Zhang, S., Gao, Q., Wen, Y., Li, M., & Wang, Q. (2021). Automatically Detecting Cognitive Engagement beyond Behavioral Indicators: A Case of Online Professional Learning Community. *Educational Technology & Society*, *24* (2), 58–72.

Automatically Detecting Cognitive Engagement beyond Behavioral Indicators: A Case of Online Professional Learning Community

Si Zhang[1*], Qianqian Gao[1], Yun Wen[2*], Mengsiying Li[3] and Qiyun Wang[2]
[1]Hubei Research Center for Educational Informationization, Faculty of Artificial Intelligence in Education, Central China Normal University, China // [2]National Institute of Education, Nanyang Technological University, Singapore // [3]School of Management, Wuhan College, China // djzhangsi@mail.ccnu.edu.cn // 2956954801@qq.com // yun.wen@nie.edu.sg // 1061224142@qq.com // qiyun.wang@nie.edu.sg
[*]Corresponding author

(Submitted August 16, 2020; Revised September 6, 2020; Accepted February 1, 2021)

ABSTRACT: Online discourse is widely used in diverse contexts of learning and professional training, but superficial interactions and digression often occur. In the face of these problems and the large-scale unstructured text data, the traditional way of learning analytics has been challenged in terms of providing timely intervention and feedback. In this paper, a workflow for automatically detecting in-service teachers' cognitive engagement in an online professional learning community is described. Discourse data of 1834 in-service teachers involved in a teacher professional development program was collected and processed using the Word2vec toolkit to generate lexical vectors. The method of vector space projection was used to calculate the new information contained in each post, cosine similarity was used to calculate topic relevance, and cluster analysis was used to explore in-service teachers' discourse characteristics. Results showed that in-service teachers' average contribution was 4.59 posts and the average length of each post was 39.47 characters in Chinese. In the mathematics online professional learning community, the average amount of new information contained in each post was 0.221 and in-service teachers' posts contained much new information in the early stages of online discourse. Most in-service teachers' posts were relevant to the discussion topic. Cluster analysis showed three different groups of posts with unique characteristics: high topic relevance with much new information, high topic relevance with little new information, and low topic relevance with little new information. Finally, limitations are discussed and future research directions are proposed.

Keywords: Learning communities, Computer-mediated communication, Evaluation methodologies, Interactive learning environments

1. Introduction

The advantages of the online professional learning community (OPLC) have been identified by previous researchers as being essential to support reflection and discussion that can improve in-service teachers' learning engagement and leverage changes in their classrooms (e.g., Xing & Gao, 2018). Discourse in online professional learning communities helps in-service teachers share teaching experiences, solve teaching problems together, and deepen understanding of the relationship between theory and practice (Tsai, 2012). Millions of in-service teachers participate in online professional development programs each year, and they spend a significant amount of time learning in OPLCs (e.g., Zhang, Liu, Chen, Wang, & Huang, 2017; Xing & Gao, 2018). Nevertheless, it would be too simple to assume that the establishment of an OPLC will automatically promotes in-service teachers' effective interaction and collaborative knowledge construction (Baran & Cagiltay, 2010). Superficial interactions and off-task messages often occur in in-service teachers' online discussions (e.g., Lantz-Andersson, Lundin, & Selwyn, 2018). Similarly, some in-service teachers do not contribute any new information, and this adversely affects the participation and sharing willingness of other members of the community (e.g., Zhang et al., 2017). Discourse in an OPLC is a social and cognitive interwoven process and the effectiveness of online discourse depends on both the iterative interaction and the active information exchange among community members (e.g., Chen, Fan, & Tsai, 2014). Meanwhile, content-related (or on-topic) discussions are more conducive to learning than off-topic discussions (e.g., Cho, 2016). According to the Community of Inquiry (CoI) framework, engaging learners in critical reflection and collaborative discussion fosters their high-order learning (Heilporn & Lakhal, 2020). The CoI framework defines three critical elements of online learning: cognitive presence, social presence, and teaching presence. Among the elements, cognitive presence is a central component in the model, which describes stages of inquiry-based learning, including problem presentation, knowledge synthesis, and evaluation (Garrison, Anderson, & Archer, 1999). Thus, the effective measures of teachers' cognitive engagement in the OPLC that support a movement toward understanding their cognitive presence and providing timely intervention and regulation are required.

Cognitive engagement refers to learners' psychological involvement in academic tasks (Fredricks, Blumenfeld, & Paris, 2004). Itis positively correlated with learners' academic performance (Wang, Wen, & Rosé, 2016). In terms of measuring cognitive engagement, the commonly used indicators include the number of posts, speed of response, or time on task (e.g., Kim, Park, Yoon, & Jo, 2016; Xie, Heddy, & Greene, 2019). It is often not enough to measure learners' cognitive engagement only by observable indicators, but combining quantitative dada and qualitative content analysis will provide more reliable results (e.g., Atapattu, Thilakaratne, Vivian, & Falkner, 2019). However, the analysis of the large-scale online discourse data (mainly textual data) raises methodological problems, including data collection and coding (Vieira, Parsons, & Byrd, 2018). Facing the sheer data volumes and the diversity of in-service teachers' language expression, the traditional way of data coding used in class discourse analysis, however, cannot address this dilemma.

In this study, we present an automatic discourse analysis approach using a language modeling technique called neural word embeddings (Word2vec) (Mikolov, Sutskever, Chen, Corrado, & Dean, 2013) to detect learners' active participation in discussions in the online professional learning community.

2. Literature review

2.1. Online professional learning community and online discourse

Many online professional learning communities have been built using social media tools (e.g., Twitter, Blog, WeChat), online learning platforms (e.g., Blackboard, Moodle), or course management systems (CMS) (e.g., Lantz-Andersson et al., 2018; Prestridge, 2019). Online professional learning communities provide learning resources for in-service teachers including video cases, online discourse, as well as reflections (e.g., Lantz-Andersson et al., 2018). Online discussion with experts or peers is useful for in-service teachers' professional development. In-service teachers have opportunities to think about teaching problems from multiple perspectives (Lee & Brett, 2015; Polizzi, Head, Barrett-Williams, Ellis, Roehrig, & Rushton, 2018). Moreover, online discussion improves the ability of teacher groups by solving teaching problems through role-playing and collaboration (Yang, 2016). In addition, online discussion helps to improve the interactions between the members in the online learning communities through questioning, interpreting, and elaborating (Cho, 2016).

Although online discourse is available in most online professional learning communities, it is primarily used as a supporting discussion between participants rather than for meaningful knowledge sharing and co-construction (e.g., Atapattu et al., 2019; Zhang et al., 2017). Prestridge (2019) argued that the critical discourse that helps to facilitate teachers' knowledge construction and transform classroom practice was only sustained for a small number of postings by teachers in a given thread. According to Gunawardena, Lowe, and Anderson (1997), social knowledge construction among learners in a constructivist learning environment needs to go through some key stages, such as restating the participant's position, negotiation or clarification of the meaning of terms, and negotiation of new statements. In this regard, effective discourse depends on negotiating existing information (given information) and constantly coming up with new information. However, repeated discussion often occurs, and some teachers enact the role of information consumer and information networker in online discussion (e.g., Prestridge, 2019; Tsiotakis & Jimoyiannis, 2016). In addition, commitment and co-construction of ideas in an online professional learning community can only be achieved when teachers are deeply engaged with discussion topics because the topic is authentic and requires active cognitive engagement and in-depth understanding of the problem (Xing & Gao, 2018). Hence, in-service teachers' online discourse data in the OPLCs provides a wealth of valuable information for educational researchers and practitioners to understand the learning process. This understanding is important because it can help teacher training managers make correct intervention policies, improve the quality of teacher training programs, and thus enhance in-service teachers' teaching abilities. However, the large-scale online discourse data in OPLCs raises methodological problems, including data coding and analysis. Facing the large-scale data volumes, the diversity of teachers' language expression, and the complexity of online discourse process, the traditional way of data collection and analysis in the general classroom, cannot deal with this dilemma.

