Visual Informatics 6 (2022) 22–33

Contents lists available at ScienceDirect

Visual Informatics

journal homepage: www.elsevier.com/locate/visinf

Towards a better understanding of the role of visualization in online learning: A review

Gefei Zhang, Zihao Zhu, Sujia Zhu, Ronghua Liang, Guodao Sun *

College of Computer Science and Technology, Zhejiang University of Technology, Hangzhou, 310023, China

ARTICLE INFO

Article history:
Available online 5 September 2022

Keywords:
Visualization in education
Online learning
Visual analytics

ABSTRACT

With the popularity of online learning in recent decades, MOOCs (Massive Open Online Courses) are increasingly pervasive and widely used in many areas. Visualizing online learning is particularly important because it helps to analyze learner performance, evaluate the effectiveness of online learning platforms, and predict dropout risks. Due to the large-scale, high-dimensional, and heterogeneous characteristics of the data obtained from online learning, it is difficult to find hidden information. In this paper, we review and classify the existing literature for online learning to better understand the role of visualization in online learning. Our taxonomy is based on four categorizations of online learning tasks: behavior analysis, behavior prediction, learning pattern exploration and assisted learning. Based on our review of relevant literature over the past decade, we also identify several remaining research challenges and future research work.

1. Introduction

Massive open online courses (MOOCs) provide learners with various learning resources with the growth of online learning. MOOCs develop rapidly and become widely used by learners over the past decade. Compared to traditional classroom learning, online learning provides learners with more flexible choices. MOOCs platforms store plenty of weblogs about learning behaviors, including structured and unstructured data containing spatio-temporal attributes.

However, data obtained from MOOCs platforms is complex, massive, and noisy, which poses a challenge for analyzing online learning datasets. Although pre-processed, data items only show who did what when and where which lacks a more intuitive and systematic analysis of the learner. Visualization techniques can be able to incorporate empirical understanding and information processing tasks of pedagogical experts. In this way, decision-makers are provided with intuitive illustrations of patterns hidden in learner activities and friendly interactions as they explore the datasets.

As far as we know, this is the first survey to explore learning from a visual analytics perspective. By analyzing the requirements of the various roles of learners, instructors, and administrators. After discussions with domain experts, we categorize the online learning tasks, we classify online learning tasks into the following four categories: *behavior analysis, behavior prediction, learning pattern exploration and assisted learning*. We summarize five data types using traditional visual analytics pipelines: *network data, text data, high-dimensional data, spatio-temporal data, and multimedia data*. Second, we extract six common visualization techniques, including *network visualization, text visualization, temporal visualization, geographic visualization, chart visualization, and glyph visualization*. We conclude with a description of the four frequent categories of interactive methods, including *selection & exploration, filtering & navigation, connection & saving, and encoding & reconfiguring*.

The major contributions of this paper are as follows. First, we provide a comprehensive survey of recent developments in online learning in visualization with the expectation that it will assist researchers in understanding the preliminary work on online learning visualization. Second, this paper provides a systematic classification of the existing literature and identifies new research challenges and trends that contribute to an enhanced understanding of the field and provides value for online learning practitioners. We develop a web-based survey browser to help users understand taxonomy in this paper. (https://zjutvis.github.io/OL4VIS).

2. Related work

This section provides an overview of the surveys related to the visualization of online learning. Visualization techniques can play an significant role in analyzing online learning data. Qu and Chen (2015) analyze six typical visual analytics systems for

* Corresponding author.
E-mail address: godoor.sun@gmail.com (G. Sun).

https://doi.org/10.1016/j.visinf.2022.09.002

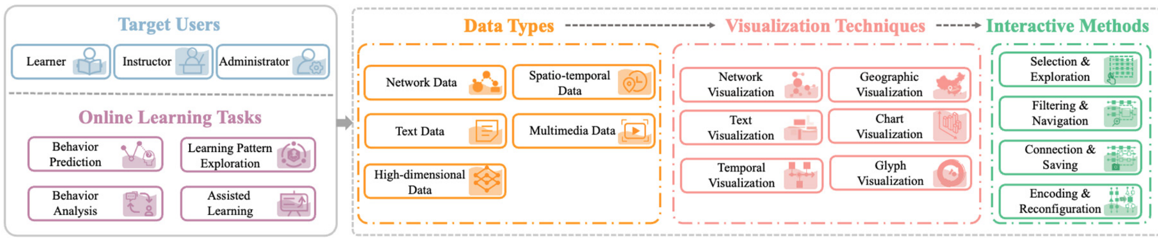

Fig. 1. Taxonomy of this survey, focusing on target users, data types, visualization techniques, and interactive methods in the visual analytics of online learning tasks. The watermark is displayed as a percentage of the total articles for this attribute.

online learning while exploring some challenges that arise when analyzing MOOCs data, and summarize possible future research directions. A more detailed classification for online learning is provided by Asli et al. (2020) and Dewan et al. (2020). They both mention monitoring dropout status, assessing course quality, exploring learner activity, and assessing learner performance. Moreno-Marcos et al. (2018) surveys the state of the art on prediction in MOOCs through a Systematic Literature Review. Results show there is strong interest in predicting dropouts in MOOCs. In contrast, our review focuses on learning analytics and revolves around visualization techniques applied to online learning. Keim et al. (2008) propose a definition and analysis pipeline for visual analytics, which inspires our formalization of the design space, which we discuss in Section 3. Some researchers attempt to gain insight into educational data in online learning. For example, Vieira et al. (2018) review existing visual learning analytics, analyzing the methods, audiences, purposes, contexts, and data sources used by designers and researchers to visualize educational data, which inspired our coding of the literature. Emmons et al. (2017) compare the various data provided by the MOOCs platform and proposed a new workflow for data analysis and visualization, which was optimized and validated in the Information Visualization MOOC (IVMOOC) at Indiana University. Some researchers (Wu et al., 2019) also focus on online learning forums and review three major platform forum studies, including performance prediction, understanding emotions, and predicting social behaviors.

Different from existing work that summarizes specific visualization related to educational data in online learning, our work aims at better understanding the role of visualization in online learning and covers a broader scope than existing surveys. We believe that this review will provide a richer resource and comprehensive overview for relevant practitioners to advance future research in this area.

3. Methodology and taxonomy

In this section, we describe our methodology for selecting papers for survey topic and create our taxonomy of online learning behaviors regarding data types, visualization techniques, and interactive analysis methods (see Fig. 1).

