Education and Information Technologies
https://doi.org/10.1007/s10639-023-11633-8

Investigating online learners' knowledge structure patterns by concept maps: A clustering analysis approach

Xiuling He[1] · **Jing Fang**[2] · **Hercy N. H. Cheng**[3] · **Qibin Men**[1] · **Yangyang Li**[2]

Received: 14 October 2022 / Accepted: 30 January 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

A deep understanding of the learning level of online learners is a critical factor in promoting the success of online learning. Using knowledge structures as a way to understand learning can help analyze online students' learning levels. The study used concept maps and clustering analysis to investigate online learners' knowledge structures in a flipped classroom's online learning environment. Concept maps ($n = 359$) constructed by 36 students during one semester (11 weeks) through the online learning platform were collected as analysis objects of learners' knowledge structures. Clustering analysis was used to identify online learners' knowledge structure patterns and learner types, and a non-parametric test was used to analyze the differences in learning achievement among learner types. The results showed that (1) there were three online learners' knowledge structure patterns of increasing complexity, namely, spoke, small-network, and large-network patterns. Moreover, online learners with novice status mostly had spoke patterns in the context of flipped classrooms' online learning. (2) Two types of online learners were found to have different distributions of knowledge structure patterns, and the complex knowledge structure type of learners exhibited better learning achievement. The study explored a new way for educators to analyze knowledge structures by data mining automatically. The findings provide evidence in the online learning context for the relationship between complex knowledge structures and better learning achievement while suggesting the existence of inadequate knowledge preparedness for flipped classroom learners without a special instructional design.

Keywords Knowledge structure · Concept map · Clustering analysis · Online learning · Flipped classroom

✉ Jing Fang
fangjing@ccnu.edu.cn

Extended author information available on the last page of the article

Published online: 20 February 2023

1 Introduction

Online learning has gradually become a very important way of learning, especially as the coronavirus disease (COVID-19) pandemic has forced higher-education activities in many countries around the world to utilize online learning or a combination of face-to-face learning and online learning (García-Morales et al., 2021). This sudden change means online learning is becoming a regular education model (Carolan et al., 2020). Learners' conceptual knowledge and academic achievement in online learning are often used as primary learning outcomes (Wei et al., 2021). These learning outcomes are assessed most often by tests, exams, etc. (Deng et al., 2019). Researchers have pointed out that understanding students' learning only by scores in most cases can only reveal the retelling of knowledge, which cannot deeply reveal the learners' learning status (Cheng et al., 2013). However, having a real comprehensive understanding of learners' learning status is crucial for educators to implement effective teaching design (Chen et al., 2021). Knowledge structures represent the organization and integration of concepts in personal memory and can reflect the form of knowledge organization of learners (Jonassen et al., 1993). Knowledge structure measures can assess learners' conceptual understanding (Weinerth et al., 2014), identify knowledge loss and misunderstanding (Kim et al., 2019), predict classroom learning and achievement (Lopez et al., 2014), distinguish learners' learning style (Hay, 2007), and describe cognitive status changes during problem solving (Hung & Lin, 2015). Recent studies have also used knowledge structures to evaluate learners' higher-order thinking abilities, such as systematic thinking (Khajeloo & Siegel, 2022). Therefore, understanding knowledge structures is considered a more comprehensive way of understanding learner cognition (Kim et al., 2019). Knowledge structures can be elicited and analyzed through a variety of visualization methods (Clariana et al., 2014; Eseryel et al., 2013; Jonassen et al., 1993; Tang & Clariana, 2017). Concept maps place more emphasis on the extraction of concepts and the establishment of knowledge relations, which is considered a suitable representation approach of knowledge structures in the actual teaching process (Jonassen & Marra, 1994). Computer-supported concept maps provide a convenient means to construct maps while recording the resulting concept maps through standardized data (Farrokhnia et al., 2019). This visualization method of knowledge structures provides the basis for the automated analysis and evaluation of knowledge structures (Weinerth et al., 2014). Current learning analysis techniques are widely used in online learning to improve teaching, learning, and educational decision-making (Brown et al., 2022). Researchers use the data mining approach to analyze learning behaviors (Tlili et al., 2021) and interactive discourse (Liu et al., 2022) to find education and learning rules. However, the data mining approach is rarely used to investigate online learners' knowledge structures.

To fill these research gaps, we used a self-developed online learning platform that supported concept mapping to collect online learners' knowledge structures and identified knowledge structure patterns through the data mining approach. Our study explored an automatic and fast method to analyze knowledge structures and find identifiable patterns of knowledge structure in online learners. The findings provided support for educators to better understand the connotation of learners' knowledge structures, and help them conduct online instructional design and evaluation.

2 Literature review

2.1 Knowledge structures as a way to understand learning

Knowledge structures are a collection of interrelated facts about specific topics within a given domain, also known as structural knowledge or cognitive structure, which provide connections between procedural and declarative knowledge and represent learners' ability to apply what they know (Clariana et al., 2013; Jonassen et al., 1993). Meaningful learning theory defines learning as occurring when individuals absorb new information in their preexisting knowledge structures and when learners choose to link new information to prior knowledge (Ausubel, 1963). Moreover, meaningful learning also leads to a more complex knowledge structure (Novak, 2010). Knowledge structures have been found to measure students' understanding or conceptual level in teaching research based on many subjects, such as science education (Schaal et al., 2010), biology (Wadouh et al., 2014), mathematics (Gogus, 2013), physics (Koponen & Pehkonen, 2010), and medicine (Hung & Lin, 2015). Previous studies have found that there are significant differences between the knowledge structures of experts and those of novices (Chi et al., 1981). For example, Gogus (2013) found differences in the conceptualization of complex mathematical problems between novices and experts by comparing their knowledge structures. Koponen and Pehkonen (2008) found that the knowledge structures of experts in physics were consistent and hierarchical, while knowledge structures of novices were fractured. The knowledge structures of experts were more closely integrated than those of novices (Ifenthaler et al., 2011). Therefore, when defining learners' ability, the integrity and good connection of knowledge structures are important considerations (Wadouh et al., 2014). As a more reliable method to determine learning results, knowledge structures can not only simply reflect the recognition and recall of concepts but also provide higher-order learning evidence to reflect learners' understanding of complex fields (Spector, 2006). Since mining and analyzing learners' knowledge structures is considered to be an important understanding of learners' learning quality (Lopez et al., 2014), it is of great significance to study the knowledge structures of online learners to help online learning participants (including teachers and students) deeply understand the level of online learning.

2.2 Using concept maps to investigate knowledge structures

Knowledge structures can be elicited and analyzed by card classification, word association, text analysis, concept maps, complex semantic network tools, and other methods (Clariana et al., 2014; Eseryel et al., 2013; Tang & Clariana, 2017). Although there have been different attempts made to elicit knowledge structures, the process of self-constructing concept maps is considered to be a more straightforward approach (Jonassen & Marra, 1994). This approach can not only visualize learners' knowledge structures but also effectively promote learners' ability improvement and learning performance, including reading ability (Clariana et al., 2014) and writing ability (Liu, 2011). Concept maps were developed by Novak and Gowin (1984) based on Ausubel's relevance theory of meaningful learning, and the key elements of concept maps include nodes, relationships, and propositions that are semantic understandings formed by nodes and relationships. The concept map can visually represent the knowledge structures of learners through its key elements (Jonassen et al., 1993). Constructivism theory mentions that the process in which learners try to solve their conflicts in how to represent and mark knowledge is called internal negotiation, while in the self-constructing concept map, learners achieve internal negotiation by creating nodes, describing node relations, and constantly reviewing node relations (Jonassen & Marra, 1994). The process of self-constructing concept maps allows teachers the opportunity to observe the extensive and integrated conceptual knowledge of students, which is difficult to observe through traditional assessment forms (such as multiple-choice questions) (Walker & King, 2003). Therefore, the process of self-constructing concept maps is a suitable method for eliciting learners' knowledge structures in online learning, realizing the representation of learners' knowledge structure, and including the learning results of learners' deep learning characteristics (Jonassen et al., 1993).