2.2. Discourse analysis based on machine learning technologies

Discourse analysis was originally proposed in the linguistics field and has been widely used in the fields of sociology, education, and psychology (Harris, 1952). In education, especially in online learning communities, text is the most commonly used data for discourse analysis (Wu, Yu, & Wang, 2018). With the growing scale of text data, discourse analysis has become a key application area of learning analytics and educational data mining.

Discourse analysis has been used to reveal the process of interaction and knowledge construction, investigate learner's behaviors and discourse content, and acquire deeper understanding of engagement in collaborative learning (e.g., Peng & Xu, 2020).

There have been two main lines of research evaluating online discourse by using the machine learning method. Some studies analyze and visualize the learners' cognitive and social presences in online learning settings (e.g., Xing & Gao, 2018). Others investigate the semantic content and processes of learner's posts in understanding interactive patterns (e.g., GaŠević, Joksimović, Eagan, & Shaffer, 2019).

The identification of new versus given information and topic relevance within a text has been frequently studied by researchers of discourse analysis. The most frequently used method to measure similarity between two texts is keyword matching, including word matching, keyword matching, andweighted keyword matching. The keyword matching method operates by finding exact matches between two texts and some disadvantages of this method include the lack of emphasis on differential importance as to how much information a particular word may carry. The strength of Latent Semantic Analysis (LSA) to measure text similarity has been demonstrated in many research studies (e.g., Chen, Chen, & Sun, 2010; Sung, Liao, Chang, Chen, & Chang, 2016). Sung et al. (2016) used 262 Chinese articles to construct a latent semantic space with 250 dimensions and the LSA was used to compute the semantic similarity between pairs of sentences, such as students' summary, expert's summary, or the source text. On account of the concept of new versus given information is related but distinct from the concept of text similarity, Hu, Cai, Louwerse, Olney, Penumatsa, and Graesser (2003) adapted the standard LSA and proposed the LSA based measure called a span method to detect given and new information in written discourse.

Word2vec is a basic word embedding technology in the field of Natural Language Processing (NLP). Word embedding originates from vector space models and aims to quantify and categorize semantic similarities between words based on their distributional properties in large-scale of text data (Li, Li, Fu, Masud, & Huang, 2016). Word2vec has been used in web search (Bing, Niu, Li, Lam, & Wang, 2018), comment sentiment classification (Zhang, Xu, Su, & Xu, 2015), and text classification (Sinoara, Camacho-Collados, Rossi, Navigli, & Rezende, 2019). Compared with other data mining methods, the accuracy of Word2vec in text classification is more than 70% (Cerisara, Král, & Lenc, 2018). The Word2vec technique has been successfully applied to detect teachers' cognitive engagement (Atapattu et al., 2019), and is at the forefront of the research field (Hashimoto, Alvarez-Melis, & Jaakkola, 2015).

3. Conceptual framework

3.1. Given and new information

Identifying given and new information in a discourse has long been regarded as an important research issue. Many researchers define given and new information from different perspectives. According to Halliday (1967), given information refers to what can be recovered anaphorically or situationally from the preceding discourse, and new information, on the contrary, is not recoverable. Chafe (1976) defines given information as "knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance" (p. 30), and new information as "what the speaker assumes he is introducing into the addressee's consciousness by what he says" (p. 30). Based on integrating previous theoretical studies, Prince (1981) develops a taxonomy that can be used to hand-code discourse text for identifying given and new information. According to Prince (1981), there are three different levels of givenness. On the first level, givenness represents the sense of predictability or recoverability which is based on the assumes that "the hearer can predict or could have predicted that a particular linguistic item will or would occur in a particular position within a sentence" (p. 226). On the second level, givenness represents the sense of saliency which is based on the assumes that "the hearer has or could appropriately have some particular thing or entity in his or her consciousness at the time of hearing the utterance" (p. 228). On the third level, givenness represents the sense of shared knowledge which is based on the assumes that "the hearer knows or can infer a particular thing" (p. 230). In addition, some other concepts, such as the theme and rhyme, the primacy and recency, are also related to what can be defined as given or new information in a discourse text. In this study, however, the distinction of given and new information is operationally defined purely in terms of semantic recoverability. The following examples (originally from Prince (1981)) illustrates how given and new information can be judged based on the criterion of recoverability.

Example1. We got some beer out of the trunk.

Example2. We got some picnic supplies out of the trunk.

In the sense of recoverability, picnic supplies in Example2 would be given information, whereas beer in Example1 would be new information, although, in some sense, they look a little bit different.

3.2. Language modeling

Taking a large corpus of text as input, Word2vec produces a vector space in which words that share common contexts in the corpus will be located in close proximity to one another. Taking the Skip-Gram approach in Word2vec as an example, the generation process of word vectors is shown in Figure 1. The goal of the Word2vec process is to learn similar word vectors for two words which have similar contexts. At last, the word vector corresponding to each word is determined.

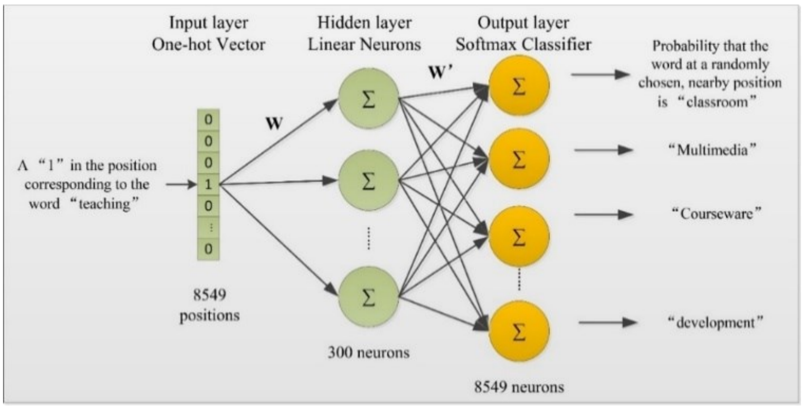

Figure 1. The generation process of word vectors (Skip-Gram approach)

4. Methodology

This study aimed to investigate the cognitive engagement, participation and discourse characteristics of the teachers in an online professional learning community. To achieve this goal, we formulated the first research question to understand teachers' nature of participation.
RQ1: What were the in-service teachers' participation characteristics in online discourse?

The second and third research questions focus on exploring teachers' cognitive engagement in the online professional learning community.
RQ2: How much new information was contained in each of the in-service teacher's post in online discourse?
RQ3: How relevant were the in-service teacher's posts to the discussion topic in online discourse?

The results would lead us to explore in-service teachers' online discourse characteristics through cluster analysis and answer the fourth research question.
RQ4. What were the in-service teachers' online discourse characteristics in terms of cognitive aspect?