3.1. Methodology

This survey aims to gain a better understanding of the role of visualization in online learning, and to provide a comprehensive review of existing research, we collected relevant papers from visualization journals, conferences, and the field of education. First, we perform a search-driven retrieval method. We select five publishers (Willey Online Library, ACM Digital Library, Springer, IEEE Xplore, ScienceDirect) to collect papers broadly. In this process, we work with a set of keywords related to online learning (e.g. online education, e-learning, MOOCs, distance learning, online learning) and visualization (e.g. visual analytics and visualization). We select five visualization conferences (IEEE VAST, IEEE InfoVis, ACM CHI, EuroVis, IEEE PacificVis), two visualization journals (IEEE TVCG, Elsevier VI), three educational journals (IEEE TLT, Elsevier C & E, Elsevier CHB), and one educational conference (ACM LAK) for secondary searches. By combining these keywords in pairs to collect a wide range of papers. We initially obtain 422 papers. Second, we engage in reference-driven literature selection with the help of Connected Papers,[1] utilizing our prior knowledge of the core technologies on the topic as a starting point to expand the scope of the work by searching for cited and referenced references. After two rounds of collecting papers in this way, we obtain 465 papers. Next, papers are further filtered and analyzed according to the following criteria. We primarily searched the titles and abstracts of each paper in the last decade.

First, we validate online learning and the paper should focus on using visualization or visual analytics techniques to assist in analyzing issues in 'Online Learning' (Misailidis et al., 2018; Xia et al., 2022). For example, Han et al. (2021) propose the HisVA system for history learning. Second, we validate the 'VIS', so we exclude work such as Brinton et al. (2014), Pérez-Álvarez et al. (2018). In the work of this survey, we not only examine visualization in online learning but also analyze the utilization of visualization techniques to assist in online learning, which allows us to work with a wider range of online learning.

In terms of coding the survey, we propose a preliminary coding approach based on our knowledge of the existing literature and discuss it with educational experts. To code the collected papers from a visual perspective, we refer to the traditional visual analytics pipeline (Keim et al., 2008), based on three key components: data types, visualization techniques, and interactive methods. The experts suggest that we divide the traditional pedagogical perspective into three perspectives: student, instructor, and administrator. Analyze whether the paper is intended to enhance student learning, improve teaching and learning for instructors or assist administrators in better managing the online learning platform. In terms of classifying online learning tasks, experts suggest three aspects of learning behavior, learning patterns, and learning predictions. However, in our analysis of the paper, we found that assisting online learning tools and the analysis of the course rather than the student could not be ignored. After further discussions with experts, we have added the category of assisting courses. This survey was independently coded by both authors and discussed the different codes. After many discussions and adjustments, we finally arrived at five data types, six types of visualization techniques, and four types of interactive methods. This classification is explained in detail in Section 3.2. Most of the papers in the collection employ interactive methods to aid visualization. However, there are some papers that employ only visualization techniques and do not mention interactive methods.

[1] https://www.connectedpapers.com/

3.2. Taxonomy

Online learning tasks. Online learning data has much finer granularity than traditional classroom education records. Online education platforms record a considerable amount of learner activity, as well as mouse click stream events.

Behavior Analysis(BA) is the task of extracting feature sets related to learning behaviors from learners' behavioral data for analysis, we classify learners' learning behaviors into collaborative learning behaviors and autonomous learning behaviors. Collaborative learning behaviors include analyzing learners' posting behavior in MOOCs forums and exploring learners' answers to questions online. Autonomous learning behaviors include the learning behaviors that learners perform by watching videos. Some researchers focus with studying learners' video clickstreams (Zhao et al., 2018).

Learning Pattern Exploration(LPE) aims to discover how learners learn in the learning process. We divide learning pattern exploration into learning path analysis and student digital portrait. The learning paths of all effective learners can be generated by clustering analysis, which can explore learners learning preferences in order to provide learners with better learning path planning (Xia et al., 2019a). Analysis of learner statement sentiment in the forum, assessment of learner problem-solving skills, and expressive learner assessments build a digital portrait of the learner (Xia et al., 2021).

Behavior Prediction(BP) includes performance prediction and exception prediction. Performance prediction includes predicting learners' grades and predicting learners' performance in the problem pool (Wei et al., 2020). Learner exception prediction includes predicting learners' dropout behaviors (Chen et al., 2016b) and learners' learning patterns (Mu et al., 2019).

Assisted Learning(AL) includes but is not limited to visualization tools made to assist teaching on the MOOCs and analysis of the course in general. Visualization tools can greatly improve the efficiency with which students acquire knowledge of more difficult imagery (Ilves et al., 2018). At the same time, the analysis of course helps instructors and administrators to have a grasp of the overall curriculum and compare the learning impact of different teaching methods (Citra and Wahyuni, 2021).

To construct a structured, comprehensive taxonomy, we define a design space for describing each online learning task. We utilized a traditional visualization analysis pipeline (Keim et al., 2008) around three key components: data types, visualization techniques, and interactive methods. In addition, we are also inspired by other surveys (Guo et al., 2021; Shi et al., 2020) to specifically categorize data types, visualization techniques, and interactive methods based on the collected papers.

Data types. According to the different data attributes, this paper summarizes five typical data types, including network data, text data, high-dimensional data, spatio-temporal data, and multimedia data. *Network data(N)* is commonly utilized for social network analysis without hierarchical relationships (He et al., 2018). *Text data(T)* is provided by the content of the learners' postings in the forum (Fu et al., 2017), including semantic information about the learners. *High-dimensional data(HD)* contains multiple independent attributes to describe the learners' attributes. *Spatio-temporal data(ST)* includes a large amount of behavioral data containing temporal information. *Multimedia data(M)* is composite data formed by text, graphics, images, and other media data. Researchers analyze MOOCs videos to predict learner performance and enable teachers to take steps to intervene quickly (Mubarak et al., 2021).

Visualization techniques. In this paper, the literatures are divided into six parts in visualization techniques.

Network Visualization(NV) is a common visualization form in online learning, which analyzes the social relationships among groups of learners. Learner group network visualizes by nodes and edges, which enable effective observation of the behavior of the learner group (Mubarak et al., 2021; Schwab et al., 2017). Common visual representations are node-link diagrams and force-guide diagrams. *Text Visualization(TeV)* mainly includes two visualization methods, one is text visualization based on word frequency statistics, the other is text visualization based on semantics. Exploring learners' posting behavior and posting content in forums requires semantic analysis of text (Chen et al., 2016a). *Temporal Visualization(TV)* emphasizes the process of developing content evolution based on time. It is common to record learners' learning sequences to analyze learners' behavioral preferences (Xia et al., 2020b). Common visual representations are stacked flow charts and chord charts. *Geographic Visualization(GeoV)* addresses the multidimensional, dynamic, and associative features implied in geospatial data for analysis. Geospatial data are typically applied to analyze learner distribution and

Table 1
Examples of papers on visualization and visual analytics of online learning are sorted by time.