In the research of using concept maps to analyze knowledge structure, researchers have focused on how to quantitatively analyze learners' concept maps. The traditional concept map evaluation method is mainly evaluated by raters based on the level and quality of propositions in the concept map (Watson et al., 2016). With the development of the automatic evaluation of concept maps, researchers can calculate the structural indicators of maps based on graph theory and network theory, including the number of nodes, the number of layers (Novak & Gowin, 1984), overall width (Hao et al., 2010), maximum layer width (Jonassen et al., 1993), etc. In addition, Clariana et al. (2013) used the centrality index of concept maps to measure the knowledge structure of learners in their study of computer-supported collaborative problem solving. Kapuza et al. (2020) proposed three measures based on the basic structure index of the concept map to analyze the concept map generated during the statistical data analysis course to understand the knowledge structure of learners.

However, Kinchin et al. (2000) warned that quantitative evaluation of concept maps should be carefully used because quantitative analysis ignores the structure of wrong concepts, and these mistakes convey the learning development trend of learners in the future (Kinchin, 2014). Therefore, the authors proposed a qualitative analysis method of map morphology to describe the concept map as three modes, namely, spoke, chain, and network (Kinchin et al., 2000), thereby further associating the morphological evolution of the concept map with the learning mode (Hay et al., 2008). Some researchers have further conducted a qualitative analysis of knowledge structure based on Kinchin's analysis method. Koponen and Pehkonen (2008) reconstructed concept maps constructed by learners into two topologies, namely, network and tree, for analysis and showed that the structure and topological features of concept maps have the ability to distinguish between novices and experts. Hung and Lin

(2015) found three types of concept maps, namely, isolated mapping, departmental mapping, and integrated mapping; they also found that more integrated mapping occurs among learners who used intervention strategies in problem-solving teaching by the empirical study. The morphological analysis of the map provides a unique perspective to compensate for the lost learning information of the quantitative evaluation map and provides a new idea for the analysis of learners' knowledge structures in online learning. However, the qualitative analysis of map forms requires great energy from teachers and it is difficult to promote in online learning. Based on previous quantitative research on concept maps, our study combined the quantitative analysis method with the morphological analysis method of map structure to explore a data mining method to analyze online learners' knowledge structures. The study identified online learners' knowledge structure patterns and further investigated the relationship between knowledge structure patterns and online learning achievement.

The following three questions are mainly investigated:

(1) Do learners' knowledge structures have identifiable patterns in online learning?
(2) How can different types of online learners be identified from the distribution of their knowledge structure patterns?
(3) What are the differences in the online learning achievement of the different types of online learners?

3 Methods

This study aims to identify meaningful patterns of knowledge structure in online learning by analyzing the features of learners' knowledge structures, and to further discover the learner types with different patterns of knowledge structures. For this purpose, we adopted clustering to automatically identify similar learners. Clustering is an unsupervised data mining algorithm that divides similar samples into the same cluster. In other words, the algorithm can uncover unique learner types with similar knowledge structure patterns. Besides, we used a statistical analysis method to compare the learning achievements of different types of online learners. By doing so, it was expected to provide insights about the relationship between knowledge structure and learning achievements in online learning environments.

3.1 Learning context

In this study, a flipped classroom was mainly adopted as the learning context. The flipped classroom is a common blended learning model, in which learners should complete online learning tasks before entering the classroom (Devi et al., 2021). Specifically, due to the impact of the COVID-19 epidemic, the learning model was implemented in a computer programming course at a university in China. More specifically, the blended learning model was supported by an online platform and an online conference system.

The course contains 14 lessons. For each lesson, the teacher provided learning materials, which covered basic knowledge through the online platform. The students were asked to complete these materials before the face-to-face sessions. Subsequently, the teacher explained the focus of the course in the face-to-face sessions. Due to the impact of COVID-19, the face-to-face sessions were organized through the Tencent Conference system in this course. Each lecture was given in two consecutive sessions, totaling 90 min. The students participated in teaching the course once a week.

3.2 Participants

Thirty-six students participated in this course, including 15 males and 21 females. They are freshmen majoring in data science and big data technology. The average age of these students was 18 years old. They had not taken other programming-related courses before participating in this course and thus were considered novice learners in the field. Prior to attending this course, all the participants had no experience in concept mapping, so they were considered unbiased in the learning instruments.

3.3 Procedure

All of the participants completed one lesson of learning per week from week 1 to week 14. The specific learning process weekly was introduced in session 3.1. The students also received training on the use of the online learning system and concept mapping in week 1. The training about concept mapping used the method proposed by Roessger et al. (2018) and mainly focused on how to extract relationships between concepts. In addition, all students participated in the exam through an additional online examination system during the middle and the end of the course. The midterm exam was held in week 9, while the final exam was delayed until three months after the course due to COVID-19.

3.4 Concept map task

The knowledge structure is hidden in the human mind and cannot be observed directly, but it can be represented through the concept map (Jonassen & Marra, 1994). Studies have suggested that low-orientation concept map tasks can better express the knowledge structures of learners (Cañas et al., 2012). A low-orientation concept map task means that someone should construct a concept map without any concept, structure, or relationship provided in advance (Ruiz-Primo, 2000). The study used a low-orientation concept map task to extract learners' knowledge structures. All the participants were asked to construct a concept map to summarize the knowledge in the unit after completing online learning.

3.5 Online learning platform

In this study, an online learning platform called "Xiaoya" was developed by our research team to support online learning. The system provides the basic functions

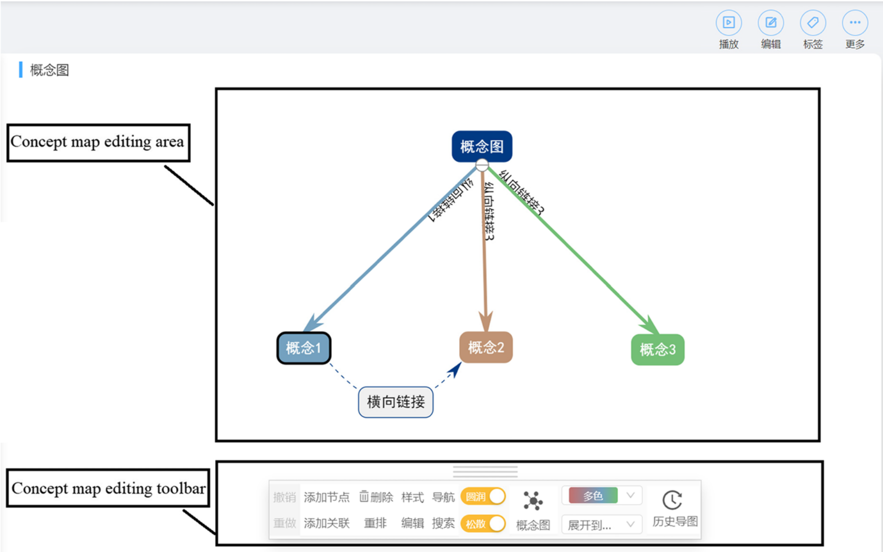


Fig. 1 The interface of constructing concept mapsin Xiaoya

of a learning management system. More specifically, the main functions of the platform include learning materials, quizzes, discussions, and notes. Through the notes function, students are able to generate document notes and concept maps. The concept maps constructed by students can include concepts, relationships, and cross-links, as shown in Fig. 1. All the concept maps constructed by the learning platform are recorded in a standard format.