4.1. Participants

In 2013, the Ministry of Education of China launched a five-year in-service teacher training program called the Information Technology Application Ability Enhancement Project for primary and secondary school teachers. They participated in the program in batches and the length of each training was about 120 hours. Many online learning platforms have been established and online learning was used in this project. In-service teachers needed to complete three tasks: watching video cases online, participating in topic-based online discussion, and submitting reflective diaries. The video cases were about how to use information technology tools to support classroom teaching. After watching the video cases, the in-service teachers in each community online discussed the contents of a video case based on their own teaching experience. In-service teachers participated in online discussion by posting, and each post is a reflection of the application of information technology in classroom teaching.

A total of 1,834 primary and secondary in-service teachers from a province in eastern China participated in the teacher professional development program for about a year. They were assigned to 31 online professional learning communities according to their subjects, with an average of 59.2 people per community. Table 1 shows the basic information of the 31 communities. After the program was completed, all online discourse data were collected and analyzed.

Table 1. Basic information of the 31 communities

Subject	Number of communities	Number of people	Subject	Number of communities	Number of people
Chinese	4	339	Mathematics	3	192
English	2	161	Geography	1	50
Chemistry	1	49	History	1	48
Art	2	48	Music	1	52
Biology	1	38	Ethics	2	95
Sports	2	111	Physics	2	54
Information technology	2	39	STEAM	3	266
Mixed group	1	42	Early childhood education	3	250

The purposive sampling method was used in this study to select one of the 31 communities to further analyze teachers' characteristics of cognitive engagement (Denzin & Lincoln, 1994). In the end, a mathematics online professional learning community was selected because it was representative (close to the average level of the population) in terms of gender, years of service and participation. The number of in-service teachers in the community was 60, and their average years of service was 19.3.

4.2. Discussion activities

In-service teachers' learning activities, such as watching video cases, participating in topic-based online discussion and submitting reflective diaries, were all carried out on the online learning platform. The interface of the online discussion activities is shown in Figure 2. The timeline of online discussion activities was as follows: (1) a teacher leader entered the name, duration, type and description of the topic-based discussion activity, and uploaded a video case, (2) all in-service teachers in the community watched the video case first and then participated in online discussions. All in-service teachers could participate in the discussion at any time. The basic learning requirement for in-service teachers was to contribute at least 5 posts during the online discourse session.

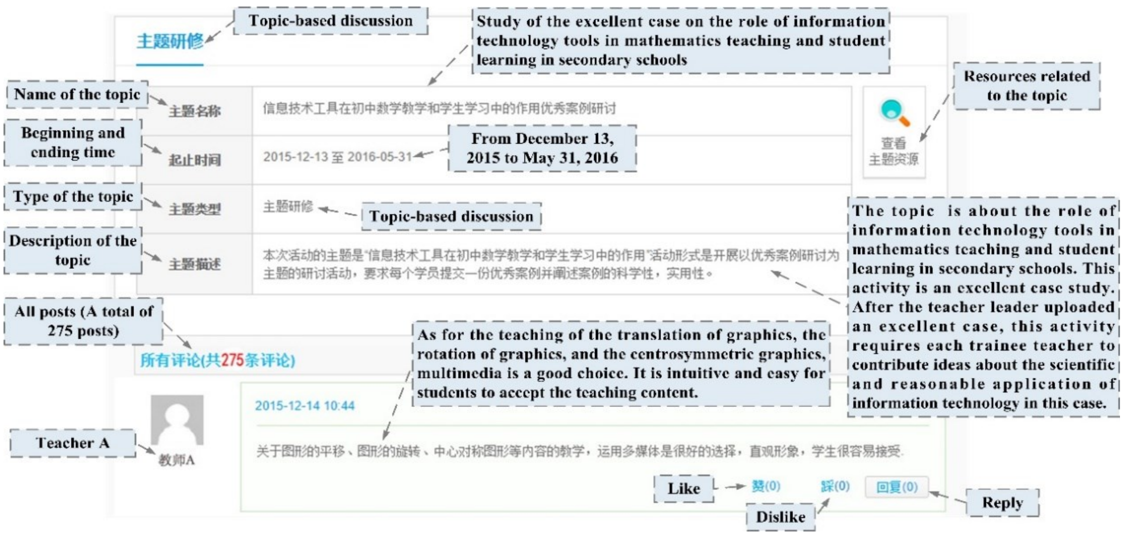

Figure 2. A screenshot of the online discussion interface

4.3. Data collection and analysis

After the teacher professional development program was completed, researchers used a web crawler tool called the Octopus to collect discourse data in all the 31 online professional learning communities. The research process consists of seven phases, as shown in Figure 3.

Phase 1: Data cleansing. All the collected data, including the name of the person who posted, the time and the text of the posts, were sorted out and stored in Excel files. All the in-service teachers' names were pseudonyms. The size of the Excel files is 3.09 M bytes. The number of posts of the data set "Application of information technology in classroom teaching" was 8418, and the average number of characters in Chinese per post was 39.47. The size of the set of lexical items in Chinese was 3859.

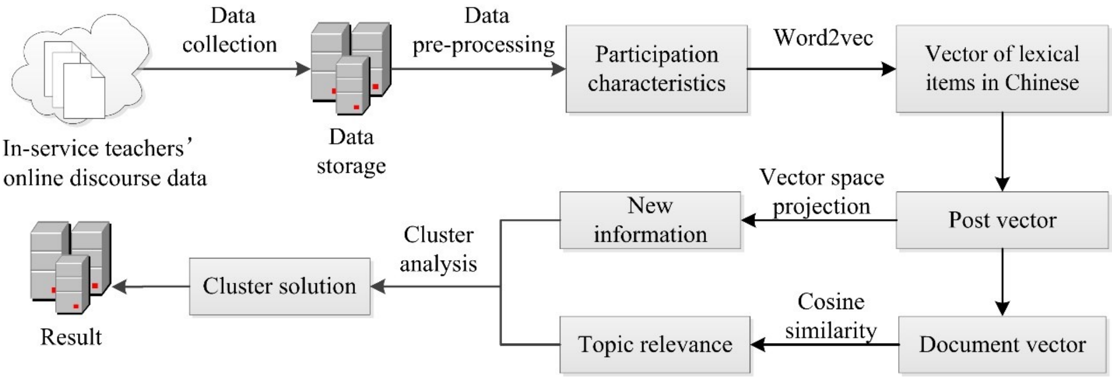

Figure 3. The process of data collection and analyses
Note. The black arrow represents the process of data analysis, and the rectangular box represents the result of data analysis.

Phase 2: Data preprocessing. The first author of this paper preprocessed the text data for all posts using the jiebaR package in R language, including Chinese segmentation, removing stop words, numbers, emojis, and special symbols. In particular, we used feature engineering methods (e.g., TFIDF, Singular Value Decomposition) to select professional terms (e.g., classroom teaching, multimedia technology) from the posts to build a special dictionary and load them into the jiebaR program to improve the accuracy of word segmentation. Secondly, the third author of this paper conducted the part-of-speech tagging using the jiebaR package in R language. In corpus linguistics, part-of-speech tagging is the process of marking up a word in a sentence (corpus) as corresponding to a particular part of speech such as nouns, verbs, adverbs, adjectives, pronouns (Kupiec, 1992). All the nouns in Chinese in the text data were retained and sorted to form a set of lexical items.

Phase 3: In-service teachers' participation characteristics in online discourse were analyzed. In addition, the time series analysis was implemented in R language version 3.5.0 using 'xts' packages.