Publication	Time	L	I	A	BA	LPE	BP	AL	N	T	HD	ST	M	NV	TeV	TV	GeoV	CV	GV	S&E	F&N	C&S	E&R
Mcgrath et al. [49]	2011		■	■	■	■		■			■	■				■		■			■		
Trimm et al. [73]	2012		■			■					■					■		■		■			
Denny et al. [16]	2013		■		■	■					■	■				■		■	■	■			
Nickels et al. [58]	2013	■						■			■								■	■	■		
Xiaohuan et al. [96]	2013		■	■	■	■					■	■				■		■		■	■		
Auvinen et al. [4]	2015	■	■				■	■			■						■	■		■			
Buder et al. [6]	2015	■			■					■	■							■		■	■	■	
Gomez et al. [24]	2015	■	■	■	■	■					■	■				■		■	■	■	■		
Karavirta et al. [34]	2015	■	■					■			■			■				■		■			■
Li et al. [42]	2015		■			■					■	■						■	■	■	■		
Liu et al. [43]	2015	■			■	■					■	■				■		■		■			
Shi et al. [66]	2015		■		■	■						■	■			■	■	■		■	■	■	
Vivian et al. [79]	2015		■	■	■	■				■	■	■			■	■		■		■		■	
Atapattu et al. [3]	2016		■		■					■	■				■				■		■		■
Bull et al. [7]	2016	■			■			■	■	■	■	■		■	■	■		■	■	■			
Chen et al. [9]	2016		■	■	■					■	■	■			■		■		■	■	■	■	
Chen et al. [11]	2016		■	■			■			■	■	■			■	■		■	■	■	■		
Jordao et al. [33]	2016	■	■				■	■	■		■	■		■		■			■	■	■	■	
Kuosa et al. [38]	2016		■		■				■	■	■	■		■	■	■		■		■	■	■	
Kwon et al. [39]	2016			■		■					■							■		■	■	■	■
Wang et al. [80]	2016		■		■						■	■	■			■		■	■	■			
Fu et al. [22]	2017		■	■	■					■	■	■		■	■	■		■	■	■		■	
Schwab et al. [65]	2017	■			■				■	■	■			■					■	■		■	■
El et al. [19]	2018			■	■				■	■	■	■		■	■			■		■	■	■	
Fu et al. [21]	2018			■	■					■		■			■	■			■	■	■	■	
He et al. [30]	2018			■	■	■				■	■	■					■	■	■	■		■	
Ilves et al. [32]	2018							■			■	■	■		■			■		■			
Martins et al. [44]	2018		■				■	■			■	■				■		■		■	■	■	
Matin et al. [45]	2018	■						■			■							■	■	■			
Misailidis et al. [51]	2018		■	■	■			■			■							■		■			
Schubert et al. [64]	2018	■	■		■	■				■								■		■			
Suntiwichaya et al. [71]	2018		■		■						■	■					■	■		■	■		
Venkatarayalu et al. [76]	2018	■						■		■						■		■		■			■
Wong et al. [86]	2018		■		■	■			■	■	■			■	■			■		■	■	■	
Xia et al. [89]	2018		■	■	■							■	■			■	■			■	■		
Zarra et al. [100]	2018		■		■					■					■			■			■		
Zhao et al. [102]	2018	■			■					■			■	■	■	■		■		■		■	■
He et al. [28]	2019		■	■		■					■	■				■	■	■		■			
He et al. [29]	2019	■			■					■	■		■			■		■		■			
Miyakita et al. [52]	2019	■			■	■						■	■			■		■		■	■	■	
Mu et al. [53]	2019		■			■	■			■		■			■			■	■	■		■	
Tubman et al. [75]	2019		■	■				■		■					■					■	■		■
Vidakis et al. [77]	2019		■		■						■	■						■		■	■		
Xia et al. [91]	2019			■		■						■						■	■	■			■
Xia et al. [93]	2019		■			■					■	■				■		■	■	■		■	
Chen et al. [10]	2020	■	■			■					■	■				■		■		■	■	■	■
Xia et al. [94]	2020		■	■		■					■	■				■		■	■	■	■		
Li et al. [41]	2021		■	■	■	■					■		■					■	■	■	■	■	
Mubarak et al. [55]	2021		■	■	■		■				■	■		■		■		■		■	■		■
Xia et al. [92]	2021		■	■	■	■					■					■		■	■	■	■	■	■
Tsung et al. [74]	2022					■	■				■	■				■		■		■		■	■
Xia et al. [95]	2022		■			■		■			■	■				■	■	■		■	■		

learner characteristics in different regions in online learning visualization (Chen et al., 2016a). *Chart Visualization(CV)* is frequently utilized to describe the distribution of learner characteristics and sequences of learning behaviors. Researchers color-code scatter plot points to express whether learners' postings are positive or negative (Wong et al., 2015; Xia et al., 2020a). *Glyph Visualization(GV)* represents the properties of different data variables through a set of visual attributes (shape, size, color, orientation). Fu et al. (2018) design group glyphs and set glyphs to describe the multifaceted nature of group features.

Interactive methods are the process of user-data conversation that enhances the user's understanding of the visualization system. By analyzing the visual tasks of online learning, we group all interactive methods into the following four categories: *Selection & Exploration(S&E)*, *Filtering & Navigation(F&N)*, *Connection & Saving(C&S)*, and *Encoding & Reconfiguration(E&R)*. Users can click and hover to explore the learning behavior of interest (Li et al., 2021). Zooming and panning are used to discover discrete points in clusters of learning behaviors and thus navigate to users with unusual behavior (Mu et al., 2019). In behavior analysis, when users need to select clusters of learners for comparative analysis, they can take snapshots to save the current class of learners' learning behaviors and compare the differences between the two classes of learners' learning behaviors by connecting (Chen et al., 2020). Users choose different encoding schemes interactively to observe the impact of each aspect on learners (Xia et al., 2019a), providing managers with different perspectives on the distribution of learners to analyze the causes of learner behavior.

Overall, based on a non-exhaustive search of existing literature review related to visualization for online learning, we summarize four types of online learning tasks, including learning behavior analysis, learning behavior prediction, learning pattern exploration, and assisted learning. For each online learning task, we have specifically broken down the common data types, visualization techniques, and interactive methods. The example papers are shown in Table 1.

4. Behavior analysis

Behavior analysis is highly significant for students, instructors, and administrators. In brief, learning behavior analysis assists students in reflecting on and improving their learning performance, helps instructors to adjust their lesson plans, supports administrators in comparing different teaching methods and better adapts platform resources. We classify learning behaviors into collaborative learning behaviors and autonomous learning behaviors. The example papers are shown in Fig. 2. We observe a large amount of *text data* in learner behavior analysis. It is due to the researchers exploring the content of students' postings in the forum. *Text visualization* is a favored visual representation for collaborative learning behaviors. In addition, *filtering & navigation* is also used for *text visualization*.