In our study, the teacher constructed a virtual curriculum space and provided learners with various forms of learning materials, including videos and slides.

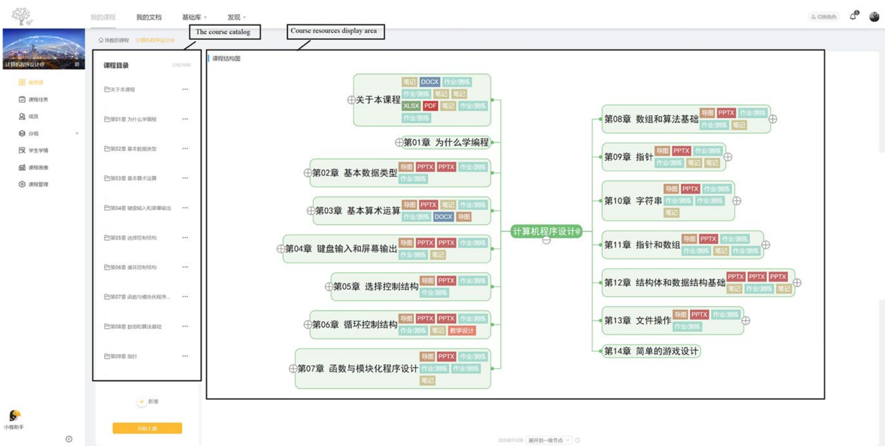
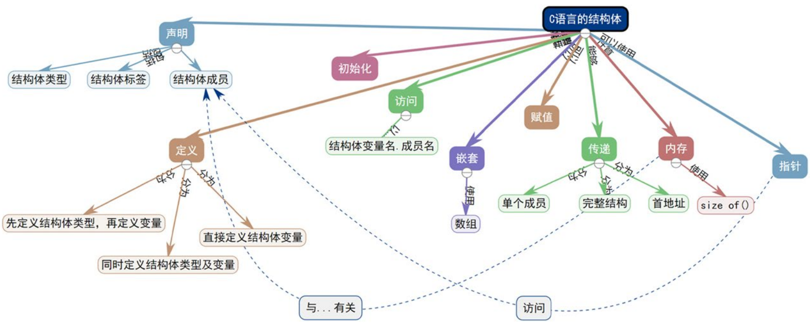


Fig. 2 Course space in the Xiaoya online learning platform

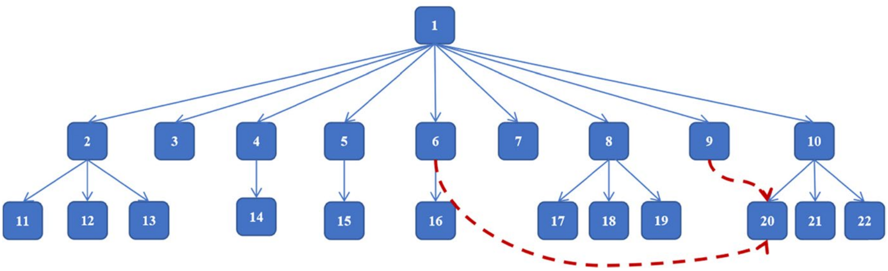
Besides, the teacher assigned concept map tasks as the questions in the quiz function. The students who took the course had to complete the materials, then construct concept maps through notes and upload them in the quiz function (Fig. 2).

3.6 Data collection

The researchers collected the concept map artifacts (N = 359) that were constructed by the 36 students from week 2 to week 12. Because the first lesson was only an introduction and lesson 13 and 14 were optional, the concept map artifacts were excluded for these three weeks. These concept map artifacts are recorded in JavaScript object symbol format through the Xiaoya online learning platform. We used Python to parse the concept map artifact in the JSON format derived from the Xiaoya platform, and each concept map artifact was reconstructed into a directed network topology (Fig. 3). The directed network topology retained all structural features in the original concept map artifact, which was used as the learner knowledge structure representation. In addition, we also collected midterm and final exam scores as learning achievements for online learners.



(i)



(ii)

Fig. 3 Constructing the directed topological network from the concept map as online learners' knowledge structures. *Note.* (**i**) Original learners' concept map. (**ii**) Reconstructed map (the content information of the concept map is removed, and only the structural features are retained)

3.7 Data analysis

Based on the three research questions posed in our study, the data analysis process (Fig. 4) was performed as follows.

To answer research question 1, the knowledge structure patterns of online learners were identified by a K-means clustering algorithm. The algorithm is an iterative solution clustering analysis algorithm that is widely used in educational data mining research (Dutt et al., 2015). Graph theory has been used in many studies on concept map analysis (Ifenthaler et al., 2011). This research summarizes the structural indicators proposed in previous concept map research and identifies 19 structural indicators (Table 1) as feature vectors for cluster analysis based on graph theory. Before calculating the indicators, the teacher checked the artifacts, and none were off-task. Besides, the topic of the concept map task may affect the scale of the map, while the maps collected in this study were

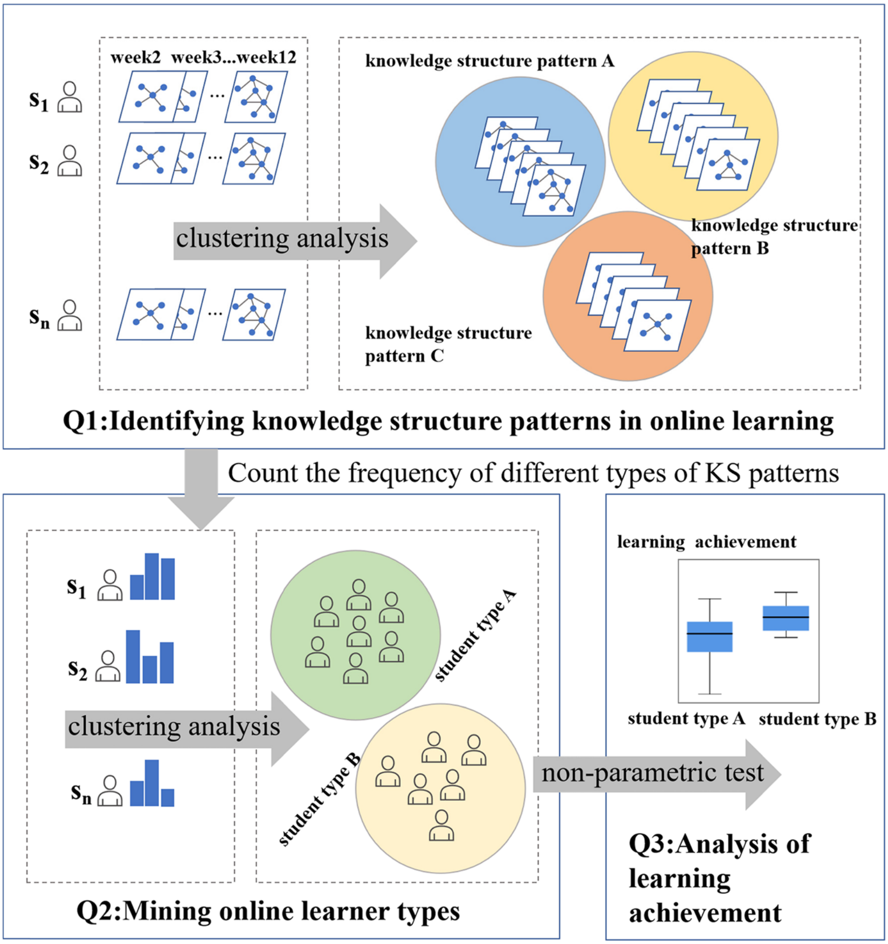


Fig. 4 Data analysis process. *Note*: KS patterns represent knowledge structure patterns

