topic modeling. Topic modeling is an unsupervised machine learning technique that detects word and phrase patterns within text data and automatically clusters word groups and similar expressions into “underlying topics” that

best characterize the overall qualitative dataset [3]. Specifically, we focus on a specific technique called the latent Dirichlet allocation topic model (LDA) [4]. Topic modeling has been used in various domains for analyzing large textual collections, such as in the education domain for analyzing student survey responses [5] and peer comments [6], in social media for sentiment analysis of tweets [7], in legal research for document summarization [8] and finding similar precedent cases [9], and in the healthcare domain for analyzing electronic health records [10].

Past studies have mostly used either manual qualitative coding (also referred to as human coding later in the paper) [11,12] or computational approaches, such as topic modeling [5], for performing qualitative analysis. They have not compared the results of the LDA topic model and human coding analysis and how well they perform in terms of capturing nuanced patterns and topics present in the dataset. This study aims to fill this gap in the literature and aims to answer the following research questions: (RQ1) What is the relative coverage of the LDA topic model and human coding in terms of the breadth of the topics/themes extracted from the text collection? (RQ2) What is the relative depth or level of detail among identified topics using LDA topic models and human coding approaches? The dataset used in this study includes student reflections aimed at developing cultural self-awareness in the context of teamwork, in preparation for a semester-long project. The overall approach to analysis is shown in the schematic diagram in Figure 1.

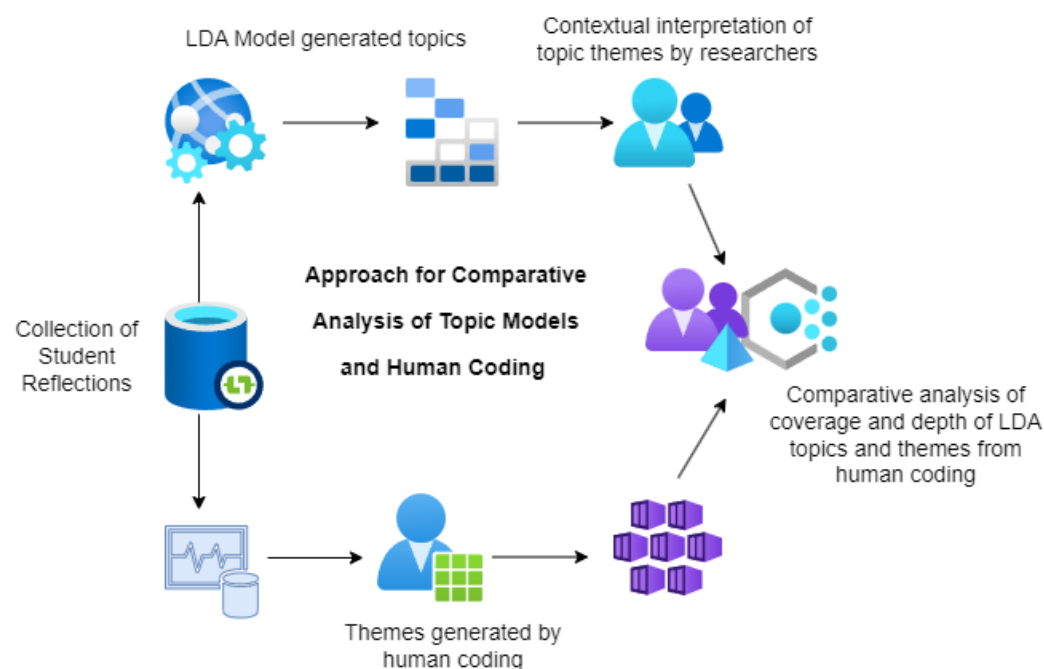


Figure 1. Overview of the comparative analysis between the LDA topic model and human coding.

2. Overview of Traditional Research Methods in Education

Education research and other disciplines in the social sciences rely on qualitative and quantitative research methods. However, differences exist regarding their philosophical groundings, which translate into differences in their goals, assumptions, and specific methods for data collection, analysis, and interpretation. For instance, while quantitative research focuses on hypothesis testing and confirmation, qualitative research focuses on hypothesis generation and understanding [13]. For these purposes, each approach uses specific types of data and analytical procedures. Quantitative research often seeks to validate an idea or theory by performing experiments and analyzing the results numerically. This approach is considered to be objective and reliable, as researchers are detached from the subject of investigation. Qualitative research often seeks to build explanations based on observations, documents, interviews, or other text or visual data. It is often used to understand thoughts or experiences with the goal of gathering in-depth insights into

social reality [12]. This approach is considered to be somewhat subjective, as the cultural meaning is central to the interpretation of the findings, which may be socially constrained. Additionally, the researcher is heavily involved in the process of data collection, data analysis, and interpretation of the findings.

The analytical procedures followed by each approach are relevant to this study. Quantitative analysis focuses on numerical data, where statistical analysis, including descriptive and inferential methods, is central to this process. Quantitative analysis tends to be more linear and straightforward [13]. Qualitative analysis, on the other hand, focuses on the emergence of conceptual categories and descriptive themes. Specifically, qualitative research methods enable thorough scrutiny of the researched topic, which is not possible in quantitative research. The thoroughness is accomplished by processes involving data reduction through detailed coding processes (herein, human coding), searching for meaning in the data by looking at them in divergent ways, and delving into the data to identify patterns and themes [12]. Qualitative analysis is highly iterative, as opposed to quantitative analysis, which is more linear. Thus, qualitative research is more time-consuming, and because of that, it is often performed with smaller samples of data than quantitative studies. Furthermore, due to the researcher's central role in the data analysis process, qualitative research tends to be seen with less rigor as insights and intuitions from the researchers are allowed a free hand.

To overcome the challenges and limitations of quantitative and qualitative research methods, researchers have consistently made the case that both methods are complementary [14]. By combining the two methods, the research findings can be strengthened with triangulation. Thus, combining methods is recommended when multiple sources of data can be taken into consideration. Thus, our stand supports the combination of qualitative and quantitative research methods whenever possible, especially in the context of education research, where rich datasets can be considered as learners interact with instructors, peers, and with technology.

However, there are occasions in which only qualitative data are available, which may pose some challenges to the researchers. As we have argued above, qualitative data are difficult to analyze, but the processes of making sense of qualitative reporting findings is also difficult. This is in part due to the complexity and richness of the data. The researcher scrutinizes the data with the goal of obtaining an in-depth understanding of the experience or phenomenon. However, that level of scrutiny is lost in the process of summarizing the findings as data need to be aggregated back to report findings, often in the form of themes. In these situations, computer-based methods can supplement, extend, or strengthen qualitative research methods. For instance, to overcome limitations in data reduction and reporting of qualitative findings, researchers have argued that computer-based analytical methods can support the process of further analyzing, interpreting, and reporting qualitative data [15]. Specifically, computer-based analytical methods can support the exploration (visualization), classification, grouping, and validation of data and patterns derived from data. These methods are highly effective only when the qualitative data have been quantified, often with the use of a rubric.