Phase 4: Lexical vector generation. The Chinese lexical items obtained in Phase 2 were sent to Word2vec to produce lexical vectors, which was implemented in R language version 3.5.0 using the approach of Skip-Gram in the text2vec package. After each lexical item was represented as a vector, each post in an online discourse was represented by an average of the vector sum of the lexical items it contains (suppose each post contained n lexical items), see equation 1. The results of this phase were the generation of lexical vectors and post vectors.

$$Vector_{post} = \frac{\sum_1^n Vector_{lexical\ item}}{n} \quad \text{Equation 1}$$

Phase 5: Calculation of new information. Each post was divided into two parts in terms of the information it contained. One part represented the given information and the other part represented the new information. Referring to the method used by Hu et al. (2003), all previous posts of the current post (post A) spanned to a vector space (see equation 2). The given information in post A was represented by the projection of the vector of post A onto the vector space, see equation 3. The result of equation 3 was a numerical value, which represented the given information contained in post A.

$$\text{Vector space} = \text{span}\ \{\overrightarrow{p_1}, \overrightarrow{p_2}, \ldots, \overrightarrow{p_{t-1}}\} \quad \text{Equation 2}$$

$$Information_{Given} = \|Projection_{vector\ space}(\overrightarrow{p_t})\| \quad \text{Equation 3}$$

Equation 3 is the norm of the projected vector. For example, in order to calculate the given information of the Post 3, the first two posts (post 1 and post 2) were spanned to a vector space (see Figure 4), and the projection of the vector of the post 3 onto the vector space represents the given information of post 3.

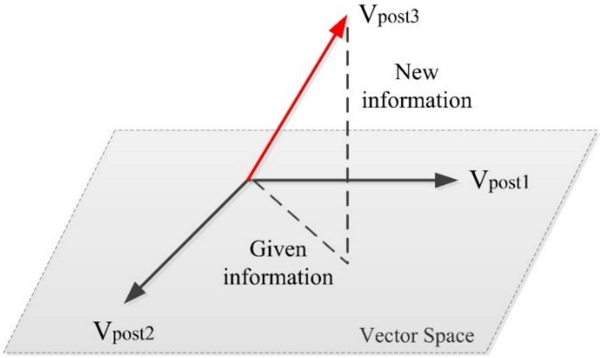

Figure 4. Projecting the vector of the third post to a vector space

The new information was represented by the projection of the vector of post 3 onto the Orthogonal Complement Space of the Vector Space (OCSVS), see equation 4. The result of equation 4 was a numerical value, which represented the new information contained in post 3.

$$Information_{New} = ||Projection_{OCSVS}(\overrightarrow{p_t})|| \quad \text{Equation4}$$

Then, the new information contained in each post was standardized, see equation 5.

$$\text{New}\ (post_i) = \frac{Information_{New}}{Information_{Given} + Information_{New}} \quad \text{Equation 5}$$

The mean value, standard deviation of the new information contained in every post and the evolution of the new information over time were also calculated. The new information analysis was implemented in R language version 3.5.0 using Matrix and limSolve packages. The Matrix package contains numerous methods for and operations on the matrix, including triangular, symmetric, and diagonal matrices, both dense and sparse and with pattern, logical and numeric entries (Bates & Maechler, 2019). The limSolve package was used to solve the linear inverse matrix of a given matrix (Soetaert, van den Meersche, & van Oevelen, 2017).

Phase 6: Calculation of topic relevance. The correlation between a post vector and the document vector (an average of the vector sum of all the posts in the discourse) indicated the extent to which the post was related to the content of the entire online discourse (Dascalu, Trausan-Matu, Dessus, & Mcnamara, 2015). Suppose an online discourse contained m posts, the representation of the document vector is shown in equation 6.

$$Vector_{document} = \frac{\sum_1^m Vector_{post}}{m} \quad \text{Equation 6}$$

The calculation of the correlation between the post vector and the document vector is shown in equation 7.

$$\text{Correlation} = \cos(\theta) = \frac{\boldsymbol{V_{post}} \cdot \boldsymbol{V_{document}}}{||\boldsymbol{V_{post}}||\,||\boldsymbol{V_{document}}||} = \frac{\sum_{i=1}^{n} \boldsymbol{V_{post}}_i \boldsymbol{V_{document}}_i}{\sqrt{\sum_{i=1}^{n} \boldsymbol{V_{post}}_i^2}\sqrt{\sum_{i=1}^{n} \boldsymbol{V_{document}}_i^2}} \quad \text{Equation 7}$$

In equation 7, cos represents the cosine distance between two vectors, which was widely used to calculate the correlation degree of two vectors. Next, the mean and standard deviation of the topic relevance of all posts and the evolution of the topic relevance over time were analyzed.

Phase 7: Identification of the cluster of posts and in-service teachers' online discourse characteristics. Before the cluster analysis, the values of new information and topic relevance were standardized, which was a common step in cluster analysis (Hastie, Tibshirani, & Friedman, 2013). Many methods can be used to implement cluster analysis, and the agglomerative hierarchical clustering method was adopted in this study for solving similar problems (e.g., Kovanović, Gašević, Joksimović, Hatala, Adesope, 2015; Wise, Speer, Marbouti, Hsiao, 2013). The cluster analysis procedure was implemented in R language version 3.5.0 using 'stats' packages. The process of cluster analysis included two steps: the first step was to determine the optimal number of clusters, and the second step was to implement cluster analysis. Two indicators, the cluster centroids and the average silhouette index, were evaluated to select the optimal number of clusters. The clustering dendrogram is shown in Figure 5. When implementing cluster analysis, the agglomerative hierarchical clustering method was used with the

Euclidean distance and Ward's agglomeration criteria. After determining the cluster of each post, we further identified in-service teachers' online discourse characteristics.

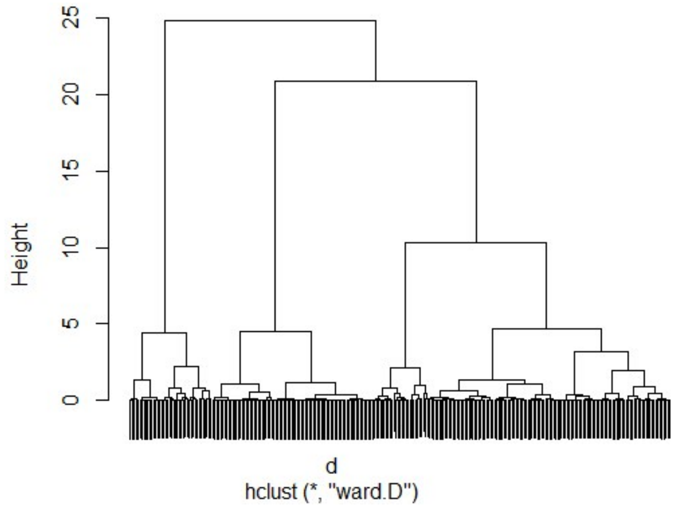

Figure 5. Dendrogram of post clustering

5. Results

5.1. Algorithm evaluation

In order to ensure the validity and reliability of the proposed algorithm, one coder in educational technology and one in teacher professional development independently rated each post on a five-point Likert scale ranging from 1 (lowest) to 5 (highest) on two indicators (new information and topic relevance). Coding results showed that the inter-rater reliability Kappa was 0.667 ($p < .01$), which demonstrated fair to good reliability. Then, two coders discussed the differences in codes and combined the results. Cohen's Kappa coefficient (K) was calculated to measure the agreement between the proposed algorithm and human in two indicators. The results showed the moderate agreement rate between the proposed algorithm and human: the coefficient (K) of new information was *0.618* and the coefficient (K) of topic relevance was *0.775*.