Table 1 Structural indicators of concept maps

Indicators	Depiction of indicators	References
V1	Number of nodes	Novak and Gowin (1984)
V2	Number of leaf nodes	
V3	Number of branches	
V4	Maximum layer width	Jonassen et al. (1993)
V5	Number of relationships	Novak and Gowin (1984)
V6	Number of cross-links	
V7	Total width	Hao et al. (2010)
V8	Average width	Morine-Dershimer (1993)
V9	Sum of information entropy	Hao et al. (2010)
V10	Average information entropy of non-leaf nodes	
V11	Number of layers	Novak and Gowin (1984)
V12	Maximum leaf node depth	Hao et al. (2010)
V13	Sum of leaf node depth	
V14	Total number of paths from the root node to the leaf nodes	
V15	Average leaf depth	
V16	Total path depth	
V17	Total number of paths	
V18	Average path depth	
V19	Number of descriptors	Roessger et al. (2018)

V19 counts the number of descriptors used to describe cross-links and relationships

from 11 different topics. Therefore, we standardized the structural metrics of the maps with standard concept maps provided by the teacher to eliminate this effect. In addition, the sum of square error (SSE) was used to determine the optimal K-clustering value (Nainggolan et al., 2019), and the Calinski-Harabaz (CH) index was used to determine the best clustering results (Rajabi et al., 2019). Both the indicators' calculation of concept maps and cluster analysis were executed automatically in Python programming language. In addition, the frequency and percentage of each knowledge structure pattern were calculated for the study semester, and the distribution of different kinds of knowledge structure patterns over time throughout the semester (12 weeks) was visualized.

To answer research question 2, we counted the frequency of different types of knowledge structure patterns identified by research question 1 during the semester (12 weeks) as clustering features. The K-means clustering analysis method was used to discover different types of online learners. Furthermore, the distribution of patterns of different types of online learners and their changes from week 2 to week 12 were also visualized.

To answer research question 3, we compared the learning achievements among the different types of online learners discovered in question 2. Due to the small number of samples in this study, a non-parametric test with the Mann-Whitney U test was taken.

4 Results

4.1 RQ1: Do learners' knowledge structures have recognizable patterns in online learning?

The concept map artifacts constructed by online learners from week 2 to week 12 (N = 359) were analyzed by the K-means algorithm. The optimal clustering results showed that there were three clusters in the directed network topology of the concept map, namely, Cluster_0, Cluster_1, and Cluster_2. The descriptive statistics of each category are shown in Table 2. Among them, Cluster_0 is the largest with 225 copies, accounting for 62.7% of the total number of submitted maps; Cluster_1 is the second largest with 106 copies, accounting for 29.5% of the total number of submitted maps; and Cluster_2 is the smallest, with 28 copies, accounting for 7.8% of the total number of submitted maps. In terms of structural indicators from V1 to V18, the three categories of Cluster_0, Cluster_1, and Cluster_2 increase in sequence. For the average value of V10, Cluster_0, Cluster_1, and Cluster_2 decrease in order because V10 and V9 have a reciprocal relationship. In terms of indicator V19, there is little difference between Cluster_1 and Cluster_2, which are both larger than Cluster_0. Overall, Cluster_0, Cluster_1, and Cluster_2 increase successively in structural complexity.

The K-means algorithm uses the centroid of each class to describe the cluster, but the centroid is not necessarily the location of the sample. Therefore, this research calculated the sample closest to the centroid as the representative of the cluster, reconstructed the corresponding concept map according to the reconstruction rules, and then visualized it for morphological analysis. The topology of the directed network reconstructed by Cluster_0, Cluster_1, and Cluster_2 is shown in Fig. 5(i)-(iii).

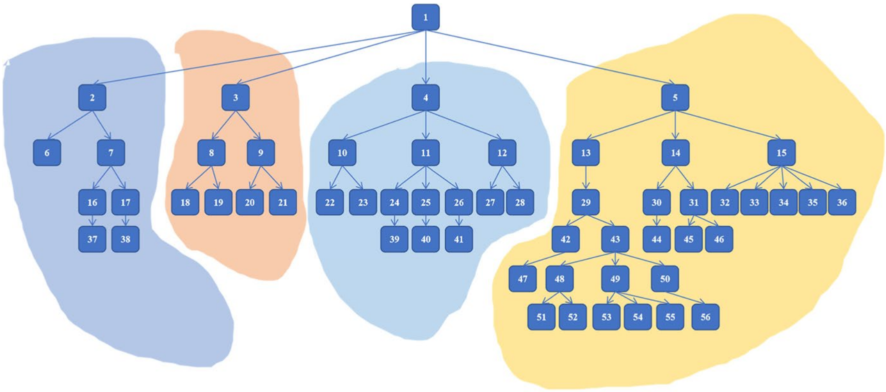
Morphological analysis of the three directed topology network visualization legends found the following. First, from the scale of the directed network topology, Cluster_2, Cluster_1, and Cluster_0 decreased in sequence. Second, from the perspective of the type of directed network topology, there were cross-links (red link) in both Cluster_1 and Cluster_2. Because of the existence of cross-links, both groups belong to the network structure. However, there was no cross-link in Cluster_0, which was essentially a tree-like structure showing a radial shape. Third, from the perspective of the aggregation of branches, nodes in the network structure gather into communities through connections, and cross-links connect small communities into larger communities. Since there were more cross-links in Cluster_2 than in Cluster_1, small communities formed larger communities due to cross-links, and the community sizes of nodes in Cluster_0, Cluster_1, and Cluster_2 increased sequentially.

Based on the descriptive statistical analysis and morphological analysis of the three categories, Cluster_0, Cluster_1, and Cluster_2 were three knowledge structure patterns that gradually increased in scale. Cluster_0 represented the tree-like structure, while Cluster_1 and Cluster_2 represented the networked structure. Therefore, this research named the three categories of knowledge structure patterns, i.e., Cluster_0, Cluster_1, and Cluster_2, the spoke pattern, the small-network pattern, and the large-network pattern, respectively.

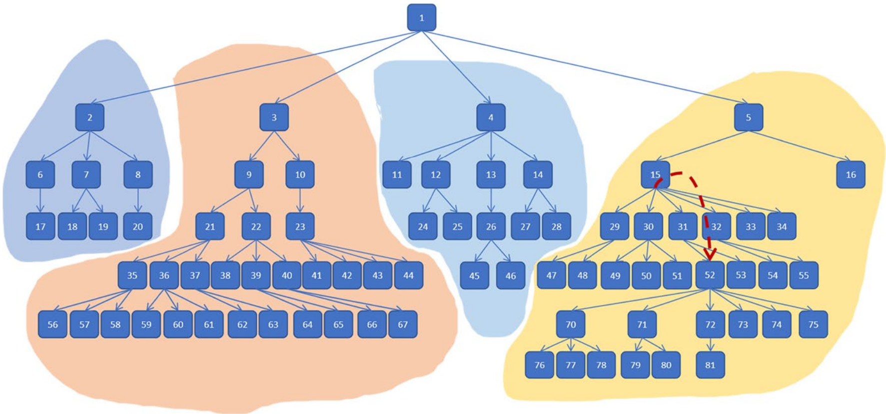
Table 2 The statistical description of structure indicators of three categories of concept maps