However, there are other methods that can support researchers in the analysis and validation of qualitative analysis. These methods involve natural language processing techniques that can support qualitative data analysis by: (a) Providing early categories that researchers may interpret and refine; and (b) Supporting the validation of the findings of topic modeling outcomes when used as training data [2]. The benefits of doing so are two-fold. First, it may allow researchers to increase the sample size of the qualitative data to something more robust, therefore increasing the generalizability of the findings. Second, it can help increase the validity of the coding process as an addition to the manual identification of codes and categories; researchers can create separate categorizations using techniques such as topic modeling. This, as a result, may increase the reliability of the analysis leading to a more trustworthy process of interpreting the findings. The following section will overview machine learning approaches for analyzing qualitative data. For the

rest of our study, we will focus on a specific natural language processing method called topic modeling. With this focus, we will also address issues associated with topic modeling, including the interpretation of themes, along with our rationale for comparing coding results via topic modeling and human coding.

3. Machine Learning and Natural Language Processing Approaches for Analyzing Qualitative Data

One of the main challenges associated with the analysis of textual data using manual qualitative coding is that it is very time- and effort-consuming [16]. Machine learning (ML) and natural language processing (NLP)-based approaches have been used to efficiently perform qualitative analysis on textual data. Some of the approaches that have been used by previous studies include: (a) Developing expert-designed rules for identifying phrases in the text that are indicative of a certain category and using NLP approaches to parse the text and detect these phrases [17]; (b) Using supervised ML approaches to learn from historically manually coded data to predict the codes or themes for textual data [18]; (c) Using unsupervised ML approaches to extract patterns from the textual data and then qualitatively analyze these patterns to determine if they are representative of any relevant themes [18,19].

While these approaches have been found to be effective in some of the previous studies, there are some challenges associated with ML-based approaches, particularly supervised ML, as discussed here. First, the supervised ML techniques are dependent on the size of the categories in the dataset, meaning they may be able to extract broad and prominent patterns from the data present in a large number of documents in the collection but may not be able to accurately identify smaller categories with a low number of cases. From the perspective of qualitative analysis, the smaller categories are also important, and the frequency of occurrence in the dataset is not the primary measure of the significance of that code/category [16]. Second, the supervised ML models that are trained on historical human-coded data and used to predict codes on a new qualitative dataset may not be very accurate if the nature of the text and frequency distribution of codes in the dataset is not very similar to the training dataset [20]. Third, the explainability and reliability of most supervised ML models are not very good as they typically function as a black box and do not offer explanations for their predictions, and the results may not be readily interpretable, which impacts the trustworthiness of the qualitative researchers [16,20].

The unsupervised ML approaches try to identify existing patterns in the data that may or may not be meaningful from a qualitative analysis perspective. Topic modeling is an unsupervised approach that generates underlying topics present in a textual collection of data, such as student responses, documents, news articles, and discussion forum posts. These topics are generated statistically based on co-occurrences of words in different documents in the collection and may or may not be interpretable. Therefore, these topics that are generated by the model need to be qualitatively interpreted to determine if they represent a coherent theme. Some of the previous studies [20,21] have compared the similarities and differences between the ML-based approaches, such as topic models, and traditional approaches of manual qualitative analysis, such as grounded theory, and some studies have also proposed an integrated framework combining the two approaches [22].

4. Methods

4.1. Context and Participants

The study was conducted in a sophomore-level “Systems Analysis and Design” course. The course provides an overview of the approaches used by today’s information system developers to discover and model the requirements of a system, and then construct and prototype an acceptable design to implement a successful system solution. During the course, the students were required to complete four major assignments, known as milestones, and a final capstone project. The students worked in teams to complete all four milestones and also for their final project.

To facilitate teamwork interaction, the course implemented transformative pedagogy as a reflective approach to promote intercultural self-awareness and its potential consequences in the context of teamwork. Specifically, transformative pedagogy allows students “to examine their assumptions critically, grapple with social issues, and engage in social action” [23]. Students apply transformative pedagogy by reflecting on their experience and putting this into action in the context of the team. Students reflect on how their cultural background may influence their communication with the team. The participants of this study were 127 sophomore-level students enrolled in a systems analysis and design course in the Spring of 2021. The class consisted of 20 female students and 107 male students. Additionally, the majority of students enrolled in this course is second-year ($n = 51$) and third-year ($n = 47$) students, followed by fourth-year students ($n = 28$) and only one first-year student.

4.2. Procedures and Data Collection Method

During Week 4 of the semester, students engaged in an activity aimed at noticing the role of their culture in their teamwork interactions. As a part of this class activity, the students were made aware of cross-cultural communication styles, the concept of power distance, and different decision-making styles. Further, the students were asked to watch a video that helped them develop their cultural self-awareness. After watching the video, the students were also engaged in reflection activities by answering the following reflection questions. Reflection Question 1 (R1): Thinking about communication styles within your own teamwork, how do you think the role of culture has influenced this process within your team? Reflection Question 2 (R2): How do you think your own cultural background may influence your teamwork interaction? All the students ($n = 127$) responded to the reflection question, and the responses of the students served as the data for the study.

4.3. Approach Used for Traditional/Manual Coding

Figure 2 depicts the steps we followed for the traditional manual coding of the data (left diagram) and the computer-based approach (right diagram). In the following sections, we describe each of these two approaches in detail.

The human coding of the data was initially performed by two independent raters on 40% of the 127 reflection responses. For this, the raters performed open coding [12], which is a traditional form of manual data coding. Open coding refers to the process of “labeling concepts, defining and developing categories based on their properties and dimensions.” [24]. After performing open coding separately, the two raters met, discussed their codes, and created a codebook. A total of 40% of the data were then re-coded along with the coding of the other 60% of the data by the two raters based on the codebook developed. The raters met again to discuss their codes and calculated the interrater reliability, which was 80%. Table 1 below represents an example of the manual coding process. First, the raters came up with initial codes; in the next steps, the codes were combined to create the categories referred to as the final code in this manuscript. Lastly, a definition was added to describe the final code.

Table 1. Example of manual coding process.

Initial Codes	Final Code	Definition
Opportunity to communicate Everyone can share ideas Freedom to express	Equal chances to speak	Everyone has equal chances to express their own ideas and thoughts.