5.2. What were the in-service teachers' participation characteristics in online discourse?

The average contribution of each in-service teacher was 4.59 posts, and the average length of posts was 39.47 characters in Chinese. Figure 6 shows the temporal characteristics of the teachers' online discourse. In-service teachers contributed a large number of posts in the early stages of online discourse. As time went by, the number of posts contributed by the in-service teachers decreased. By the end of the online discourse, in-service teachers' contribution increased slightly.

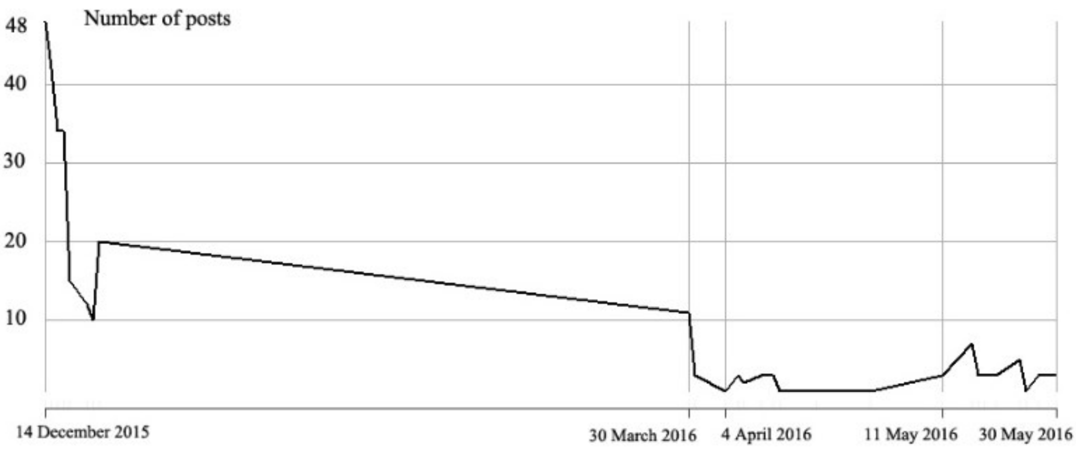

Figure 6. Change of the number of posts over time

5.3. How much new information was contained in each of the in-service teacher's post in online discourse?

In-service teachers of the mathematics community contributed a total of 275 posts, and the average length of posts was 41.138 characters in Chinese. The minimum value of new information (= 0) indicated that some posts did not contain any new information, while the maximum value (= 1) indicated that the information contained in a post was entirely new information. The mean of new information contained in the posts was 0.221.

Figure 7 further shows how the new information contained in the posts changed over time. The dotted line in Figure 7 represented the mean amount of new information. In-service teachers' posts contained much new information in the early stages of online discourse. As time went on, the new information contained in the posts gradually decreased. It was worth noting that posts with new information of 0 appeared at all stages of the online discourse and appeared intensively in the later stage of the online discourse.

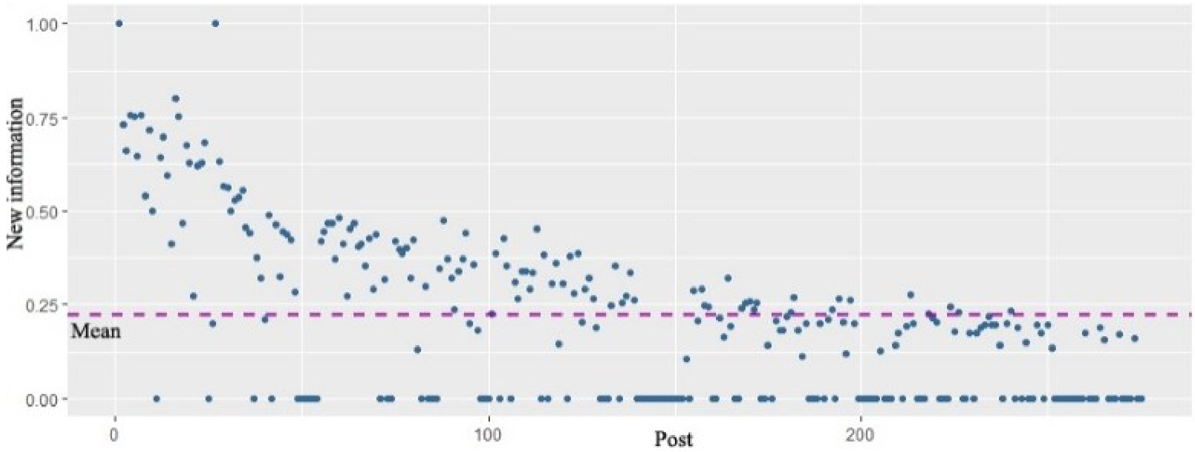

Figure 7. Change of the amount of new information over time

5.4. How relevant were the in-service teacher's posts to the discussion topic in online discussion discourse?

The minimum value of topic relevance (= 0) indicated that a post had nothing to do with the discussion topic, while the maximum value (= 0.943) indicated that the content of a post was highly relevant to the discussion topic. The mean of relevance was 0.708, indicating that most posts were very relevant to the discussion topic.

Figure 8 further shows how the topic relevance changed over time. The dotted line in Figure 8 represents the mean of the topic relevance. The content of in-service teachers' posts was highly related to the discussion topic, which appeared in all stages of online discourse. It was rare that the content of in-service teachers' posts was completely unrelated to the discussion topic.

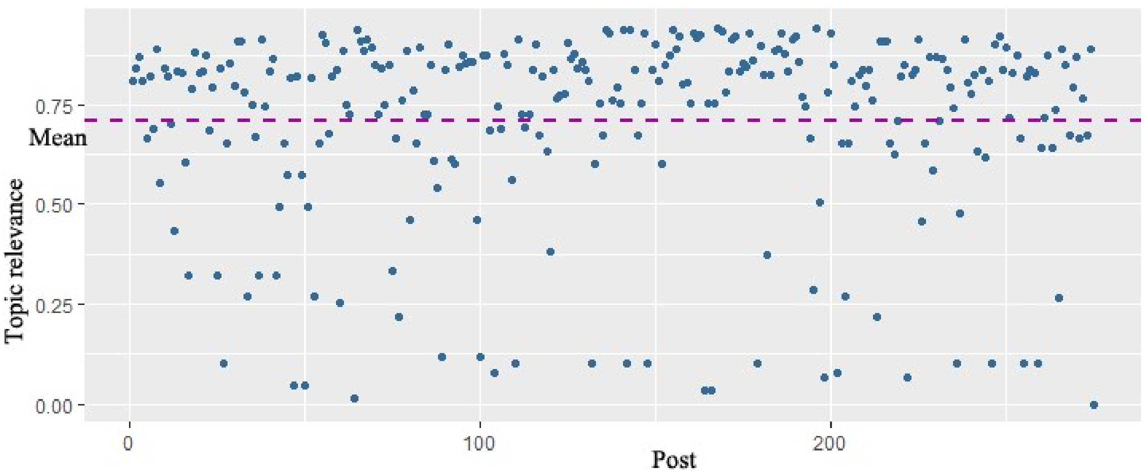

Figure 8. Change of the topic relevance over time

5.5. What were the in-service teachers' online discourse characteristics in terms of cognitive aspect?

In order to determine the optimal number of clusters, we compared the differences of the cluster centroids and the average Silhouette index of different clustering solutions. Although the two-cluster solution had the highest average Silhouette index, we decided to use the three-cluster solution, where one cluster of the two-cluster solution was divided into two smaller sub-clusters for the following reasons. First, in the two-cluster solution, one cluster included 83.6% of the posts (see Table 2) and a basic distinction between related or not related to the discussion topic which had very little practical significance. Second, we can see the difference between the two sub-clusters from the cluster centroids of the three-cluster solution, which are associated with in-service teachers' cognitive characteristics and were aligned with literatures on learners' cognitive engagement (e.g., Ding, Er, Orey, 2018; Ding, Kim, & Orey, 2017). Based on the above considerations, this study used the three-cluster solution (see Table 3).