4.1. Data types

As the behavior of students learning to interact on online forums is concerned with the transfer and exchange of information, the network is frequently seen when students interact (Wong et al., 2018). *Network data* is available to describe the relationship between keywords in the forum (Hsu et al., 2017; Wong and Li, 2016), students and forum information (Wong et al., 2018), the hierarchical relationship between forum topics (Bull et al., 2016; El-Assady et al., 2018; Kuosa et al., 2016), interactions between students (Tervakari et al., 2014), learning behaviors between students the transformation (Coffrin et al., 2014), and the interaction between students and problems (Wei et al., 2020). *Text data* such as keywords and forum content assist instructors in understanding student learning behavior and knowledge acquisition. As it provides information on text clusters including student emotions (Schubert et al., 2018), study notes (Nakayama et al., 2012) and more. *Text data* plays an active role in forum content research (Atapattu et al., 2016; El-Assady et al., 2018; Wei et al., 2020; Wong and Li, 2016). The analysis of text data can provide instructors with keywords for students in the forum to understand misconceptions (Hsu et al., 2017) and monitor students' semantic information in group work (Vivian et al., 2015). In addition,researchers explore the course information (Chen et al., 2016a), course content (Schwab et al., 2017; Tervakari et al., 2014; Zhao et al., 2018), and learning materials (Kuosa et al., 2016). *Spatio-temporal data* are crucial for describing student learning behavior. *Spatio-temporal data* include log data about student interaction behavior (Fu et al., 2018; Gómez-Aguilar et al., 2015; Mazza and Dimitrova, 2007; Wong et al., 2015), video click streams (He et al., 2019b; Mubarak et al., 2021; Shi et al., 2015; Wang et al., 2016), problem solving trajectories (Wei et al., 2020; Xia et al., 2021, 2020b), mentioning exchange data. Liu et al. (2014). Meanwhile, *spatio-temporal data* include spatial data such as students' geographical distribution (Chen et al., 2016a; He et al., 2018; Suntiwichaya et al., 2018). *High-dimensional data* is used to analyze student learning behavior include student profiles (Xia et al., 2021), the number of posts (Gómez-Aguilar et al., 2015), and student grades (Xiaohuan et al., 2013). *Multimedia data* is one of the essential data for online learning. Students' main learning behaviors depend on multimedia data to be generated. (Hasnine et al., 2021; Muñoz-Merino et al., 2015), course videos (Miyakita et al., 2019; Shi et al., 2015; Xia and Wilson, 2018), invigilation video (Li et al., 2021), interactive slides,course notes (Nakayama et al., 2012). These behavioral data are well worth studying.

4.2. Visualization techniques

Collaborative learning behaviors refers to the learning behavior of students interacting or discussing with others, as opposed to independent learning behavior. For this survey, we categorize student discussion in forums (Oliveira et al., 2010) and group work (Wong et al., 2018) as collaborative learning. Node-link diagrams are common network visualization views. The node-link visualization supports the discovery of connections in discussion threads from forum interactions (Wong et al., 2018). The ThreadReconstructor (El-Assady et al., 2018) was proposed by El-Assady et al. The tool visualizes a threaded conversation in a discussion forum, with the relationship between replies to two posts represented by two nodes and the arc between them. *Text visualization* is applied to reveal the topic of a forum post (Atapattu et al., 2016). The keywords that appear more frequently are amplified through the word cloud so that users can quickly discover valid information (Hsu et al., 2017). Log data with a temporal attribute can be displayed with a spiral timeline (Gómez-Aguilar et al., 2015) to show the global time pattern of the course projected each week. *Chart visualization* includes bubble charts (Zarra et al., 2018) and pie charts (McGrath, 2011) to depict discussion topics and user session clusters. A novel embedded glyph visualization that responds to students' problem-solving logic, engagement, and problems encountered (Xia et al., 2021). To explore user groups in forum discussions.

Autonomous learning behaviors are student-led behaviors that result in continuous change (improvement and sublimation of knowledge and skills, methods and processes, emotions and values) for the individual through reading, listening, and inquiry. There are some specific examples of independent learning, such as students watching videos (He et al., 2019b), answering questions (Minematsu et al., 2020; Xia et al., 2020a), and game-based

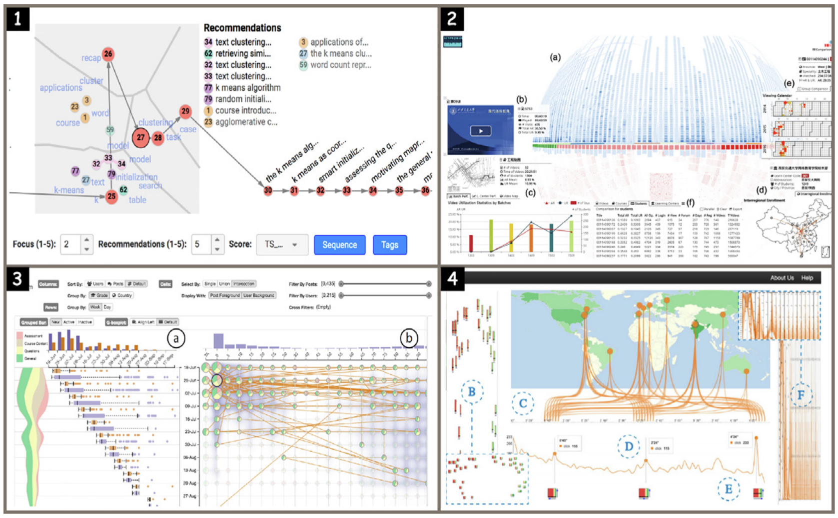

Fig. 2. Visualizations of learning behavior analysis. (1) The links between lecture videos can be represented as a network (Zhao et al., 2018). (2) VUSphere (He et al., 2018) applies chart visualizations to analyze student video utilization. (3) iForum system (Fu et al., 2017) is designed to visualize MOOCs forums. (4) PeakVizor (Chen et al., 2016a) visualizes the interaction peaks in the MOOCs video clickstream.

learning (Vidakis et al., 2019). *Text visualizations* can present basic course information to assist instructors in exploring student learning behaviors to keep abreast of course content (Chen et al., 2016a). Animated narrative visualizations (Wang et al., 2016) allow presenting video clickstream data. Xia and Wilson (2018) use heatmaps presenting student engagement in course videos. *Geographic visualization* is typically utilized to locate the geographic location of active users in the system (He et al., 2018; Suntiwichaya et al., 2018). *Chart visualization* such as bar charts (Xiaohuan et al., 2013), line charts (He et al., 2019b) and scatter diagrams (Muñoz-Merino et al., 2015) are common in the act of exploring independent learning. They are commonly used to describe the efficiency of students' learning. Learners' behavior can be presented using a glyph visualization. For example, Li et al. (2021), design a suspect type of overall risk glyph to assess student behavior in examinations.