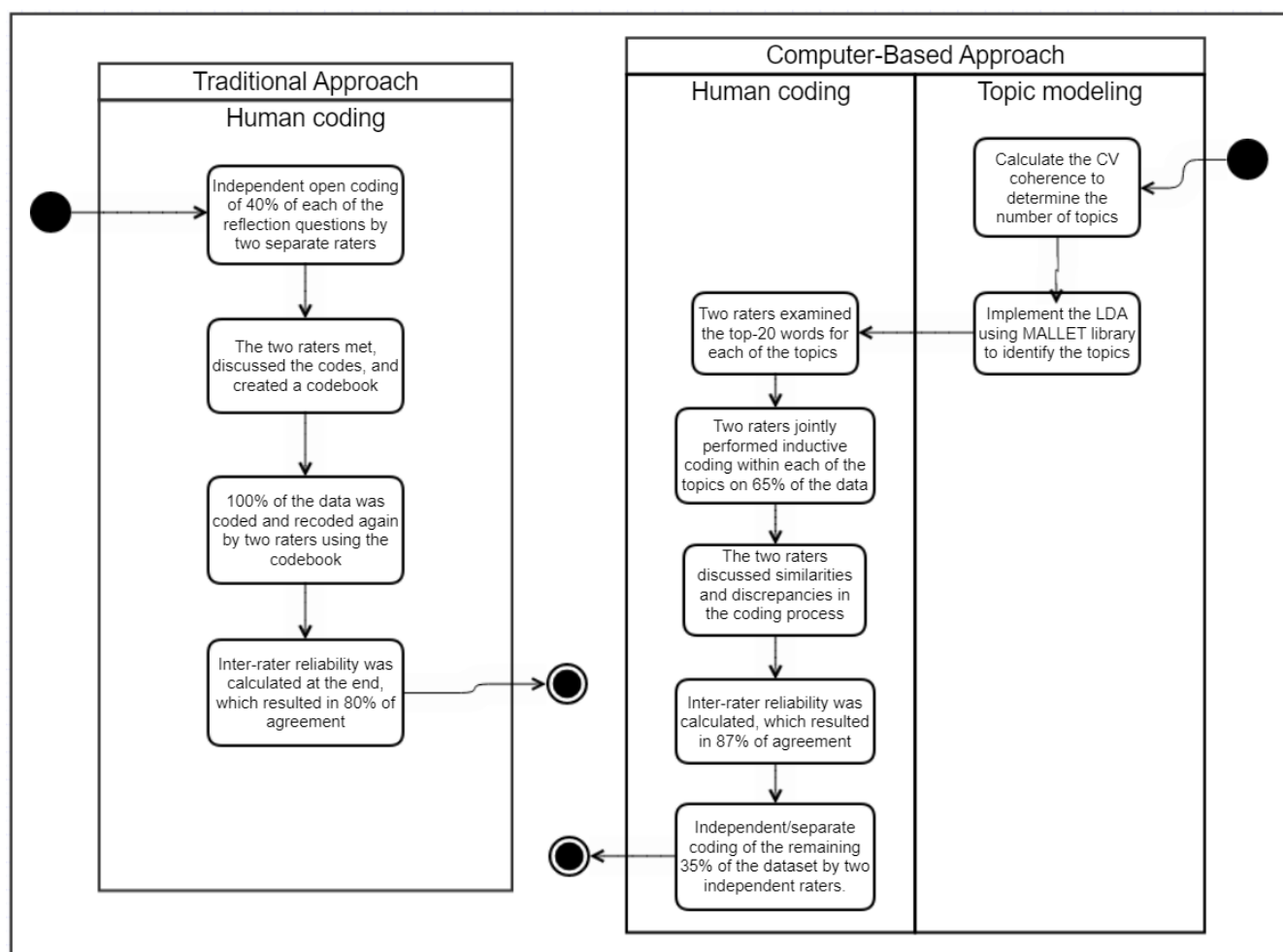


Figure 2. Steps followed for the traditional approach (left diagram) and the computer-based approach (right diagram).

4.4. Approach Used for LDA Topic Modeling

For the computer-based approach of coding the data, we used LDA, the most commonly used topic model, for determining the prominent topics from the collection of student reflections. LDA is an unsupervised generative probabilistic method for modeling a textual collection, which models each document in the collection as a mixture (probabilistic distribution) of a given number of underlying topics, and each topic is modeled as a mixture (probabilistic distribution) of words in the collection [2,25–27].

The LDA topic model requires the number of topics as an input variable. To determine the optimum number of topics for the collection, the measure CV (content vector) coherence was used as it has been found to be well-correlated with human judgment by other studies [27–29]. The CV coherence measure uses content vectors' representation of words based on their co-occurrences and coherence scores, which are calculated using normalized pointwise mutual information (NPMI) and cosine similarity. Some of the other measures used to evaluate the optimum number of topics include perplexity, UMass coherence, and UCI coherence [27,29]. Similar to the approach used in previous studies [5], the CV coherence value for the different input topics was calculated for the dataset (collection of reflections) using the PyLDAvis Python library [30], and the number of topics associated with the maximum CV coherence value was selected as the optimum number of topics for the collection. In this study, two collections were analyzed, a collection of reflections corresponding to R1 and a collection of reflections for R2. Separate LDA topic models were developed for each collection.

The LDA topic model was implemented in this study using the MALLET (MACHINE Learning for Language Toolkit) library [31], as it is computationally efficient, includes implementation of Gibbs sampling with document–topic hyperparameter optimization, and provides output in a format that is suitable for conducting post hoc analysis of the generated topics to determine the theme associated with each topic [3]. The MALLET LDA topic model provides the following output: (a) The list of topics with associated weights indicating how prominently that topic is present in the entire collection; (b) The list of top 20 words representing the topic arranged in decreasing order of weightage/importance; (c) Proportion of topics present in each document in the collection represented by the topic weight associated with that document. These outputs were analyzed qualitatively to determine the theme of each topic generated by the LDA model. Following a similar approach used in previous studies [5], the initial theme of the topic was determined by examining the top 20 words associated with the topic (as outputted by MALLET) in the context of the question asked to the students (Questions 1 or 2), and then the theme was refined after reading the top 10 documents in which this topic had the strongest presence, as indicated by the topic weight. This second step in the topic theme determination process was accomplished by sorting the documents in descending order of weights associated with each topic using the MALLET output (c) mentioned above and then examining the top 10 documents. We chose the top 10 documents for topic theme interpretation as the total size of the dataset was 127 and the representation of the topic in the document started decreasing as the weight of the topic decreased.

The themes determined for the LDA-generated topics for the two collections (R1 and R2 reflections) were then compared with the codes obtained from manual coding analysis. The optimum number of topics was determined for each of the collections (reflections for R1 and reflections for R2) using the CV coherence approach. The CV coherence values calculated for a different number of input topics using the PyLDAvis package are shown in Figure 3.

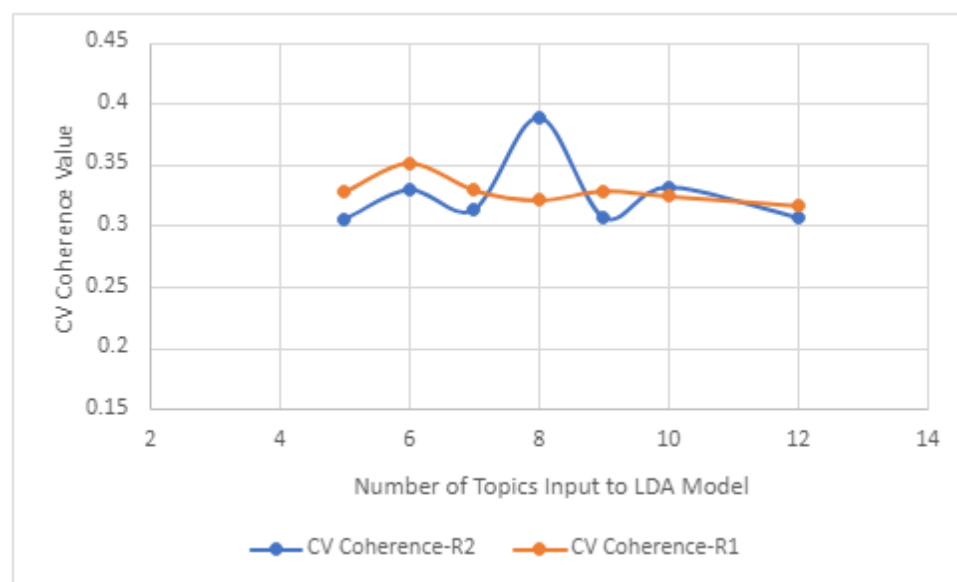


Figure 3. CV coherence values corresponding to the number of topics for LDA models R1 and R2.