Table 2. Characteristics of the two-cluster solution

Cluster #	Cluster centroids (Relate, New)	Number of posts (%)	Characteristics
Cluster 1	(0.803, 0.220)	230 (83.6%)	High relevance, little new information
Cluster 2	(0.221, 0.221)	45 (16.4%)	Low relevance, little new information

Table 3. Characteristics of the three-cluster solution

Cluster #	Cluster centroids (Relate, New)	Number of posts (%)	Characteristics
Cluster 1	(0.764, 0.484)	75 (27.3%)	High relevance, much new information
Cluster 2	(0.810, 0.104)	160 (58.2%)	High relevance, little new information
Cluster 3	(0.193, 0.190)	40 (14.5%)	Low relevance, little new information

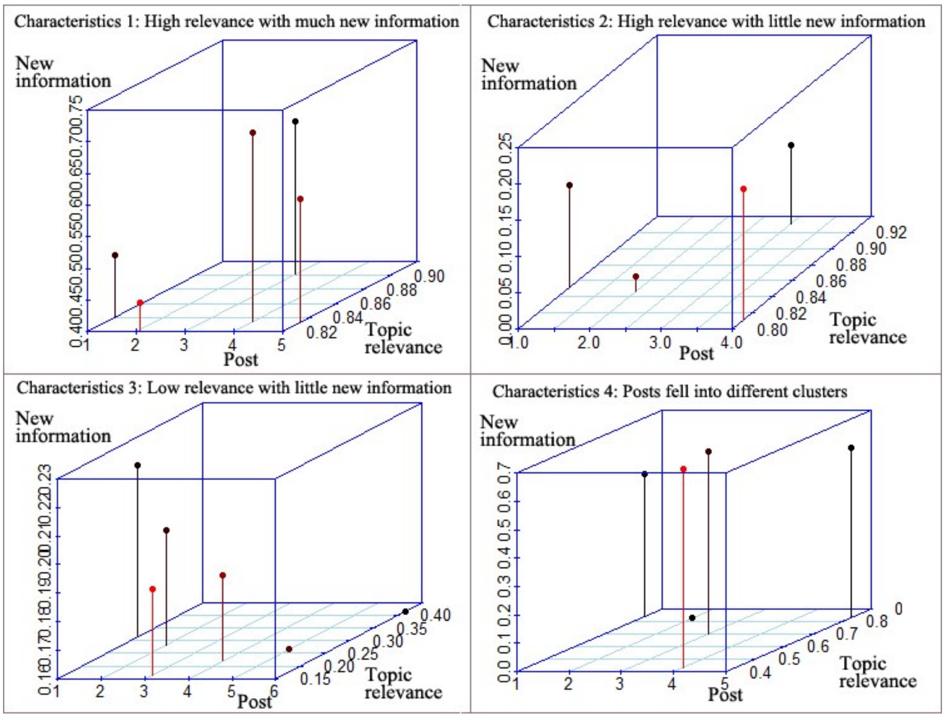

Figure 9. Four types of online discourse characteristics

Combining the results of Figures 7 and 8, it is easy to observe the following results: cluster 1 appeared primarily in the early stage of online discourse; cluster 2 existed in the whole process of online discourse; and cluster 3 occasionally appeared in the online discourse. Four typical types of in-service teachers' online discourse characteristics are shown in Figure 9. The in-service teachers with the first type of characteristics always contributed posts that were related to online discussion topics and contained much new information. However, the in-service teachers with the second type of characteristics always contributed posts that were related to online discussion topics but with little new information. In-service teachers with the third type of characteristics contributed many posts, but most of the posts were of low topic relevance and little new information. In terms of the posts contributed by in-service teachers with the fourth type of characteristics, some posts were highly related to the discussion topic but with zero new information, while some with low topic relevance but much new information included.

6. Discussion

In-service teachers contributed an average of 4.59 posts and the average length of each post was 39.47 characters in Chinese. Moreover, in-service teachers' contribution mainly appeared in the early stage of online discourse. This finding was inconsistent with other studies. In Liu's (2012) study, an online video-case discussion community was used to foster in-service teachers' professional development and the results showed that in-service teachers posted an average of 13.72 messages per person. In another study, teachers' contribution across the 6 weeks of online synchronous discussion was quite consistent (Chen, Chen, & Tsai, 2009). Two possible reasons may be related to in-service teachers' low and unbalanced contribution. The first was the scoring rules of the teacher professional development program. Many in-service teachers contributed 5 posts in succession in order to complete the learning tasks as soon as possible, and stopped posting once the basic learning requirements were met. The second reason was due to in-service teachers' busy work schedule. When faced with the contradiction between work and learning, in-service teachers often expected to finish their learning task as early as possible. Due to the constant changes of participants, a good discourse culture in the online professional learning community had not been formed and it was difficult for in-service teachers to have high-level reflection and collaborative knowledge construction (Lee & Brett, 2015).

New information contained in learners' contribution is conductive to the accumulation and continuation of ideas (Keles, 2018), as well as the promotion of collaborative knowledge building. The results indicated that the online discourse activity presented in this study had encouraged and facilitated in-service teachers' active and constructive contributions related to classroom practice to some extent. In-service teachers' posts contained much new information in the early stages of the online discourse, while in the later stage of the online discourse, the amount of new information became very small. From a technical perspective, as the online discourse progressed, most of the information related to the topic had been discussed. If an in-service teacher repeated something similar, there would be little new information in the post. However, from a pedagogical perspective, this situation should be avoided. An effective online professional learning community relies on the continuous, active participation and contribution of team members to solve emerging problems and generate new knowledge. From this point of view, a good community atmosphere had not been formed in the mathematics community. Two possible reasons may be related to this situation. On the one hand, although online discourse environments facilitated collaboration among in-service teachers, they still lacked necessary discussion skills, such as reviewing existing information, or comparing inconsistencies between posts. On the other hand, the online discourse platform could not give hints to in-service teachers about what had been discussed. Therefore, the necessary scaffolding and technical support can help the formation and development of the online professional learning community (e.g., Marbouti & Wise, 2016).

Whether the content of posts was related to the discussion topic not only influences the cohesion of the online discourse, but also influences the shared culture of the online learning community (Tsiotakis & Jimoyiannis, 2016). The results showed that most of in-service teachers' posts were relevant to the discussion topic, indicating that the cohesion of in-service teachers' online discourse was high. This finding was correlated with previous studies. The teachers in the online professional learning community actively exchanged information, and provided and sought advices (e.g., Lantz-Andersson et al., 2018; Trust, Krutka, & Carpenter, 2016). In an organized online discourse, teachers concentrated on the discussion topic much more frequently (e.g., Keles, 2018). This behavior is of value to the online professional learning community, as it indicated that in-service teachers were actively thinking about the topic of online discourse and related it with their own professional knowledge and experience.