Cluster	N	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19
Cluster_0	225	0.90	0.97	0.82	1.07	0.90	0.19	0.84	0.93	1.02	1.41	0.80	0.70	0.57	0.75	0.72	0.54	0.72	0.70	0.07
Cluster_1	106	2.20	2.01	2.16	1.89	2.26	0.63	2.17	0.99	1.81	0.80	1.37	1.51	2.45	1.92	1.29	2.81	2.18	1.26	0.38
Cluster_2	28	4.98	4.45	4.93	3.58	5.19	1.41	4.99	1.01	3.16	0.58	1.87	2.49	8.08	4.55	1.75	10.26	5.55	1.76	0.35

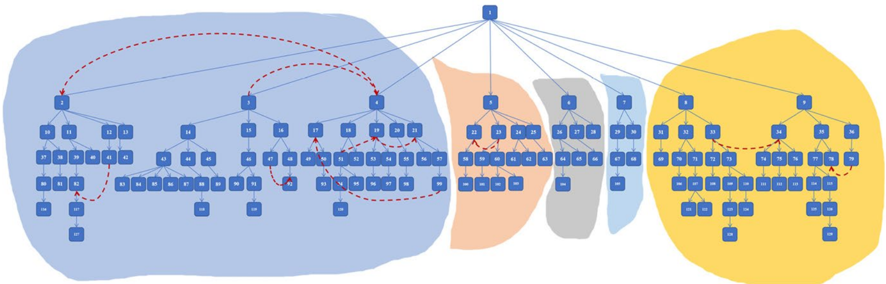
N indicates the number of concept maps in this cluster



(i)



(ii)



(iii)

Fig. 5 Directed network topology of representative concept maps of three clusters. *Note.* (**i**) Cluster_0: spoke pattern; (**ii**) Cluster_1: small-network pattern; (**iii**) Cluster_2: large-network pattern

4.2 RQ2: How can different types of online learners be identified from the distribution of their knowledge structure patterns?

The knowledge structures of learners reflect the organizational form of domain knowledge in their minds, which is not immutable (Ifenthaler et al., 2011). Thus, it is of great significance to analyze and study the changes in learners' knowledge structure patterns over a long period of time to reveal their learning (Jonassen et al., 1993). Research question 2 further used the frequency of learners' occurrence of different structural patterns over the course of the study semester as a description of learners' characteristics and again discovered different types of online learners through cluster analysis. Since not all the learners who participated in the experiment completed the 11 concept map tasks, we added a nonpattern to describe a situation in which a concept map was not submitted. Finally, we constructed a feature vector containing four dimensions (i.e., statistical frequency in non, spoke, small-network, and large-network) and used the K-means algorithm to conduct clustering analysis on 36 learners.

The optimal clustering results showed that there were two clusters, namely, Cluster_a and Cluster_b. The descriptive statistics of each cluster are shown in Table 3. In Cluster_a, the mean frequency of learners' occurrence of the four dimensions was non ($n=1.75$), spoke ($n=7.69$), small-network ($n=1.50$), and large-network ($n=0.06$). In Cluster_b, the mean frequency of learners' occurrence of the four dimensions was non ($n=0.45$), spoke ($n=4.90$), small-network ($n=4.30$), and large-network ($n=1.35$). We visualized the distribution of online learners in Cluster_a and Cluster_b in 4 dimensions (Fig. 6). There were more large-network patterns and small-network patterns in Cluster_b than in Cluster_a, while there were more spoke patterns and non patterns in Cluster_a than in Cluster_b. We defined Cluster_a as simple knowledge structure online learners and Cluster_b as complex knowledge structure online learners, thereby combining the characteristics of the two clusters in the statistical analysis results and visual analysis results.

To further analyze the characteristics of the knowledge structures of the two types of online learners changing over time, we visualized the distribution changes of the knowledge structures of the two types of learners from Week 2 to Week 12 (Fig. 7). For simple knowledge structure online learners, the small-network pattern appeared from Week 4, and the large-network pattern appeared from Week 10. Moreover, the total number of small-network and large-network patterns was relatively high in Weeks 5, 6, 10, and 11. In addition, the nonpattern appeared in simple knowledge structure online learners every week except Week 6, and nonpattern counts increased in Week 5, Week 8, Week 10, and Week 12. For complex knowledge structure online learners, the small-network pattern appeared from Week 2, and the large network pattern appeared from Week 5. Moreover, the total number of small

Table 3 The statistical description of four patterns distributions of two learners' clusters

Cluster	N	Non	Spoke	Small-network	Large-network
Cluster_a	16	1.75	7.69	1.50	0.06
Cluster_b	20	0.45	4.90	4.30	1.35

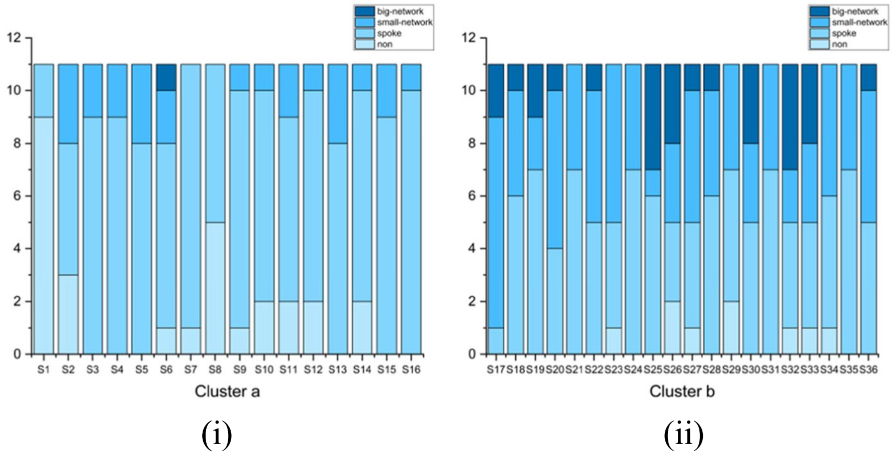


Fig. 6 Patterns' distribution histogram of two learner clusters

network and large network patterns showed two continuous and regular evolves from Week 2 to Week 6 and from Week 8 to Week 11. Similar to simple knowledge structure online learners, complex knowledge structure online learners showed the peak of small network and large network patterns at Weeks 5, 6, 10, and 11.

4.3 RQ3: What are the differences in the online learning achievement of the different types of online learners?

The Mann-Whitney U test results showed in Table 4. In terms of midterm scores, the complex type learners scored significantly higher grades than the simple type learners ($P=0.03$). Meanwhile, in terms of final scores, the complex group still had higher scores than the simple group, but the difference was not significant ($P=0.46$).

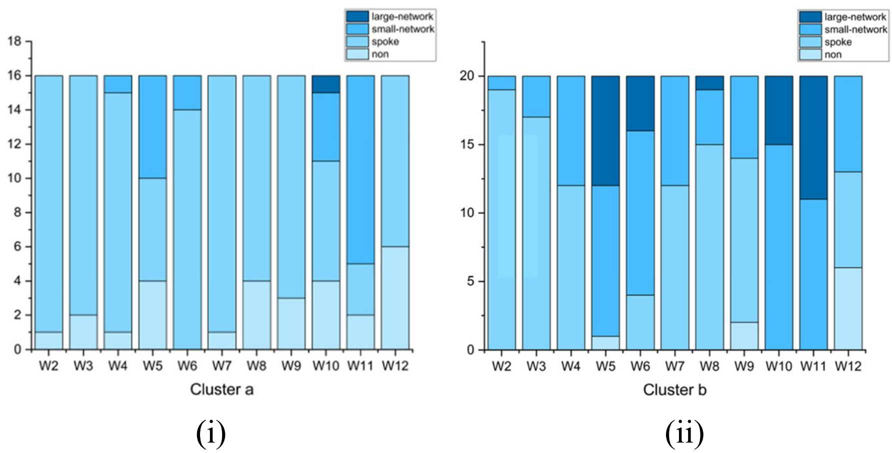


Fig. 7 Patterns' distribution histogram of two learner clusters by weeks

Table 4 Mann-Whitney U test results of online learning achievement between two types of online learners