As shown in Figure 3, the peak value of CV coherence for the collection of responses R1 corresponds to the number of topics = 6, and for collection R2, the peak value of CV coherence happens at the number of topics = 8. The number of topics corresponding to the highest CV coherence value is considered to be optimal [27,28]. Therefore, these numbers of topics were used as input for developing the LDA topic models for collections R1 (optimal number of topics = 6) and R2 (optimal number of topics = 8). The topics generated from

the LDA topic models were then analyzed qualitatively using the approach mentioned in Section 4.

Once the topics were generated automatically for each of the reflection questions, two raters jointly performed inductive coding on 65% of the data and for each of the topics. The raters performed independent coding for Topics 1, 2, 3, and 4 for Reflection Question 1, and Topics 1, 2, 3, 4, and 5 for Reflection Question 2. It is important to note that the raters for this phase of the coding were different individuals and were unaware of the codes generated by the raters, as described in Section 4.3. They met to discuss the similarities and discrepancies. Based on the discussion, the code book was created, and raters re-performed the coding and met again to calculate the interrater reliability. At this point, the two raters performed inter-rater reliability, which was 87%. Later, Topics 5 and 6 were coded independently by Rater 1 for Reflection Question 1, and Topics 6, 7, and 8 were coded independently by Rater 2 for Reflection Question 2.

5. Results

5.1. Students' Perceptions of Potential Influences of Their Cultural Backgrounds in Their Teamwork Communication Styles

Table 2 represents the codes and definitions for each code that emerged from the manual coding process. For the manual coding process, student reflections were read and coded by two coders. The process followed for the manual coding is shown in Figure 2 and explained in Section 4.3. Table 2 also presents representative quotes from students' reflections exemplifying each code. The codes in Table 2 describe the students' perceptions of the influence of culture on teamwork communication. From the results, we can infer that the students demonstrated or increased their level of awareness on how culture may have an impact on teamwork communication and team dynamics, such as leadership. It may have also raised the level of awareness regarding psychosocial factors, such as equity, respect, trust, and understanding of others. On the other hand, some students felt that culture had no obvious influence on team communication. Some students also felt that demographic factors, such as gender ratio, could influence the teams' communication style, and some students also preferred online communication over face-to-face communication.

Table 2. Codes generated from manual coding for Reflection Question 1.

Codes	Definition	Representative Quote
Leadership as needed	People stand out and lead the conversation when no one else tries to lead the conversation or take a leadership role when required.	"I think culture definitely plays a role in this process within our team because everyone has a different communication style despite the fact we are close in age and studies the same major while being in the same college . . . I could see that some members came from a culture where speaking up and speaking directly and have strong leadership skills, but someone members came from a culture that feels the opposite."
Unnoticeable cultural differences	There is no obvious influence that cultural background has on communication due to the similar culture that group members have.	"I don't think our group is very different in the communication styles we grew up with two of our team members are from within an hour of the University, and two of us are from close to the same place in the United States."
Equal chances to speak	Everyone has equal chances to express their own ideas and thoughts.	"In my team, I can be pretty safe to assume that all of us have been raised in an American culture where a shift is being observed of more confrontation and less hierarchy. These two concepts can be observed in our team as we will interject when possible in order to add more to each thought, and we will treat each other equally."

Table 2. Cont.

Codes	Definition	Representative Quote
Respect for others' ideas	Show respect for each other's different ideas and thoughts.	"Although each of us may have had a similar upbringing and thinking about situations, there are still some cultural differences that influence communication styles and help when solving problems we run into. In our group, we respect when other people are speaking and listen to what they have to say."
Trust in teammates	Showing trust for teammates.	"For me, I rely on constant communication and precise communication. I also rely on people holding true to their word. I always grew up where if you say something, you would do it, so when my teammates say they will do something I trust that they will get it done."
Understanding others' backgrounds	Understanding group members' cultural backgrounds.	"I think that culture has influenced our process within our team is our work ethic. Some people in the group come from a culture that prides itself on work ethic and getting stuff done early. This leads to some coming off really strong when communicating to the team because they want to get stuff done early rather than late."
Gender ratio within teams	The gender ratio would also influence the communication styles within a team.	"I think that the role of culture has chosen who takes charge during our meeting times. We have 3 males and 1 female in our group. The people who are speaking the most in the group are the males. We need to be aware that [female student name] has great ideas and need to give her the opportunity to share those ideas. Historically culture has said that males are the leaders and that is just not the case."
Online communication is more comfortable	Communication through online platforms helps individuals to communicate with more comfort.	"My team communicates over Teams or online platforms. Instead of choosing to meet in-person, we decided, especially in these times to meet over an online platform as it was easy for all team members."

Table 3 represents the topic model results and the themes that were generated from the topics, as well as the corresponding manual coding codes for each topic model theme. Topic T1_R1 has the largest weight. This topic discusses the role of culture in shaping an individual's perception. Therefore, the theme that emerged for this topic is *culture helps to see people from different lenses*. Under this theme, students have discussed how having a prior understanding of other cultures helps to understand people from diverse backgrounds. For example, one of the students mentioned, "having an understanding of cultural differences has positively influenced our team, as each of our group members comes from a different cultural background. It has influenced our conversation and how we each try to interact with each other, as we each have different views on how things should work." This theme also aligns with the *equal chance to speak*, *respect for others' ideas*, and *trust on teammates* codes from Manual Coding Phase 1.

T2_R1 has the second largest weight, and the theme for this topic is that *understanding culture promotes bonding*. For example, one student discussed the role of culture in promoting team bonding and interaction. The student said, "my team tends to talk about a lot of topics unrelated to the work while we are working together. These conversational topics are related to the other teammates' cultures because my group has some diversity in culture, such as American, Korean, and Indian. While this conversation may seem distracting at first, it helps the teammates feel more friendly and open to giving ideas. I believe this practice is based upon American culture which is formed by people with various backgrounds, and it is natural to be embracing and curious about the differences." Additionally, it is important to note that none of the codes from the manual coding aligned with this theme.

Table 3. Topic model results and themes for Reflection Question 1.