The results of cluster analysis showed three different groups of posts. The second group of posts, those highly related to the discussion topic but containing little new information, accounted for the largest number. This group of posts appeared mainly in the middle and latter stages of the online discourse. The first group of posts, which closely related to the discussion topic and containing much new information, came in the second place. This group of posts appeared mainly in the early stage of the online discourse. The posts that were less relevant to the discussion topic and contained little new information appeared sporadically in the online discourse. The results of the cluster analysis suggested that in-service teachers' behaviors and the extent that they contributed new knowledge might be influenced by teachers' tight schedules, the activity design of the online discussion, and the support strategies of the online learning platform. Therefore, it is necessary to design different intervention strategies for in-service teachers with different cognitive characteristics in the OPLCs. For example, for teachers who consistently contribute posts that are highly relevant to the topic and contain a lot of new information, more rewards and opportunities to show their prestige are necessary. For in-service teachers who consistently contribute posts that are unrelated to the topic and have no new information, they need to be prompted. In addition, for teachers who consistently contribute posts that are relevant to the topic but contain less new information, they need to be encouraged to read their colleagues' posts and contribute innovative ones, for instance.

This study had two major contributions. The first one was that it proposed a completely automatic evaluation approach for detecting in-service teachers' cognitive engagement in the OPLC. The method used in this study was highly automatic. It helped introduce an intelligent supervision system in the online learning community to guide and intervene learners' online discourse. If learners contribute little new information (e.g., always repeating what has been discussed), community information sharing culture, problem solving process, and collaborative knowledge construction will be affected. At this point, it is necessary for the online learning platform to remind learners of what is being discussed and encourage learners to think from different perspectives. On the other hand, if the posts contributed by learners are off-topic or the topic relevance is very low, it will affect the cohesion of the discussion among community members. At this point, it is necessary for the online learning platform to intelligently determine the topic relevance of the posts. The method presented in this paper established a technical basis for the implementation of the intervention. The second was that it identified in-service teachers' online discourse characteristics from the cognitive aspects of interaction, which was helpful to identify the learner roles emerging in online learning settings, such as core members (those who actively contribute new information and have a high degree of topic relevance) and marginal roles (those who contribute a small amount of posts and have little new information). Further design of intervention strategies for different roles, such as incentive strategies for core members and prompt strategies for marginal members, will help core and marginal members of the online learning communities to better monitor and regulate the online learning process.

This study had a few important implications. First, when evaluating in-service teachers' learning engagement in an online learning community, we should pay attention to both the quantity and quality of their contribution. Previous studies used teachers' behaviors, such as the number of posts, replies, and resources downloaded to represent teachers' learning engagement in online discussions and did not notice that many posts with no new information appeared (Xing & Gao, 2018). Therefore, just counting in-service teachers' behaviors did not represent their real performance in online discussions. Second, this study had important implication for building an OPLC. Since in-service teachers in the OPLC had different online discourse characteristics, teacher educators need to design different supervision and intervention strategies. Third, this study provided a bridge between a theoretical description of leaner's interaction in the online learning environment and a computational method of the interaction. Alternative analyses or interpretations were definitely possible given the types of analysis conducted in this study. For example, the computational method could be used to evaluate the performance of different online discussion pedagogy in online learning settings (e.g., MOOCs), or the learning performance of different learning groups under the same online discussion pedagogy.

7. Conclusion, limitations, and future study

This study integrated learning analytics and educational data mining to build a workflow for solving a key issue in the OPLC: to detect in-service teachers' cognitive engagement in online discussion activity. Results showed that in-service teachers' posts contained more new information in the early stages of online discourse and the average amount of new information contained in each post was 0.221. When using the calculation method of topic relevance proposed in this paper, most of in-service teachers' posts were highly relevant to the discussion topic. In addition, cluster analysis generated three different groups of posts with unique characteristic. Based on the three groups of posts, it was easy to identify in-service teachers' online discourse characteristics. The workflow and research results presented in this study were not only applicable to the discourse analysis in the OPLC, but also contribute to the establishment of an intelligent online learning environment (e.g., MOOCs) which can automatically monitor and provide feedback on learners' discourse.

There were two major limitations in this study. Firstly, the research method proposed in this paper tended to ignore some small patterns that may be important. For example, the posts that were not highly relevant to the discussion topic but contained a lot of new information were ignored. In fact, such posts may contain innovative ideas that were simply not appreciated by the community members. Secondly, this study only analyzed the online discourse characteristics of a mathematics community. When researchers and teachers apply the conclusions of this study, they need to consider the context of the study. Future research directions are to: (1)

explore the relationship between the quantity of new information in online discourse and in-service teachers' perception, collaborative learning quality and learning score; (2) design different intervention strategies for in-service teachers with different types of online discourse characteristics; and (3) deeply understand the factors influencing in-service teachers' online discourse characteristics through interviews and questionnaire surveys.

Acknowledgement

This paper was supported by the National Natural Science Foundation of China (Grant No. 62077016). This work was also conducted as part of a research project that was funded by the School of Educational Information Technology, Central China Normal University of China (Project No. CCNUSEIT202008).

Statements on open data, ethics and conflict of interest

Data can be accessed by contacting the author (saved in a personal repository). Ethical approvals were gained from the hosting institution. This research has no conflicts of interest.

References

Atapattu, T., Thilakaratne, M., Vivian, R., & Falkner, K. (2019). Detecting cognitive engagement using word embeddings within an online teacher professional development community. *Computers & Education*, *140*, 103594.

Baran, B., & Cagiltay, K. (2010). The Dynamics of online communities in the activity theory framework. *Educational Technology & Society*, *13*(4), 155-166.

Bates, D., & Maechler, M. (2019). *Matrix: Sparse and Dense Matrix Classes and Methods*. Retrieved from https://cran.r-project.org/web/packages/Matrix/index.html

Bing, L., Niu, Z.-Y., Li, P., Lam, W., & Wang, H. (2018). Learning a unified embedding space of web search from large-scale query log. *Knowledge-Based System*, *150*, 38-48.

Cerisara, C., Král, P., & Lenc, L. (2018). On the effects of using Word2vec representations in neural networks for dialogue act recognition. *Computer Speech & Language*, *47*, 175-193.

Chafe, W. L. (1976). Givenness, contrastiveness, definitensess, subjects, topics, and point of view. In C. N. Li (Eds.), *Subject and Topic* (pp. 25-55), New York, NY: Academic Press.

Chen, J.-M., Chen, M.-C., & Sun, Y., S. (2010). A Novel approach for enhancing student reading comprehension and assisting teacher assessment of literacy. *Computers & Education*, *55*(3), 1367-1382.

Chen, Y., Chen, N.-S., & Tsai, C.-C. (2009). The use of online synchronous discussion for web-based professional development for teachers. *Computers & Education*, *53*(4), 1155-1166.

Chen, H.-L., Fan, H.-L., & Tsai, C.-C. (2014). The Role of community trust and altruism in knowledge sharing: An Investigation of a virtual community of teacher professionals. *Educational Technology & Society*, *17*(3), 168-179.

Cho, H. (2016). Under co-construction: An Online community of practice for bilingual pre-service teachers. *Computers & Education*, *92-93*, 76-89.

Dascalu, M., Trausan-Matu, S., Dessus, P., & McNamara, D. (2015). Discourse cohesion: A Signature of collaboration. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge* (pp. 350-354). New York, NY, United States: Association for Computing Machinery. doi:10.1145/2723576.2723578

Denzin, N. K., & Lincoln, Y. S. (1994). *Handbook of qualitative research*. Thousand Oaks, CA: Sage Publications.