4.3. Interactive methods

Visual analytics of learning behavior takes *selection & exploration* as the first step in the analysis. Users mark information of interest by clicking on (Mazza and Botturi, 2007; Schubert et al., 2018) and swiping (Gibbs et al., 2006; Mazza and Dimitrova, 2003). *Filtering & navigation* allow users the flexibility to analyze the data of interest, including performing zooms (Mazza and Dimitrova, 2003; Oliveira et al., 2010) and pans (Wong et al., 2018). Buder et al. (2015) evaluate three visual filters for navigating the forum content to help learners find the forum content of interest efficiently and quickly. *Connection & saving* method is usually implemented in systems with multiple coordinated views (He et al., 2018; Xia et al., 2021), and iForum supports interactive linking to associate the matrix diagram with the top bar (Fu et al., 2017). When the administrator has a basic understanding of the learning behavior, coding and reassignment is performed. Atapattu et al. (2016) explore different visualization views by changing variables.

5. Learning pattern exploration

Learning model exploration aims to find ways in which students learn across a range of learning behaviors. Learning pattern exploration can be further divided into learning path analysis and student digital profiling. Student path analysis is based on the learning process of students in online learning platforms (Zou et al., 2020), including behavior pattern analysis in the writing process (Liu et al., 2014) and process analysis in the systematic question and answer (Xia et al., 2020a). Student digital profile is used to provide a summary description of students by tagging learners and facilitating teachers or administrators to accurately provide instructional support and services to different groups of learners. Specifically, we categorize issues into solving skills (Xia et al., 2021), student performance (Coffrin et al., 2014; Xu et al., 2021), grades (Li et al., 2015; Xiaoya et al., 2009), and emotions (Schubert et al., 2018) as indicators that make up the digital portrait of students. The example papers are shown as Fig. 3. *Spatio-temporal data* and *high-dimensional data* account for a large portion of learning model exploration. Therefore, *temporal visualization* and *chart visualization* are used for exploring learning paths and student digital portrait analysis.

5.1. Data types

Network data refers to the trajectory of students' participation in numerous learning activities (Coffrin et al., 2014; Tervakari et al., 2014). *Text data* associated with the learning mode is mainly designed to describe basic student information and learning content, such as student ID (Mu et al., 2019) and practice questions (Xia et al., 2021). *High-dimensional data* enriches the appearance of students and contributes to the construction of a digital portrait of them (Xia et al., 2022; Xiaohuan et al., 2013; Xiaoya et al., 2009) explore the impact of emotions on academic performance and categorize students' emotions during learning into six positive emotions (enjoyment, hope, pride, confidence, excitement, and interest) and six negative emotions (anxiety, anger, shame, hopelessness, boredom and frustration). *Spatio-temporal data* is essential for describing learning paths (He et al., 2018; Liu et al., 2014; McGrath, 2011). Hasnine et al. (2021) analyze multimedia data such as student participation in interactive lecture videos to extract students' emotions, detect student engagement and enrich the digital portrait of students.

5.2. Visualization techniques

Learning paths refer to the routes and sequences of learning activities. In online learning, student learning paths are commonly explored to analyze student learning preferences (Cui

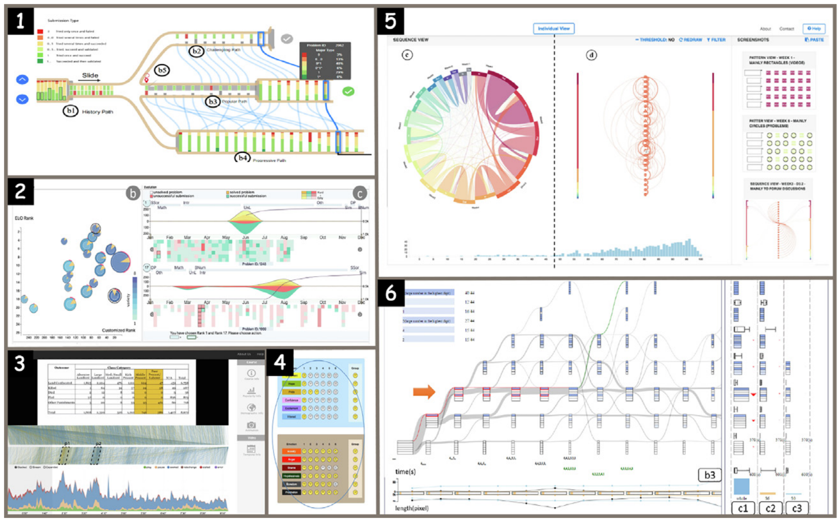

Fig. 3. Visualizations of learning pattern exploration. (1) BlockLens describes changes in learning sequences over time using sankey diagrams (Tsung et al., 2022). (2) Researchers (Xia et al., 2020b) use glyph visualization to compare different learning paths. (3) The calendar view (Shi et al., 2015) shows the popularity of the selected video over time. (4) Ruiz et al. (2016) design the emoticons to capture students' attention. (5) ViSeq system (Chen et al., 2020) explore learning patterns with chord diagrams. (6) Xia et al. (2021) deform a diagram to describe the logic of student problem solving.

et al., 2021; Li et al., 2015), learning styles (Williams and Conlan, 2007) and to extract salient learning paths to generalize a category of learners (Chen et al., 2020). *Network visualization* provides a clear picture of the relationships among learning behaviors. The utility of state transition diagrams as well reveals how differences in curriculum and assessment design affect student engagement patterns (Coffrin et al., 2014). Learning paths are normally achieved and visualized by converting time series into symbolic representations. LearnerVis (He et al., 2019a) visualizes the learning process of time characteristics, it allows users to customize student groups to compare differences in student engagement and time management. *Geographical visualization* is applied to show the distribution of learners (Shi et al., 2015). Nakayama et al. (2012) use line charts to visualize the extent of coverage of instructor terminology to understand students' learning paths in the course. Students are represented by glyphs (Li et al., 2015). Researchers design a student score based on a musical pentatonic score (Chang et al., 2022) to help users gain a general impression of students.

Student digital portrait is an abstraction of labeled student models based on student learning data. Student portraits are divided into student group portraits and individual student portraits (Vivian et al., 2015). The student group portrait is based on a clustering system that clusters the behavioral characteristics of different groups and provides multi-dimensional data analysis overlooking the characteristics of the student group (Xiaohuan et al., 2013). The individual student profile is the 'big picture' of the students' personal appearance so that administrators can tailor their teaching to the students' individual labeling characteristics and ultimately achieve personalized management (Chen et al., 2020). *Network visualization* contributes to the description of the students' image in social relationships (Tervakari et al., 2014). Graphical layouts can identify and visualize groups of students with similar characteristics and thus explore patterns of student success, failure, and repetition. By visualizing time-series data such as students' logbook data and learning behaviors, students' probability of dropping out of school (Wortman and Rheingans, 2007) and their performance (Gómez-Aguilar et al., 2015; Xia et al., 2022; Tervakari et al., 2014; Trimm et al., 2012) can be analyzed. Ruiz et al. (2016) design innovative square-based visualizations to show the balance of negative/positive emotions for individuals and groups.