Topic Id	Topic Weight	Top 20 Words	Theme	Manual Coding Phase 1 Codes
T1_R1	5.51729	culture team communication group members work influenced role people cultures communicate process speak styles don't similar things cultural time feel	Culture helps to see people from different perspectives.	Equal chances to speak, respect for others' ideas, and trust in teammates.
T2_R1	0.70183	open member Chinese the role day Korean terms note handle found past manner outspoken talking American gave leader prefer mesome longer	Understanding culture promotes bonding.	None.
T3_R1	0.65222	understanding lot teamwork problem job exposed bit pretty rest there are makes barrier differently big their culture older personalities completely except our conditions	Culture has helped to develop an understanding of people from diverse backgrounds.	Understanding others' backgrounds.
T4_R1	0.6157	conversation states teamwork directly persons respect interrupt give students mind current comfortable contribute lack reason due moment case equal worked	The culture has influenced how we talk to one another on the team.	Understanding others' backgrounds.
T5_R1	0.56384	males works aspect worry south aspects decided online expected significant speaking likes conflict hasn't issues great project extremely communicate discussed	Culture can foster stereotypes in leadership roles.	Leadership as needed, gender ratio within teams.
T6_R1	0.552	scrum complete upbringing times learned tasks situations topics messaging groupmate grow age create collective decision solving meetings talk conflict groups	Culture can cause communication challenges.	None.
		None.	None.	Unnoticeable cultural differences.
		None.	None.	Online communications are more comfortable.

T3_R1 discusses the role of culture in understanding people from different backgrounds. The theme that emerged from the topics was *culture can help develop an understanding of people from diverse backgrounds*. For example, a student mentioned how being cognizant of cultural differences has helped his team members to develop mutual respect and understanding for one another. The student said, "I believe that culture influenced my team to respect each other and give everyone a chance to speak up. We are 4 team members, 2 are from the US, and 2 from Saudi Arabia, so it really balances well." This theme for T3_R1 aligns with the manual code of *understanding others' backgrounds*.

T4_R1 has the fourth largest weight, and the theme for this topic is that *culture has influenced how we talk to one another in the team*. For example, one student discussed the role of culture in influencing team communication. The student said, "I think that culture has influenced the communication styles in my team in many ways. Since I have team members from various cultures, some members of my team have a more structured way

of communicating and treat each meeting in a very formal structured way. Whereas other members treat the meeting as casual and open.” This theme for T4_R1 aligns with the manual code of *understanding others’ backgrounds*.

T5_R1 has the fifth largest weight, and the theme for this topic is that *culture can foster stereotypes in leadership roles*. For example, a student described how culture played an important role in promoting gender stereotypes among team members when leading team meetings. The student said, “I think that the role of culture has chosen who takes charge during our meeting times. We have 3 males and 1 female in our group. The people who are speaking the most in the group are the males. We need to be aware that [female student name] has great ideas and need to give her the opportunity to share those ideas. Historically culture has said that males are the leaders, and that is just not the case.” This theme for T5_R1 aligns with the manual code of *leadership as needed and gender ratio within teams*.

T6_R1 has the sixth largest weight, and it is also the last topic for Reflection Question 1. The theme for this topic is *culture can cause communication challenges*. The students described for this theme how coming from different cultural backgrounds can cause communication challenges. For example, a student mentioned: “I think that as mostly Americans, we use distinctly direct forms of communication. Whenever someone is behind on work or their work is due soon, we will mention them and try and gather information about how they’re coming along to ensure that it’ll be done by the due date. Some people are afraid to speak up, though, because they might interpret it as rude.” Additionally, it is important to note that none of the codes from manual coding aligned with this theme.

If we compare the results of the manual coding of Phase 1 with the themes that emerged through topic modeling in Phase 2, we see a 75% match. The 75% match indicates that out of the eight codes generated from manual coding, six codes aligned with themes generated from topic modeling. For example, T1_R1 represents three manual coding themes: equal chances to speak, respect for others’ ideas, and trust in teammates. It is also important to note that the themes for manual coding and topic modeling were generated through qualitative analysis. We also observed that topic modeling did not identify a separate topic for *online communication* and students *not perceiving any influence* of their cultural backgrounds on their interactions with their teammates. However, topic modeling identified an additional theme of culture that can cause communication challenges that were missed during the manual coding process in Phase 1.

5.2. Students’ Perceptions of Potential Influences of Their Cultural Backgrounds on Their Interactions with Their Teammates

Table 4 shows the codes for students’ perceived potential influences of their cultural backgrounds on their interactions with their teammates. It also presents representative quotes from students’ reflections exemplifying each code. Overall, students believe that their interactions with their teammates are influenced by their family, schooling, and other upbringing experiences. These influences play an important role in shaping students’ teamwork experiences in communication; for instance, listening rather than speaking, and showing respect for others. Sharing responsibilities among all team members influences their level of involvement in the group project, as does being more task-oriented and showing more individualistic views.

Table 5 represents the topic model results and themes that were generated from the topics and also manual coding codes that align with the topic model theme for the Reflection Question 2. T1_R2 has the largest weight for Reflection Question 2. The theme that emerged for this topic is *family background and upbringing helped students to develop teamwork skills*. In this theme, students have described the role of their upbringing and family background in helping them develop teamwork skills. For example, a student described how being from Vietnam and being a student in the US has influenced his teamwork skills. The student said, “I am from Vietnam, where building rapport before starting to work together is somewhat expected. Since I have been in the US, I think my style of communication is a mix of the two cultural values. I like to get started on work as soon as possible, but I would like to

build rapport with my teammates along the way.” Moreover, the manual coding codes that match this theme are *upbringing experiences and parental influence*.

Table 4. Codes generated from manual coding for Reflection Question 2.

Codes	Definition	Representative Quotes
Parental influence	Some of a person’s qualities are inherited from parents and family members. Parents’ views about life and tasks influence their children’s attitudes and behavior.	“As a child of immigrant parents from China, I grew up under parents who communicated with me in a very direct and simple way. Growing up in essentially a Chinese household, I think this has contributed to my team communication style, where I am very upfront with other members on what needs to be done and when. Obviously being direct sometimes comes with a negative connotation given how stern it sounds, but I think my American culture has made it so that I can be more upfront with others while being wary of how my comments can be perceived by others.”
Listening rather than speaking	Certain individuals prefer to listen to others rather than express their own opinions.	“I think that my cultural background has influenced our teamwork interaction. I am used to being in team environments where people do not like to talk a lot but listen.”
Shared responsibilities among teammates	Each participant has an equal role and responsibility.	“I come from a background that prioritizes equality between members, so I believe that everyone in our team should have the same amount of say when it comes to decisions and contribute equally to projects.”
Respect for others’ ideas	Showing respect for each other’s different ideas and thoughts.	“With my cultural background, I believe it causes me to be very respectful of what others have to say in my team. I think making sure I treat everyone with respect is crucial in a team setting.”
Task-Oriented and individualistic views	Certain students use direct communication and feel more at ease and efficient while working alone. Additionally, some individuals are more concerned with completing the work than with team collaboration.	“Since I personally come from the southern United States, I might have a slightly different culture than my teammates. It’s more acceptable to be more open with people, so I may be being more direct with my teammates than they are used to.”
Upbringing experiences	The location of one’s upbringing or the setting in which an individual grew up.	“Personally, I grew up in a pretty traditional family environment consisting of partially progressive partial American dreamer family culture. There is a lot of emphasis placed on the importance of education and the value of hard work and close relationships with family. Therefore, this may influence teamwork interaction because it encourages me to update team members on what I have done and ensure that I complete my roles on time and keep up with our team tasks.”