	Simple type mean scores ($n=16$)	Complex type mean scores ($n=20$)	Mann-Whitney U	p
Midterm score	71.38	79.90	92.50	0.03*
Final score	61.34	65.59	136.50	0.46

$^{*}p<0.05$

5 Findings and discussion

Based on graph theory, this study explored a new method of analyzing the knowledge structure of online learners through data mining by using structural indicators of maps as characteristics. This research identified knowledge structure patterns of online learners through a K-means clustering algorithm and revealed the morphological characteristics of online learners' knowledge structure. One of the main findings of the research was the identification of three knowledge structure patterns of online learners, namely, spoke patterns, small network patterns, and large network patterns. Statistical analysis and morphological analysis were used to further explore the characteristics of the three structural patterns. The results showed that, first, the scale of the three patterns increased in turn. Second, the spoke pattern did not have cross-links and represented the tree structure, while the small network pattern and the large network pattern both contained cross-links and represented the network structure. Third, the difference in cross-links affected the degree of node aggregation in the three patterns. As the number of cross-links among spoke patterns, small network patterns, and large network patterns increased, the scale of node aggregation in knowledge structure patterns also increased. The structural patterns automatically identified by the data mining method in this research were basically consistent with the concept map structural patterns found by the previous qualitative analysis. Kinchin et al. (2000) first proposed three forms of spoke, chain, and net using the qualitative analysis of concept maps. In the study of Kinchin et al. (2000), a spoke was a tree structure with a single level in which many concepts were directly connected to the central concept, and a chain was a linear sequence of understanding in which each concept was only connected to the concept above or below, and the net was a highly integrated and hierarchical network. Hung and Lin (2015) further divided map morphology into isolated mapping, departmental mapping, and integrated mapping when applying map morphology to the study of knowledge structure in problem-solving learning. In the isolated mapping, several single concepts were linked to the core concepts. In the departmental mapping, several single concepts were combined into conceptual units, but no network was formed. In the integrated mapping, the whole map formed a network, including not only units for concept formation but also horizontal connections between units. In the qualitative analysis of map morphology, the connectivity of map structure is considered to be an important indicator (Kinchin et al., 2005), which has been verified in the identification of knowledge structure patterns using data mining methods.

This study had an additional finding based on the identification of three online learners' knowledge structure patterns. In the preclass learning in flipped classrooms without additional teaching design, most learners showed the knowledge structure of spoke patterns, while only a few learners showed large-network patterns. In fact, this distribution pattern was consistent with the pattern found in the previous qualitative evaluations of map morphology. The measurement of knowledge structure usually includes the number of nodes in the structure and the relationship between nodes (Clariana et al., 2013). The relationship between nodes promotes the formation of coherent networks with isolated knowledge (Koponen & Pehkonen, 2010). Novices present spoke patterns of knowledge structures so that new information can be added to existing knowledge structures in a timely manner (Kinchin et al., 2000). Domain experts tend to be more aggregated than novices in terms of knowledge structures (Ifenthaler et al., 2011). The knowledge structure of network attributes often reflects learners' good understanding of knowledge (Joseph et al., 2017). The knowledge structure collected in this research came from the preclass learning of learners in a flipped classroom. Before participating in the experiment, the learners had not received any knowledge learning experience related to this course and were thus considered novices in this knowledge field. The distribution characteristics of knowledge structure patterns found in the research were in line with the novice status of learners and reflected the real situation that learners were not well prepared in the preclass learning of flipped classrooms without especially teaching intervention. In research on teaching strategies based on problem solving, Hung et al. found that teaching intervention strategies promoted the evolution of learners' knowledge structures from isolated mapping to integrated mapping, and the evolution of knowledge structure reflected the differentiation and integration of learners' concepts (Hung & Lin, 2015). Therefore, teachers should attach importance to the preclass teaching design of flipped classrooms because the potential benefits of the flipped classroom may be weakened due to learners' insufficient preclass preparation (Chuang et al., 2018).

Another main finding of the study was to identify two types of online learners with different distributions among the three identified online knowledge structure patterns. We found some common characteristics and differences worthy of further discussion by comparing and analyzing the performance of different types of knowledge structure pattern distributions between the two types of learners. Compared with the simple knowledge structure online learners, the complex knowledge structure online learners appeared with large network patterns and small network patterns earlier, and there were two-periodic evolution increases in knowledge structure complexity from Week 2 to Week 6 and from Week 8 to Week 12. The patterns transformed from spoke to net include the development of the learner's learning, and the net patterns represent a high level of understanding of the topic (Kinchin et al., 2000). The development of knowledge structure reflects changes in individual knowledge and understanding and can be used as evidence of meaningful learning (Jonassen et al., 1993), which can distinguish between deep learning and shallow learning in the learning process of learners (Hay, 2007). For the complex knowledge structure, online learners with gradual improvement regarding the connectivity of knowledge structures with the progress of learning evidence the development of knowledge structure patterns from spoke patterns to large network patterns. The increased connectivity in knowledge structure during learning reflects the occurrence of meaningful learning by learners (Joseph et al., 2017). In addition, the two types of online learners simultaneously showed relatively more complex knowledge structures in Weeks 5, 6, 10, and 11. After reviewing the teaching content, it was found that the above four teaching units belong to the teaching focus of this course and contain more knowledge points, which may lead to the main reason for the improvement of the complexity of the knowledge structure patterns of the two types of online learners.

A comparison of the learning achievement of the two types of learners found that the complex type learners had outperformed the counterpart in both the midterm and final exams. However, this difference was not significant in the final exam scores. One possible explanation was that the final exam did not fully reflect the learner's knowledge at the end of the course due to the delay in the exam. The performance of some students might be interfered by other factors. Even so, in general, the results of the learning achievement analysis showed that complex type learners achieved better grades than simple type learners. Through meaningful learning, the learner completes the integration of new knowledge with prior knowledge (Novak, 2010), forming a richer and more aggregated knowledge structure (Ifenthaler et al., 2011). The differences in learning achievement between the two types of learners were consistent with the differences in the distribution of knowledge structure patterns. Consistency findings also support the validity of automatic analysis of knowledge structures through this data mining approach.

6 Conclusion and limitations

Concept maps constructed by learners represent their knowledge structures and reflect learners' deep understanding of knowledge (Jonassen & Marra, 1994). Analyzing these knowledge structures to better understand learners' learning can promote teachers' teaching transformation from content to understanding (Kinchin et al., 2000). However, online learners' knowledge structures have not been widely analyzed; there is especially a lack of analysis using the data mining method. This study explored a data mining method to analyze the knowledge structures of learners and realized the automatic analysis of the morphological characteristics of knowledge structures through a clustering algorithm. First, three online learners' knowledge structure patterns were identified, and the distribution characteristics of these three knowledge structure patterns were found in the pre-class of the flipped classroom without any special teaching interventions. Second, based on the identified knowledge structure patterns of online learners, two types of learners, namely, complex knowledge structures and simple knowledge structures, were further discovered based on the long-term changes in the knowledge structure patterns of learners. Finally, through statistical analysis, it was verified that complex knowledge structure online learners had better learning achievement than simple knowledge structure online learners.