Ding, L., Er, E., & Orey, M. (2018). An Exploratory study of student engagement in gamified online discussions. *Computers & Education*, *120*, 213-226.

Ding, L., Kim, C., & Orey, M. (2017). Studies of student engagement in gamified online discussions. *Computers & Education*, *115*, 126-142.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research*, *74*(1), 59-109.

Garrison, D. R., Anderson, T., & Archer, W. (1999). Critical inquiry in a text-based environment: Computer conferencing in higher education. *The Internet and Higher Education*, *2*(2-3), 87-105.

GaŠević, D., Joksimović, S., Eagan, B. R., & Shaffer, D. W. (2019). SENS: Network analysis to combine social and cognitive perspectives of collaborative learning. *Computers in Human Behavior*, *92*, 562-577.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*, *17*(4), 397-431.

Halliday, M. A. K. (1967). Notes on transitivity and theme in English. *Journal of Linguistics, 3*, 199-244.

Harris, Z. (1952). Discourse analysis. *Language*, *28*(1), 1-30.

Hashimoto, T. B., Alvarez-Melis, D., & Jaakkola, T. S. (2015). Word, graph and manifold embedding from Markov processes. *Computing Research Repository.* Retrieved from https://arxiv.org/pdf/1509.05808.pdf

Hastie, T. J., Tibshirani, R. J., & Friedman, J. H. (2013). *The Elements of statistical learning: Data mining, inference, and prediction*. New York, NY: Springer.

Heilporn, G., & Lakhal, S. (2020). Investigating the reliability and validity of the community of inquiry framework: An Analysis of categories within each presence. *Computers & Education*, *145*, 103712.

Hu, X., Cai, Z., Louwerse, M., Olney, A., Penumatsa, P., & Graesser, A. (2003). A Revised algorithm for latent semantic analysis. In *Proceedings of the 18th international joint conference on Artificial intelligence* (pp. 1489-1491). San Francisco, CA: Morgan Kaufmann Publishers Inc.

Keles, E. (2018). Use of Facebook for the community services practices course: Community of inquiry as a theoretical framework. *Computers & Education*, *116*, 203-224.

Kim, D., Park, Y., Yoon, M., & Jo, I.-H. (2016). Toward evidence-based learning analytics: Using proxy variables to improve asynchronous online discussion environments. *Internet and Higher Education*, *30*, 30-43.

Kovanović, V., Gašević, D., Joksimović, S., Hatala, M., & Adesope, O. (2015). Analytics of communities of inquiry: Effects of learning technology use on cognitive presence in asynchronous online discussions. *The Internet and Higher Education*, *27*, 74-89.

Kupiec, J. (1992). Robust part-of-speech tagging using a hidden MarKov model. *Computer Speech & Language*, *6*(3), 225-242.

Lantz-Andersson, A., Lundin, M., & Selwyn, N. (2018). Twenty years of online teacher communities: A Systematic review of formally-organized and informally–developed professional learning groups. *Teaching and Teacher Education*, *75*, 302-315.

Lee, K., & Brett, C. (2015). Dialogic understanding of teachers' online transformative learning: A Qualitative case study of teacher discussions in a graduate-level online course. *Teaching and Teacher Education*, *46*, 72-83.

Li, J., Li, J., Fu, X., Masud, M. A., & Huang, J. Z. (2016). Learning distributed word representation with multi-contextual mixed embedding. *Knowledge-Based Systems*, *106*, 220-230.

Liu, M.-H. (2012). Discussing teaching videocases online: Perspectives of preservice and inservice EFL teachers in Taiwan. *Computers & Education*, *59*(1), 120-133.

Marbouti, F., & Wise, A. F. (2016). Starburst: A New graphical interface to support purposeful attention to others' posts in online discussions. *Educational Technology Research and Development*, *64*, 87-113.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems*. Red Hook, NY, United States: Curran Associates Inc. Retrieved from https://arxiv.org/pdf/1310.4546.pdf

Peng, X., & Xu, Q. (2020). Investigating learners' behaviors and discourse content in MOOC course reviews. *Computers & Education*, *143*, 103673. doi:10.1016/j.compedu.2019.103673

Polizzi, S., Head, M., Barrett-Williams, D., Ellis, J., Roehrig, G. H., & Rushton, G. T. (2018). The Use of teacher leader roles in an online induction support system. *Teaching and Teacher Education*, *75*, 174-186.

Prestridge, S. (2019). Categorising teachers' use of social media for their professional learning: A Self-generating professional learning paradigm. *Computers & Education*, *129*, 143-158.

Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-255), New York, NY: Academic Press.

Sinoara, R. A., Camacho-Collados, J., Rossi, R. G., Navigli, R., & Rezende, S. O. (2019). Knowledge-enhanced document embeddings for text classification. *Knowledge-Based Systems*, *163*, 955-971.

Soetaert, K., van den Meersche, K., & van Oevelen, D. (2017). *limSolve: Solving Linear Inverse Models*. Retrieved from https://cran.r-project.org/web/packages/limSolve/

Sung, Y.-T., Liao, C.-N., Chang, T.-H., Chen, C.-L., & Chang, K.-E. (2016). The Effect of online summary assessment and feedback system on the summary writing on 6th graders: The LSA-based technique. *Computers & Education*, *95*, 1-18.

Trust, T., Krutka, D. G., & Carpenter, J. P. (2016). "Together we are better": Professional learning network for teachers. *Computers & Education*, *102*, 15-34.

Tsai, I.-C. (2012). Understanding social nature of an online community of practice for learning to teach. *Educational Technology & Society*, *15*(2), 271-285.

Tsiotakis, P., & Jimoyiannis, A. (2016). Critical factors towards analyzing teachers' presence in online learning communities. *The Internet and Higher Education*, *28*, 45-58.

Vieira, C., Parsons, P., & Byrd, V. (2018). Visual learning analytics of educational data: A Systematic literature review and research agenda. *Computers & Education*, *122*, 119-135.

Wang, X., Wen, M., & Rosé, C. P. (2016). Towards triggering higher-order thinking behaviors in MOOCs. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge* (pp. 398–407). New York, NY, United States: Association for Computing Machinery. doi:10.1145/2883851.2883964

Wise, A. F., Speer, J., Marbouti, F., & Hsiao, Y. -T. (2013). Broadening the notion of participation in online discussions: Examining patterns in learners' online listening behaviors. *Instructional Science*, *41*(2), 323-343.

Wu, P., Yu, S., & Wang, D. (2018). Using a learner-topic model for mining learner interests in open learning environments. *Educational Technology & Society*, *21*(2), 192-204.

Xie, K., Heddy, B. C., & Greene, B. A. (2019). Affordances of using mobile technology to support experience-sampling method in examining college students' engagement. *Computers & Education*, *128*, 183-198.

Xing, W., & Gao, F. (2018). Exploring the relationship between online discourse and commitment in Twitter professional learning communities. *Computers & Education*, *126*, 388-398.

Yang, S.-H. (2016). Conceptualizing effective feedback practice through an online community of inquiry. *Computers & Education*, *94*, 162-177.

Zhang, S., Liu, Q., Chen, W., Wang, Q., & Huang, Z. (2017). Interactive networks and social knowledge construction behavioral patterns in primary school teachers' online collaborative learning activities. *Computers & Education*, *104*, 1-17.

Zhang, D., Xu, H., Su, Z., & Xu, Y. (2015). Chinese comments sentiment classification based on Word2vec and SVM. *Expert Systems with Applications*, *42*(4), 1857-1863.