Topic T2_R2 has the second largest weight, and it discusses the impact of culture on communication. The theme for this topic is *my cultural background has made me less communicative*. Under this theme, students describe how their family background, upbringing or their personal preferences have affected their communication styles and have made them less verbally expressive. One student said, “I think my own cultural background has influenced me to be quieter and more reserved. I find it hard to relate to people, and I always have felt like somewhat of an outsider, even in my own circles. So I think it may be why I don’t like posting messages or responding until it’s absolutely necessary. I also do not like being in group meetings longer than the bare minimum.” Moreover, the manual coding codes that match this theme are *upbringing experiences, parental influence, and listening rather than speaking*.

Table 5. Topic model results and themes for Reflection Question 2.

Topic Id	Topic Weight	Top 20 Words	Theme	Manual Coding Phase 1 Codes
T1_R2	0.38247	leadership comfortable found process master roles adapt complete promotes change confident role kind students states ive taking start point good	Family background and upbringing helped students to develop teamwork skills.	Upbringing experiences, parental influence.
T2_R2	3.47678	background team cultural culture work teamwork people influence group interaction don't communication tend speak members family raised make things influences	My cultural background has made me less communicative.	Upbringing experiences, parental influence, and listening rather than speaking.
T3_R2	0.48292	part years leader social problems thing focused active listen experience issues loud diverse teamwork interaction greatly time considered line paced fast	My family background taught me collaboration.	Parental influence.
T4_R2	0.48133	time working individual understanding heavily Mexico contribute style work differences related told attending aggressive independent accustomed USA focused interacting high	Personal experiences influence teamwork interaction.	Upbringing experiences.
T5_R2	0.39661	household bit parents made end everyone style task Kenyan hierarchy nontraditional negative provide toes environment school ive opinions large grew	Teamwork means sharing responsibilities.	Shared responsibilities among teammates.
T6_R2	0.37257	communicate share Turkish generally submitted realized setting efficiently assume culturally works video interaction because interactions makes likes raised awkward reason tend	High school experiences shaped my teamwork skills.	Upbringing experiences.
T7_R2	0.35795	point cultural side advocate finish polish effectively personally teams complicated technology responding hard music subcultures play leader taught culture opposite	My cultural background has taught me to respect others.	Respect for others' ideas.
T8_R2	0.34233	follow affect groupmate slightly sort observing working create role are more learned born leads positive personal decision needed power create confusion methods very open	My culture has taught me to be open and direct.	Task-oriented and individualistic views

Topic T3_R2 has the third largest weight. It discusses the positive impact of family background. The theme that emerged for this topic is *my family background taught me*

collaboration. One of the students shared her experience. She mentioned, “My own cultural background can influence my teamwork in a positive way. I come from an African background, specifically Kenya, and being raised in a Kenyan household, I was taught to care for others and keep up for the household, and my younger siblings at a very young age and to be selfish in my family was something we were all taught not to do. So as I implemented the culture, I was raised with into my teamwork and collaborative skills, something important to me. I make sure everyone’s opinions are heard and that we all understand the task we are given.” Moreover, the manual coding code that matches this theme is *parental influence*.

Topic T4_R2 has the fourth largest weight. It describes the impact of personal experiences on teamwork. The theme that emerged for this topic is that *personal experiences influence teamwork interaction*. For example, a student mentioned how his personal experiences have influenced his teamwork skill. He said, “I try to use my personal experiences to bring a perspective to the team. I come from a cultural background that places a strong emphasis on equality and freedom of expression, so the interactions I have had with teammates until now have always been unrestricted. As a result, I have grown accustomed to this method of interaction, and it has set the standard for what I consider to be effective teamwork interactions.” Moreover, the manual coding code that matches this theme is *upbringing experiences*.

Topic T5_R2 has the fifth largest weight. The theme for this topic is *teamwork means sharing responsibilities*. In this theme, students have mentioned how sharing responsibilities and being open to suggestions helps with completing the work. For example, a student said, “I believe my own cultural background may influence my teamwork interaction because similar to my team members. We all share cultural backgrounds with similar concepts and are able to work much more effectively with each other. We are open to new ideas and speak up when we need to add some input. We also make sure to include each other as well so that everyone’s ideas are considered during the creation of the project.” Moreover, the manual coding code that matches this theme is *shared responsibilities among teammates*.

Topic T6_R2 has the sixth largest weight. The theme for this topic is *high school experiences shaped my teamwork skills*. Under this theme, students described the role of their high school experiences in shaping their teamwork skills. For example, a student described how his high school experience of being on a sports team had help him lead the team. The student said, “I think my cultural background coming from a small town in northeast Indiana has made me confident in who I am and willing to stand up and be a leader like I was in high school on sports teams. I think that this would be good for the team because having driven leaders on the team is good.” The manual coding code that matches this theme is *upbringing experiences*.

Topic T7_R2 has the seventh largest weight. The theme for this topic is *my cultural background has taught me to respect others*. Under this theme, students discussed the role of their cultural background in making them patient and respectful. For example, a student said, “I have always been taught to respect others and wait until they finish talking. I apply the same to my team.” The manual coding code that matches this theme is *respect for others’ ideas*.

Topic T8_R2 has the eighth largest weight. The theme for this topic is *my culture has taught me to be open and direct*. Students under this theme discussed how their culture has made them open to seeking and sharing ideas. Additionally, they mentioned how their culture has made them direct and individualistic when expressing their opinion. For example, a student mentioned, “As said prior, I come from a culture focused on self-sufficiency. I enjoy collaboration and teamwork, but I am also good at working on my own. Our group meetings are typically just check-ins where we share what we worked on by ourselves. As the project progresses, I will continue to reach out and keep up open communication.” Moreover, the manual coding code that matches this theme is *task-oriented and individualistic views*.

If we compare the results of the manual coding Phase 1 with the results of themes generated for topic models, we find that there was an 100% agreement between the topic

model themes and the manual coding codes for Phase 1. However, it is also important to note that the topic model helped us conduct a deeper analysis of the data; for example, if we consider the code *upbringing experiences* we see that it generated topics related to the impact of personal experiences and high school experiences on teamwork. The topic model helped us learn more about the impact of the *upbringing experiences* of students.