Different learning strategies may lead to different knowledge structures in a regular format (Kim et al., 2019), and observing the morphological differences in knowledge structures provides us with a unique perspective with which to observe the changes in learners' online learning to measure whether learning strategies truly produce effects. Generally, our study proposed a new method for analyzing the knowledge structure patterns of online learners by data mining and initially explores the correlation between knowledge structure patterns and learning achievements. This innovative analysis approach provided a convenient way for educators in online teaching to understand the knowledge structure of students. Furthermore, the findings in this study on the distribution of knowledge structure patterns in flipped classroom learning context provided empirical evidence for the need of pre-class instructional design.

Although this research preliminarily explored a new means of using data analysis to analyze the knowledge structure of online learners, it still had some limitations. First, this research did not distinguish semantic correctness when analyzing the structural features of concept maps. This is mainly because our study aims to study the online knowledge structure from the perspective of structural features. Previous studies emphasized that invalid concepts in concept maps had also revealed learners' thinking processes (Kinchin et al., 2000), and concept map assessment studies focused on accuracy had been criticized for ignoring important details (Kinchin, 2014). Moreover, before cluster analysis, the teacher checked all the concept maps to ensure none were off-task. For these reasons, our study can still provide valid and reliable findings from the perspective of morphological features of knowledge structures. In future studies, we can further evaluate the accuracy of concept maps through natural language processing, and compare the impact of accuracy on understanding the knowledge structure. Second, the online learning context studied in this paper is preclass autonomous learning or mixed learning based on flipped classrooms. The research finds that the characteristics of the knowledge structure mode of learners' online learning are consistent with this teaching context, and further research is needed to generalize this research conclusion to other online learning contexts.

Data availability Research data are not shared.

Declarations

The study and the data used in the study have been approved by the ethics Committee of our institution.

Informed consent All authors have given their informed consent to this paper.

Conflict of interest We have no conflicts of interests to disclose.

References

Ausubel, D. P. (1963). *The psychology of meaningful verbal learning*. Oxford.

Brown, A., Lawrence, J., Basson, M., & Redmond, P. (2022). A conceptual framework to enhance student online learning and engagement in higher education. *Higher Education Research & Development, 41*(2), 284–299. https://doi.org/10.1080/07294360.2020.1860912

Cañas, A. J., Novak, J. D., & Reiska, P. (2012). Freedom vs. restriction of content and structure during concept mapping -possibilities and limitations for construction and assessment. *Proc. of the Fifth Int. Conference on Concept Mapping*, *2*(1), 247–257. Retrieved February 14, 2023, from http://cmc.ihmc.us/cmc2012Papers/cmc2012-p192.pdf

Carolan, C., Davies, C. L., Crookes, P., McGhee, S., & Roxburgh, M. (2020). COVID 19: disruptive impacts and transformative opportunities in undergraduate nurse education. *Nurse Education in Practice, 46*, 102807. https://doi.org/10.1016/j.nepr.2020.102807

Chen, Z., Jiao, J., & Hu, K. (2021). Formative assessment as an online instruction intervention: student engagement, outcomes, and perceptions. *International Journal of Distance Education Technologies, 19*(1), 50–65. https://doi.org/10.4018/IJDET.20210101.oa1

Cheng, A. C., Jordan, M. E., & Schallert, D. L. (2013). Reconsidering assessment in online/hybrid courses: knowing versus learning. *Computers & Education, 68*, 51–59. https://doi.org/10.1016/j.compedu.2013.04.022

Chi, M. T. H., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science, 5*(2), 121–152. https://doi.org/10.1207/s15516709cog0502_2

Chuang, H. H., Weng, C. Y., & Chen, C. H. (2018). Which students benefit most from a flipped classroom approach to language learning?: flipped classroom does not fit all students. *British Journal of Educational Technology, 49*(1), 56–68. https://doi.org/10.1111/bjet.12530

Clariana, R. B., Engelmann, T., & Yu, W. (2013). Using centrality of concept maps as a measure of problem space states in computer-supported collaborative problem solving. *Educational Technology Research and Development, 61*(3), 423–442. https://doi.org/10.1007/s11423-013-9293-6

Clariana, R. B., Wolfe, M. B., & Kim, K. (2014). The influence of narrative and expository lesson text structures on knowledge structures: alternate measures of knowledge structure. *Educational Technology Research and Development, 62*(5), 601–616. https://doi.org/10.1007/s11423-014-9348-3

Deng, R., Benckendorff, P., & Gannaway, D. (2019). Progress and new directions for teaching and learning in MOOCs. *Computers & Education, 129*, 48–60. https://doi.org/10.1016/j.compedu.2018.10.019

Devi, B., Sharma, C., & Lepcha, N. (2021). Blended learning—A global solution in the age of COVID-19. *Journal of Pharmaceutical Research International*, 125–136. https://doi.org/10.9734/jpri/2021/v33i41B32351

Dutt, A., Aghabozrgi, S., Ismail, M. A. B., & Mahroeian, H. (2015). Clustering algorithms applied in educational data mining. *International Journal of Information and Electronics Engineering*. https://doi.org/10.7763/IJIEE.2015.V5.513

Eseryel, D., Ifenthaler, D., & Ge, X. (2013). Validation study of a method for assessing complex ill-structured problem solving by using causal representations. *Educational Technology Research and Development, 61*(3), 443–463. https://doi.org/10.1007/s11423-013-9297-2

Farrokhnia, M., Pijeira-Díaz, H. J., Noroozi, O., & Hatami, J. (2019). Computer-supported collaborative concept mapping: the effects of different instructional designs on conceptual understanding and knowledge co-construction. *Computers & Education, 142*, 103640. https://doi.org/10.1016/j.compedu.2019.103640

García-Morales, V. J., Garrido-Moreno, A., & Martín-Rojas, R. (2021). The transformation of higher education after the COVID disruption: emerging challenges in an online learning scenario. *Frontiers in Psychology, 12*, 616059. https://doi.org/10.3389/fpsyg.2021.616059

Gogus, A. (2013). Evaluating mental models in mathematics: a comparison of methods. *Educational Technology Research and Development, 61*(2), 171–195. https://doi.org/10.1007/s11423-012-9281-2

Hao, J. X., Kwok, R. C. W., Lau, R. Y. K., & Yu, A. Y. (2010). Predicting problem-solving performance with concept maps: an information-theoretic approach. *Decision Support Systems, 48*(4), 613–621. https://doi.org/10.1016/j.dss.2009.12.001

Hay, D. B. (2007). Using concept maps to measure deep, surface and non-learning outcomes. *Studies in Higher Education, 32*(1), 39–57. https://doi.org/10.1080/03075070601099432

Hay, D., Kinchin, I., & Lygo-Baker, S. (2008). Making learning visible: the role of concept mapping in higher education. *Studies in Higher Education, 33*(3), 295–311. https://doi.org/10.1080/03075070802049251

Hung, C. H., & Lin, C. Y. (2015). Using concept mapping to evaluate knowledge structure in problem-based learning. *BMC Medical Education, 15*(1), 212. https://doi.org/10.1186/s12909-015-0496-x

Ifenthaler, D., Masduki, I., & Seel, N. M. (2011). The mystery of cognitive structure and how we can detect it: tracking the development of cognitive structures over time. *Instructional Science, 39*(1), 41–61. https://doi.org/10.1007/s11251-009-9097-6

Jonassen, D. H., Yacci, M., & Beissner, K. (1993). *Sturctural knowledge: techniques for representing, conveying, and acquiring structural knowledge*. Routledge. https://doi.org/10.4324/9780203052563

Jonassen, D. H., & Marra, R. M. (1994). Concept mapping and other formalisms as mindtools for representing knowledge. *Research in Learning Technology, 2*(1), 50–56. https://doi.org/10.3402/rlt.v2i1.9573