5.3. Interactive methods

Instructors select and explore data to discover student learning path characteristics (Denny, 2013; Shi et al., 2015) and support users to retrieve detailed course outlines after selecting a course (Mu et al., 2019). By filtering, navigating, and reconfiguring various visual forms, learning patterns that are not easily detected are uncovered. Gómez-Aguilar et al. (2015) utilize advanced interactive filtering to display enhanced information about student engagement. Moreover, the filtering allows for quick filtering out of unwanted information. The administrator selects parameters for parallel coordinate mapping, sorts the attributes by clicking on the axis heads, and reorders them by dragging the axes (Kwon and Lee, 2016). MOOCad (Mu et al., 2019) supports associative interaction between the exploration view and the personal path view. BlockLens (Tsung et al., 2022) and ViSeq (Chen et al., 2020) include a snapshot panel that allows administrators to easily capture both holistic and individual sequence views. In addition, they both support switching between individual and multiple student learning modes. The visual analytics system proposed by Xia et al. (2019b) supports both correlation analysis across attributes and detailed visualization of user mouse movements.

6. Behavior prediction

Behavior prediction has significant implications for learners. Retention rates for MOOCs are poor due to the diversity of online learners (Moreno-Marcos et al., 2018). We can divide behavior predictions in MOOCs into two categories, including student performance prediction and learner exception prediction. Student performance prediction focuses on certificate earner (Cobos and Palla, 2017), score prediction (Zou et al., 2020), and learning path prediction (Xu et al., 2021). Student exception prediction includes dropout and mood (Shi et al., 2021). The example papers are shown in Fig. 4. In behavior prediction the data are more derived from *spatio-temporal data*, utilizing *temporal visualization* to analyze the learning sequences over time.

6.1. Data types

Behavior prediction analysis is commonly used in sequence datasets for anomaly detection (Mu et al., 2019), automatic assessment (Li et al., 2020), and procrastination prediction (Xu et al.,

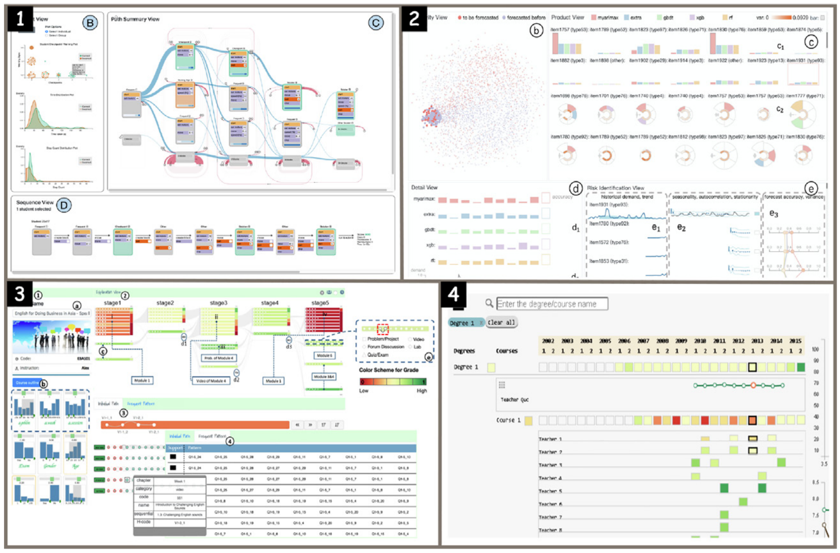

Fig. 4. Visualizations of learning behavior prediction. (1) PeerLens (Xia et al., 2019a) recommends customized and adaptable practice question sequences for individual learners. (2) The glyph employs a circular design with a fan-encoded prediction model (Sun et al., 2020). (3) The learning sequences are analyzed by an anomaly detection algorithm (Mu et al., 2019). (4) Martins et al. (2018) analyze the effects of presenting students' illustrations of their behavior in an online learning environment.

2021). MOOCs platforms develop rapidly and produce a great number of datasets. Researchers (Li et al., 2020; Mu et al., 2019; Wei et al., 2020; Weiand and Manssour, 2015) use types of techniques to extract activities in MOOCs data. These activities form sequence datasets, which hidden the characteristics and regular patterns of students' behaviors. Analyzing these patterns can help teachers implement formation teaching measures to improve the quality of student performance (Tsung et al., 2022). The characteristics of students' interactions and the similarity between online questions are taken into consideration (Wei et al., 2020) to strengthen the prediction quality of students performance. Behavior prediction analysis is widely used for high-dimensional datasets to implement teaching tasks and enhance learning. To enhance teaching and learning, Zou et al. (2020) use machine learning techniques and clustering algorithms to predict students' score levels for carrying an early warning. Facing the evolution academic path of students and courses, Auvinen et al. (2015) discover hidden patterns and trends to help teachers rapidly analyze and improve the teaching–learning process. Behavior prediction analysis is ubiquitous in network datasets for clustering and relationship mining. Wong and Li (2016) use a text-mining analytical technique and construct KeyGraphs to predict students' academic performances level.

6.2. Visualization techniques

Network visualization assists teachers in monitoring learner behavior to predict learner performance and take timely intervention. Mubarak et al. (2021) employ a bipartite graph layout (Zhao et al., 2021) to draw the node-link graph, where the learners and video nodes represent two separate groups in the graph, and each edge represents their interaction (i.e. the learners' behaviors in the video clickstream data). Each cluster of nodes is marked with a different color to indicate group membership and how the cluster interacts with a particular video (Mubarak et al., 2021). Wortman et al. set the start of each course as a 'virtual' node and drew a visual view of the network by linking data from the previous course to the next course for understanding students' propensity to drop out (Wortman and Rheingans, 2007). *Text visualization* is commonly applied in the predictive analysis of MOOCs forum data (Chen et al., 2016b; Mu et al., 2019). Wong and Li (2016) utilize a software tool called Polaris to visualize KeyGraphs. The visualization of learning sequences is a crucial part of predictive analysis. Some researchers have described changes in learning sequences over time using sankey diagrams with circular links (Tsung et al., 2022). Auvinen et al. utilize heat maps to show predictions of success based on student behavior (Auvinen et al., 2015). Kayanda and Machuve (2020) display primary and secondary school dropouts through geographic visualization. where users can move the slider to see the dropouts by year and hover over specific areas to see specific numbers. *Chart visualization* is an indispensable part of predictive analysis. Bars are provided to represent the dropout rates of boys and girls in each region (Kayanda and Machuve, 2020), the different distributions of learners with the same attributes (e.g. average time spent studying per week) (Mu et al., 2019; Xu et al., 2021), statistics showing the total number of posts per day (Chen et al., 2016b), and the characteristics mentioned on the course videos (Mubarak et al., 2021). Shi et al. chart the relationship between help-seeking emotions and students' enrollment performance in online learning (Shi et al., 2021). The advantages of visualizing the learning process are evident, as it can create conditions for learning and reduce dropout rates by providing feedback to students about learning behaviors, learning content, learning activities, and learning communities. Examples include radar charts presenting task schedules, course learning, and line charts presenting learning engagement trend charts (Zou et al., 2020). Researchers design *glyph visualizations* to depict the predictive performance of different models. The bar-based and doughnut-based designs used in Dropoutseer represent weekly learning patterns and the average number of videos viewed by students, with modifications to the common pie chart used to represent different click popularity and distribution performance along the timeline (Chen et al., 2016b).