6. Discussion and Practical Implications

The study compared the efficacy of manual coding and the topic model technique supplemented with manual coding. The study used a two-step process to conduct the qualitative analysis, as shown in Figure 2. In the first phase, the reflection data for both questions were separately analyzed qualitatively by two raters using a manual human coding process. In the next step, the data were analyzed using topic modeling combined with human coding, and the topics pertaining to each reflection question were reported. For the first reflection question, six topics were generated, and for the second reflection question, eight topics were generated. Then, two raters (in this case, different individuals from the Phase 1 raters) again qualitatively analyzed the data for each topic using manual coding. It is also important to note that the two raters in Phase 2 were unaware of the codes generated in the first phase of the manual coding process.

When approaching the first research question (RQ1) regarding the relative coverage of the LDA topic model and human coding in terms of the breadth of the topics/themes extracted from the text collection, we can conclude that the coverage was comparable. The results of the study indicated that for Reflection Question 1, there was a 75% match between the results of manual coding performed using the topic modeling and manual coding performed in Phase 1. The topic modeling helped us identify an additional theme, *communication challenges*, which was missed during the manual coding in Phase 1. On the other hand, the topic modeling approach missed the codes for *unnoticeable cultural differences* and *online communications are more comfortable* generated by manual coding approach. For Reflection Question 2, the matching percentage was 100%. In this case, the topic modeling approach helped us conduct a deeper analysis of data for both reflection questions. For example, the manual coding identified the codes as *upbringing experiences*, but the topic model helped us understand the specific instances of previous experiences, such as high school experiences, personal experiences, family background, etc. Based on the comparable results obtained from topic modeling and human coding approaches of qualitatively analyzing two datasets of student reflections, we can infer that topic modeling was an effective and comparable approach.

Regarding the second research question (RQ2) considering the relative depth or level of detail among the identified topics using LDA topic models and human coding approaches, we can conclude that the topic modeling approach provided a comparable coverage. Although the topic modeling approach provided comparable coverage, it also provided more refined topics, as compared to the codes generated by the manual coding approach. However, these findings may not be generalizable for different types of datasets, as the collection in this study consisted of long-form documents, and the count of documents was not very large. It is important to note that human interpretation of the themes of the topics identified by the LDA topic model is crucial in making sense of the topics generated. It is also important to highlight that the manual coding and interpretation of topic themes were performed by different researchers, so there may have been some subjective variation among them and their levels of expertise might also have varied.

We also observed that topic modeling is relatively efficient compared to human coding, especially when the dataset is large. The increase in time and effort needed for qualitative analysis will be considerably higher for the manual coding approach compared to the topic model approach. While the exact time taken for analyses using manual coding and topic model theme interpretation was not recorded, based on the effort reported by the researchers, we estimate that manual coding took approximately 15 h to develop one code, while for the LDA analysis, the time taken for theme determination by analyzing the top

words and documents was about one hour per topic. It should also be noted that the manual coding and LDA topic-theme analysis were performed by different researchers; therefore, there may be variations in the level of expertise of the coders for the two approaches, so an accurate comparison was not possible. However, the effort estimates indicated that LDA can help in reducing the time taken for qualitative analysis and cover a larger sample size quickly. The LDA algorithm can help to quickly identify prominent topics that could be further interpreted qualitatively by a human, thus guaranteeing that the main topics from the textual collection will not be missed. On the other hand, in traditional human coding, the identification of codes are dependent on the level of expertise and experience of the human coder, and some important codes may be missed if the collection of documents is relatively large.

Overall, combining the human coding and topic modeling approaches can serve as an effective method for conducting a qualitative data analysis of a large collection of documents. It is also important to note that the efficacy of the topic modeling algorithms is high when the dataset is larger; therefore, it may not be efficient to use topic modeling for smaller datasets. Additionally, when conducting human coding, it is also important to acknowledge the level of expertise of the coders. Analysis conducted by novice coders might not be as effective as when the themes were generated by experienced coders. To some extent, the topic modeling approach ensures good coverage and depth of themes identified by using the algorithm to scan through all the documents in the collection and then generating the topics, which is not guaranteed in the manual coding approach as it is more dependent on the expertise of human coders.

7. Conclusions

This study compared two approaches for analyzing qualitative data; a traditional approach following human coding and a computer-based approach supplementing topic analysis with human coding. Our findings suggest that both methods are comparable in terms of: (a) The relative coverage in terms of the breadth of the codes extracted from the text collection; and (b) The relative depth or level of detail among the identified topics. This is an important contribution because, to our knowledge, no study has been conducted that compares the validity of both methods on the same dataset. Although this is an important contribution, we also recognize that the findings have limitations. Specifically, the sample was relatively small (although in qualitative research, it can be considered large), and it only focused on the case of short reflection. Future work could consider other types of data formats, such as responses to questionnaires with shorter answers or essays with longer responses. The length and context of other formats in the data may impact both the hand coding and the computer-based approach. Despite its limitations, the findings of this study highlight the value of a computer-based approach for supplementing qualitative analyses. Our findings also highlight the importance of a human in the loop, either contributing with the categories upfront or providing meaning once the topics are generated by the computer-based approach. That is, the computer-based approach can supplement, but not replace, existing qualitative methods. As previous research [32] has repeatedly suggested, human in the loop is a critical step for ensuring the trustworthiness, validity, and reliability of the findings, as no computer-based approach is capable of interpreting or providing insights under a conceptual or theoretical lens.

A potential approach that could overcome the limitations of our findings may be to separately apply the two approaches presented in this study with a manageable sample size to increase the validity and trustworthiness of the analysis in uncovering the codes. Then, once the analysis is validated by both approaches, the sample size can be increased and the computer-based approach can be applied to strengthen the generalizability of the findings. This proposed combined approach could also take into consideration the potential subjectivity and level of expertise of the human coders.

Author Contributions: Conceptualization, G.N. methodology, G.N., A.J. and A.J.M.; validation, A.J. and H.C.; formal analysis, G.N., A.J., H.C., Y.Z. and A.C.; investigation, G.N., A.J., H.C., Y.Z. and A.C.; resources, A.J.M.; data curation, A.J., H.C., Y.Z. and A.C.; writing—original draft preparation, G.N., A.J. and A.J.M.; writing—review and editing, G.N., A.J. and A.J.M.; supervision, A.J. and A.J.M.; project administration, A.J. and A.J.M.; funding acquisition, A.J.M. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the U.S. National Science Foundation under award numbers #2113991 and #2219271. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of NSF or the U.S. Government.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Purdue University (protocol code IRB-2021-1181 and date of approval 08-20-2021).