Joseph, C., Conradsson, D., Nilsson Wikmar, L., & Rowe, M. (2017). Structured feedback on students' concept maps: the proverbial path to learning? *BMC Medical Education, 17*(1), 90. https://doi.org/10.1186/s12909-017-0930-3

Kapuza, A., Koponen, I. T., & Tyumeneva, Y. (2020). The network approach to assess the structure of knowledge: Storage, distribution and retrieval as three measures in analysing concept maps. *British Journal of Educational Technology, 51*(6), 2574–2591. https://doi.org/10.1111/bjet.12938

Khajeloo, M., & Siegel, M. A. (2022). Concept map as a tool to assess and enhance students' system thinking skills. *Instructional Science*. https://doi.org/10.1007/s11251-022-09586-5

Kim, K., Clarianay, R. B., & Kim, Y. (2019). Automatic representation of knowledge structure: enhancing learning through knowledge structure reflection in an online course. *Educational Technology Research and Development, 67*(1), 105–122. https://doi.org/10.1007/s11423-018-9626-6

Kinchin, I. M., Hay, D. B., & Adams, A. (2000). How a qualitative approach to concept map analysis can be used to aid learning by illustrating patterns of conceptual development. *Educational Research, 42*(1), 43–57. https://doi.org/10.1080/001318800363908

Kinchin, I. M., De-Leij, F. A. A. M., & Hay, D. B. (2005). The evolution of a collaborative concept mapping activity for undergraduate microbiology students. *Journal of Further and Higher Education, 29*(1), 1–14. https://doi.org/10.1080/03098770500037655

Kinchin, I. M. (2014). Concept mapping as a learning tool in higher education: a critical analysis of recent reviews. *The Journal of Continuing Higher Education, 62*(1), 39–49. https://doi.org/10.1080/07377363.2014.872011

Koponen, I. T., & Pehkonen, M. (2008). Physics concepts and laws as network-structures: Comparisons of structural features in experts' and novices' concept maps. In *Concept Mapping: Connecting Educators-Proceedings of the 3rd International Conference on Concept Mapping*, *2*, 540–547. Retrieved February 14, 2023, from http://cmc.ihmc.us/cmc2008Papers/cmc2008-p179.pdf

Koponen, I. T., & Pehkonen, M. (2010). Coherent knowledge structures of physics represented as concept networks in teacher education. *Science & Education, 19*(3), 259–282. https://doi.org/10.1007/s11191-009-9200-z

Liu, P. L. (2011). A study on the use of computerized concept mapping to assist ESL learners' writing. *Computers & Education, 57*(4), 2548–2558. https://doi.org/10.1016/j.compedu.2011.03.015

Liu, Z., Kong, X., Liu, S., Yang, Z., & Zhang, C. (2022). Looking at MOOC discussion data to uncover the relationship between discussion pacings, learners' cognitive presence and learning achievements. *Education and Information Technologies, 27*, 8265–8288. https://doi.org/10.1007/s10639-022-10943-7

Lopez, E. J., Shavelson, R. J., Nandagopal, K., Szu, E., & Penn, J. (2014). Ethnically diverse students' knowledge structures in first-semester organic chemistry. *Journal of Research in Science Teaching, 51*(6), 741–758. https://doi.org/10.1002/tea.21160

Morine-Dershimer, G. (1993). Tracing conceptual change in preservice teachers. *Teaching and Teacher Education, 9*(1), 15–26. https://doi.org/10.1016/0742-051X(93)90012-6

Nainggolan, R., Perangin-angin, R., Simarmata, E., & Tarigan, A. F. (2019). Improved the performance of the K-Means cluster using the sum of squared error (SSE) optimized by using the elbow method. *Journal of Physics: Conference Series*, *1361*(1), 012015. https://doi.org/10.1088/1742-6596/1361/1/012015

Novak, J. D., & Gowin, D. B. (1984). *Learning how to learn*. Cambridge University Press.

Novak, J. D. (2010). *Learning, creating, and using knowledge: Concept maps as facilitative tools in schools and corporations*. Routledge.

Rajabi, Z., Shehu, A., & Purohit, H. (2019). User behavior modelling for fake information mitigation on social web. In: Thomson, R., Bisgin, H., Dancy, C., Hyder, A. (Eds), *Social, Cultural, and Behavioral Modeling. Lecture Notes in Computer Science* (pp. 234–244). https://doi.org/10.1007/978-3-030-21741-9_24

Roessger, K. M., Daley, B. J., & Hafez, D. A. (2018). Effects of teaching concept mapping using practice, feedback, and relational framing. *Learning and Instruction, 54*, 11–21. https://doi.org/10.1016/j.learninstruc.2018.01.011

Ruiz-Primo, M. A. (2000). On the use of concept maps as an assessment tool in science: what we have learned so far. *Revista Electrónica de Investigación Educativa, 2*(1), 29–53.

Schaal, S., Bogner, F. X., & Girwidz, R. (2010). Concept mapping assessment of media assisted learning in interdisciplinary science education. *Research in Science Education, 40*(3), 339–352. https://doi.org/10.1007/s11165-009-9123-3

Spector, J. M. (2006). A methodology for assessing learning in complex and ill-structured task domains. *Innovations in Education and Teaching International, 43*(2), 109–120. https://doi.org/10.1080/14703290600650368

Tang, H., & Clariana, R. (2017). Leveraging a sorting task as a measure of knowledge structure in bilingual settings. *Technology Knowledge and Learning, 22*(1), 23–35. https://doi.org/10.1007/s10758-016-9290-z

Tlili, A., Wang, H., Gao, B., Shi, Y., Zhiying, N., Looi, C. K., & Huang, R. (2021). Impact of cultural diversity on students' learning behavioral patterns in open and online courses: a lag sequential analysis approach. *Interactive Learning Environments*, 1–20. https://doi.org/10.1080/10494820.2021.1946565

Wadouh, J., Liu, N., Sandmann, A., & Neuhaus, B. J. (2014). The effect of knowledge linking levels in biology lessons upon students' knowledge structure. *International Journal of Science and Mathematics Education, 12*(1), 25–47. https://doi.org/10.1007/s10763-012-9390-8

Walker, J. M. T., & King, P. H. (2003). Concept mapping as a form of student assessment and instruction in the domain of bioengineering. *Journal of Engineering Education, 92*(2), 167–178. https://doi.org/10.1002/j.2168-9830.2003.tb00755.x

Weinerth, K., Koenig, V., Brunner, M., & Martin, R. (2014). Concept maps: a useful and usable tool for computer-based knowledge assessment? A literature review with a focus on usability. *Computers & Education, 78*, 201–209. https://doi.org/10.1016/j.compedu.2014.06.002

Watson, M. K., Pelkey, J., Noyes, C. R., & Rodgers, M. O. (2016). Assessing conceptual knowledge using three concept map scoring methods: assessing conceptual knowledge using concept map scoring. *Journal of Engineering Education, 105*(1), 118–146. https://doi.org/10.1002/jee.20111

Wei, X., Saab, N., & Admiraal, W. (2021). Assessment of cognitive, behavioral, and affective learning outcomes in massive open online courses: a systematic literature review. *Computers & Education, 163*, 104097. https://doi.org/10.1016/j.compedu.2020.104097

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Xiuling He[1] · Jing Fang[2] · Hercy N. H. Cheng[3] · Qibin Men[1] · Yangyang Li[2]

[1] National Engineering Laboratory For Education Big Data, Central China Normal University, Wuhan City, Hubei Province, China

[2] National Engineering Research Center for E-learning, Central China Normal University, Wuhan City, Hubei Province, China

[3] Center for General Education, Taipei Medical University, Taipei City, Taiwan