6.3. Interactive methods

To select groups of students based on one or more similarities, researchers typically use queries to assist users in exploring interactively (Wortman and Rheingans, 2007). MOOCad (Mu et al., 2019) permits users to learn sequences with attribute constraints by querying them and linking all views together to assist users in identifying the corresponding elements. Researchers have found that providing timely feedback (e.g. highlighting, fading in and out, and highlighting) after swiping and filtering data attributes can be beneficial to users who are unfamiliar with the system (Chen et al., 2016b). The researcher collapsed the data to save space by clicking on a line that would expand to show details of all students (Weiand and Manssour, 2015). Filtering apply to analyze all attributes of the clustered learners (Chen et al., 2016b). EduVis (Jordão et al., 2016) allows information filters to be added and removed at any stage. To enable administrators to focus on specific choices (Martins et al., 2018), a filtering mechanism has developed by Martins et al. Similarly, the filtering mechanism enables searches for specific students or resources (Weiand and Manssour, 2015) Connection and saving reduce pointless switching back and forth. To facilitate comparisons between individual students and multiple students, BlockLens (Tsung et al., 2022) allows instructors to switch between 'select individual' and 'select group' modes. To assist users in differentiating predicted metric performance, DFSeer (Sun et al., 2020) utilizes a circular glyph design as the default option. Users can interactively switch to a bar chart based glyph design. At the same time, DFSeer provides three sliders for users to adjust the weight, variance, and number of prediction accuracy. Reallocation facilitates predictive analytics from a different point of view (Weiand and Manssour, 2015; Chen et al., 2016b). The axes in the DropoutSeer system are ranked according to their importance in the current prediction model and also support the user to adjust the order of the axes (Chen et al., 2016b). In the analysis of student clustering, the user is allowed to reconfigure the clustering results.

7. Assisted learning

Assisted learning is designed to support students in applying visual tools to enhance learning and assist instructors and administrators in analyzing the status of courses and projects. At the same time, visualization effectively contributes to students learning lessons that are not easily understood, such as algorithms (Karavirta and Shaffer, 2015; Matin et al., 2018), and molecular structures (Nickels et al., 2013). We observe that assisted learning tends to target both instructor and learner, and online learning tools play a role in both the instructor's teaching and the learners' learning process. Most of the data in assisted learning task is *high-dimensional*, and the researchers design *glyph visualization* to improve the readability of subject knowledge.

7.1. Data types

Text data provides an evaluation of the online platform by online learners (Tubman et al., 2019) and includes a huge amount of teaching resources (Venkatarayalu, 2018; Zhang and Sun, 2008). *High-dimensional data* consists of multiple attributes containing information about the context of the course. List of courses, list of students, course professor, and grades for course as examples of high-dimensional data to assist analysis (Martins et al., 2018). *Spatio-temporal data* provides details of timestamps and IP addresses. MOOCViz (Dernoncourt et al., 2013) categorizes students according to the IP addresses from which they access the website and explores how different groups of students in the course utilize different types of resources. Zhang et al. visualize knowledge on a range of multimedia data, such as text, video, audio, flash, and PowerPoint, to assist students in easily understanding the relationships between concepts (Zhang and Sun, 2008). Kim and Xia (2022) select MOOC videos from five online learning platforms and propose guidelines for designing mobile-friendly MOOC.

7.2. Visualization techniques

The node-link diagram visualizes the course structure. To assist administrators in understanding the course model, Jordão et al. (2016) propose a solution based on two linked views, one based on node-links and the other on a multi-matrix representation. *Text visualization* for analysis of problem sets (Ilves et al., 2018), MOOCs contributions (Tubman et al., 2019). Ilves et al. conducted an experimental analysis of text visualizations and graphical visualizations, and the study show that graphical visualizations had better learning outcomes (Ilves et al., 2018). *Temporal visualization* reveals the overall learning progression of the course (Anderson et al., 2014; Jordão et al., 2016; Martins et al., 2018). For example, Xia et al. (2022) present two line charts side by side to compare and assist students in understanding their overall learning progress. *Geographic visualization* shows the ratio of the number of students who receive a certificate in each country/region for both courses to the number of students enrolled in the course (Dernoncourt et al., 2013). Heat map visualization of error frequencies to show the frequency of learning activities (Mazza and Dimitrova, 2007; Xia et al., 2020a, 2022) and to understand the important concepts in the course. Additional chart visualizations, such as linear charts, are provided to track changes in the distribution of tutors and tutors Xia et al. (2022). Novel *glyph visualization* increases students' motivation and interest in learning (Anderson et al., 2014; Matin et al., 2018; Paiva et al., 2018). For example, Anderson et al. design badges as a motivational tool for participation in the MOOCs, and the study found that the use of badges increased student forum participation (Anderson et al., 2014).

7.3. Interactive methods

A variety of visual tools assist users in learning (Matin et al., 2018; Paiva et al., 2018; Tubman et al., 2019; Venkatarayalu, 2018) and exploring course information (Ilves et al., 2018; Jordão et al., 2016). When the administrator clicks on the node of interest, relevant data that was not originally visualized is displayed. Martins et al. (2018) analyze student course and degree information by selecting a course by clicking and dragging, which would cause all other courses to disappear, thus allowing the administrator to focus on a specific selection. Filtering is commonly apply to filter the data (Martins et al., 2018; Tubman et al., 2019). RLens (Xia et al., 2022) provides a filter panel from different perspectives (e.g. tutor, topic, session, date, etc.) to facilitate continuous tracking of learners' progress. PresentaBALL (Nickels et al., 2013) supports user-driven navigation, including redirection to other text sections and activation of modeling functions such as highlighting sections of the molecule, changing the representation to a more detailed level, or moving the camera to a predefined viewpoint for pressure. A few visualization tools support associative saving. The system proposed by Martins et al. (2018) allows users to analyze cross-sectional data already to understand the evolution of degrees and courses over time. Additionally, the system also supports drop-down lists to filter degree information. The user interactively adjusts parameters to change the visualization (Venkatarayalu, 2018), for example by providing data to the algorithm to change (Karavirta and Shaffer, 2015). Tubman et al. (2019) regroup the MOOC contributions into a word cloud, when users click on a word, the word cloud is redrawn according to the selected word.