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Crowston, K.; Allen, E.E.; Heckman, R. Using natural language processing technology for qualitative data analysis. *Int. J. Soc. Res. Methodol.* **2012**, *15*, 523–543. [\[CrossRef\]](#)
2. Gillies, M.; Murthy, D.; Brenton, H.; Olaniyan, R. Theme and Topic: How Qualitative Research and Topic Modeling Can Be Brought Together. *arXiv* **2022**, arXiv:2210.00707.
3. Chauhan, U.; Shah, A. Topic Modeling Using Latent Dirichlet Allocation: A Survey. *ACM Comput. Surv.* **2021**, *54*, 145. [\[CrossRef\]](#)
4. Blei, D.M.; Ng, A.Y.; Jordan, M.I. Latent Dirichlet Allocation. *J. Mach. Learn. Res.* **2003**, *3*, 993–1022.
5. Nanda, G.; Douglas, K.A.; Waller, D.R.; Merzdorf, H.E.; Goldwasser, D. Analyzing Large Collections of Open-Ended Feedback from MOOC Learners Using LDA Topic Modeling and Qualitative Analysis. *IEEE Trans. Learn. Technol.* **2021**, *14*, 146–160. [\[CrossRef\]](#)
6. Nanda, G.; Wei, S.; Katz, A.; Brinton, C.; Ohland, M. Work-in-Progress: Using Latent Dirichlet Allocation to Uncover Themes in Student Comments from Peer Evaluations of Teamwork. In Proceedings of the 2022 ASEE Annual Conference & Exposition, Minneapolis, MN, USA, 25–29 June 2022.
7. Yang, S.; Zhang, H. Text Mining of Twitter Data Using a Latent Dirichlet Allocation Topic Model and Sentiment Analysis. *Int. J. Comput. Inf. Eng.* **2018**, *12*, 525–529.
8. Kumar, R.; Raghuvver, K. Legal Document Summarization Using Latent Dirichlet Allocation. *Int. J. Comput. Sci. Telecommun.* **2012**, *3*, 8–23.
9. Li, G.-K.J.; Trappey, C.V.; Trappey, A.J.; Li, A.A. Ontology-Based Knowledge Representation and Semantic Topic Modeling for Intelligent Trademark Legal Precedent Research. *World Pat. Inf.* **2022**, *68*, 102098. [\[CrossRef\]](#)
10. Lebeña, N.; Blanco, A.; Pérez, A.; Casillas, A. Preliminary Exploration of Topic Modelling Representations for Electronic Health Records Coding According to the International Classification of Diseases in Spanish. *Expert Syst. Appl.* **2022**, *204*, 117303. [\[CrossRef\]](#)
11. Chowdhury, M.F. Coding, sorting and sifting of qualitative data analysis: Debates and discussion. *Qual. Quant.* **2015**, *49*, 1135–1143. [\[CrossRef\]](#)
12. Elliott, V. Thinking about the coding process in qualitative data analysis. *Qual. Rep.* **2018**, *23*, 2850–2861. [\[CrossRef\]](#)
13. Cleland, J. Exploring versus Measuring: Considering the Fundamental Differences between Qualitative and Quantitative Research. In *Researching Medical Education*; Wiley Online Library: Hoboken, NJ, USA, 2015; pp. 1–14.
14. Allwood, C.M. The Distinction between Qualitative and Quantitative Research Methods Is Problematic. *Qual. Quant.* **2012**, *46*, 1417–1429. [\[CrossRef\]](#)
15. Evans, M.S. A computational approach to qualitative analysis in large textual datasets. *PLoS ONE* **2014**, *9*, e87908. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Chen, N.-C.; Drouhard, M.; Kocielnik, R.; Suh, J.; Aragon, C.R. Using Machine Learning to Support Qualitative Coding in Social Science: Shifting the Focus to Ambiguity. *ACM Trans. Interact. Intell. Syst.* **2018**, *8*, 9. [\[CrossRef\]](#)
17. Crowston, K.; Liu, X.; Allen, E.E. Machine Learning and Rule-Based Automated Coding of Qualitative Data. In Proceedings of the American Society for Information Science and Technology, Pittsburgh, PA, USA, 22–27 October 2010; Volume 47, pp. 1–2.
18. Rosenberg, J.M.; Krist, C. Combining Machine Learning and Qualitative Methods to Elaborate Students’ Ideas about the Generality of Their Model-Based Explanations. *J. Sci. Educ. Technol.* **2021**, *30*, 255–267. [\[CrossRef\]](#)
19. Marathe, M.; Toyama, K. Semi-automated coding for qualitative research: A user-centered inquiry and initial prototypes. In Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems, Montreal, QC, Canada, 21–26 April 2018; pp. 1–12.

20. Baumer, E.P.; Mimno, D.; Guha, S.; Quan, E.; Gay, G.K. Comparing Grounded Theory and Topic Modeling: Extreme Divergence or Unlikely Convergence? *J. Assoc. Inf. Sci. Technol.* **2017**, *68*, 1397–1410. [[CrossRef](#)]
21. Muller, M.; Guha, S.; Baumer, E.P.; Mimno, D.; Shami, N.S. Machine Learning and Grounded Theory Method: Convergence, Divergence, and Combination. In Proceedings of the 19th International Conference on Supporting Group Work, Sanibel Island, FL, USA, 13–16 November 2016; pp. 3–8.
22. Nelson, L.K. Computational Grounded Theory: A Methodological Framework. *Sociol. Methods Res.* **2020**, *49*, 3–42. [[CrossRef](#)]
23. Jaiswal, A.; Patel, D.; Zhu, Y.; Lee, J.S.; Magana, A. A Reflection on Action Approach to Teamwork Facilitation. In Proceedings of the 2022 ASEE Annual Conference & Exposition, Minneapolis, MN, USA, 25–29 June 2022.
24. Khandkar, S.H. *Open Coding*; University of Calgary: Calgary, AB, Canada, 2009.
25. Kherwa, P.; Bansal, P. Topic Modeling: A Comprehensive Review. In *EAI Endorsed Transactions on Scalable Information Systems* **2019**, *7*, e2. [[CrossRef](#)]
26. Jelodar, H.; Wang, Y.; Yuan, C.; Feng, X.; Jiang, X.; Li, Y.; Zhao, L. Latent Dirichlet Allocation (LDA) and Topic Modeling: Models, Applications, a Survey. *Multimed. Tools Appl.* **2019**, *78*, 15169–15211. [[CrossRef](#)]
27. Chang, J.; Gerrish, S.; Wang, C.; Boyd-Graber, J.; Blei, D. Reading Tea Leaves: How Humans Interpret Topic Models. In *Advances in Neural Information Processing Systems 22: Proceedings of the 23rd Annual Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 7–10 December 2009*.
28. Mimno, D.; Wallach, H.; Talley, E.; Leenders, M.; McCallum, A. Optimizing Semantic Coherence in Topic Models. In Proceedings of the 2011 Conference on Empirical Methods in Natural Language Processing, Edinburgh, Scotland, UK, 27–31 July 2011; pp. 262–272.
29. Röder, M.; Both, A.; Hinneburg, A. Exploring the space of topic coherence measures. In Proceedings of the 8th ACM International Conference on Web Search and Data Mining, Shanghai, China, 2–6 February 2015; pp. 399–408.
30. Welcome to PyLDAvis's Documentation PyLDAvis 2.1.2 Documentation. Available online: <https://pyldavis.readthedocs.io/en/latest/> (accessed on 4 March 2023).
31. Topic Modeling. Available online: <https://mimno.github.io/Mallet/topics.html> (accessed on 4 March 2023).
32. Rahman, S.; Kandogan, E. Characterizing practices, limitations, and opportunities related to text information extraction workflows: A human-in-the-loop perspective. In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems, Orleans, LA, USA, 29 April–5 May 2022.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.