8. Discussion and challenges

In this section, we discuss our findings and challenges regarding data types, visualization techniques, and interactive methods across in the four online learning tasks.

Data types. *High-dimensional data* can effectively analyze online learning tasks from multiple perspectives, providing a variety of features for detecting online learning tasks. Learning behavior is often presented through data with temporal and spatial attributes, it is often used in combination with other data types. The above two data types account for a relatively high percentage of all data. *Network data* are mostly used for behavioral analysis and are widely available in social networks between instructors–students, and students–students. *Multimedia data*, which consists mainly of video data, is an important component of MOOCs platforms. Analysis of media data can include learner learning activities and identify anomalies for a timely response. A large amount of *text data* exists in MOOCs forums and learning materials, and the mining of text data can reveal hidden learning patterns.

Challenges of data type. There are limitations to the data currently available for the analysis of online learning. Obtaining more comprehensive information about learners remains a significant challenge, and as a result, learners cannot be contextualized to provide a more in-depth explanation of learner behavior.

Visualization techniques. Among visualization, *chart visualization* and *temporal visualization* are more common, which also coincide with more high-dimensional and temporal data. Researchers use chart visualization to represent learner learning behaviors, and *chart visualization* can visually explain learners' learning patterns. *Temporal visualizations* are popular in showing learner learning behaviors over time because it maximizes the explanation of learner learning habits. In comparing visualization techniques applicable to online learning tasks, we found that *glyph visualization* is suitable for visualizations for behavior analysis and learning pattern exploration. Intelligent visualization design can better satisfy the needs of users in exploring online learning.

Challenges of visualization techniques. Online learning tasks can contain multiple heterogeneous temporal events. For example, MOOCs platform data is diversified and includes text, images, and videos. These data record each type of event at different sampling rates and do not display different patterns of events. It brings challenges to aggregate data from different sources. Most existing visualization techniques choose to display a selection of data, which is challenging to perform integrated analysis of multiple learning tasks. To solve this problem, a visual analytics framework needs to be developed.

Interactive methods. *Selected & exploration* has been the most popular interactive analysis method in visual analytics of online learning tasks. Most assisted learning tools to provide the user with an overview and user drill down to the details as required. The second popular interactive method is *filtering & navigation*, which fits the visual information seeking mantra: "Overview first, zoom and filter, then details-on-demand" (Shneiderman, 2003). Observing different online learning tasks, visualization systems that study learning pattern exploration often use filtering to mine learners for significant learning patterns. *Connection & saving* is employed in comparing differences in learning behavior across learners.

Challenges of interactive methods. As the complexity of data is exponentially increased, interactive methods play an important role in visual analytics techniques. *Encoding & reconfiguration* allows to leverage and process data, which is a deeper level of data-human interaction. *Encoding & reconfiguration* is of interest to the visualization community. However, few of these interactions have been applied to current research. For visual analytics, it is a challenge to effectively utilize interactive methods.

Overall, most online learning visualization research is designed to help instructors and administrators better understand the recommendation and predicted dropout rates. Fewer visualization works are designed to help learners understand their learning for improvement. Most studies in visualization of online learning have focused on behavior analysis, followed by exploring learning patterns and assisted learning, with relatively few studies focusing on behavior prediction.

9. Future work

With the rapid growth of the internet, online learning is becoming increasingly popular and many researchers have realized the importance of applying visualization techniques to understand online learning. However, there are still challenges to be addressed in future research. We discuss several directions that we believe are important and promising.

Learner Isolation. In our discussions with experts, they mention that a large part of the high dropout rate is due to learner isolation. Online learning amplifies the isolation of learners and it is essential to address how learners feel isolated in the learning process. We can utilize visualization techniques and interactive methods to help learners feel less isolated and increase retention of online learning.

Multiple Forecast. For online learning dropout prediction, new machine learning techniques (Yuan et al., 2021; Shah et al., 2021) and predictive models can enhance the predictive power. Researchers can consider additional predictive outcomes such as predicting learning efficiency (whether students learn at a good pace), and learner expectations. More importantly, most researchers focus on univariate predictions and cannot predict multiple variables and establish links between them.

Instructor Behavior Analysis. We found that most of the analysis of online learning behaviors was directed at student behavior. There is less analysis of the instructor's clickstream on the online platform. The instructor's repost rate on the forum and the time spent uploading the course have an impact on the learners. In future work, more analysis of faculty members could be conducted to help administrators evaluate faculty members.

Enhanced Interpretability. Researchers provide personalized advice for MOOCs learners, most of which ignore interpretability and cannot inform learners why they are receiving such advice, or empower users to modify it. For example, Zhao et al. (2018) assist flexible learning through semantic visual exploration and opportunity sequences for MOOCs video recommendation. However, this study only applies to MOOCs videos that have a predetermined order and the user's choice does not contribute to the next recommendation for system.

10. Conclusion

In this paper, we analyze relevant papers on the latest advances and have conducted an in-depth study of online learning tasks based on expert advice. Our survey presents trends and preferences for online learning tasks, data types, visualization techniques, and interactive methods. We also discuss challenges and future work in this paper which provides a forward-looking perspective to inspire readers. Finally, we design a web-based exploration browser for users to facilitate locating papers of their interests. We believe that our work provides a good overview of the use of visualization in online learning.

CRediT authorship contribution statement

Gefei Zhang: Conceptualization, Writing - Original draft. **Zihao Zhu:** Writing - Reviewing and Proofreading. **Sujia Zhu:** Writing - Reviewing and Editing. **Ronghua Liang:** Guidance, Reviewing, Proofreading. **Guodao Sun:** Supervision, Idea evaluation, Writing - Reviewing and Editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Ethical Approval

This study does not contain any studies with human or animal subjects performed by any of the authors. All data used in the study are taken from public databases that were published in the past.

Acknowledgments

This work is partly supported by the National Natural Science Foundation of China (61972356, 62036009). Guodao Sun is the corresponding author. We would like to thank Professor Jihong Ding for her valuable opinions, Wenjing Wang for her design and discussion about taxonomy, and all anonymous reviewers for their feedback.

References

Anderson, A., Huttenlocher, D., Kleinberg, J., Leskovec, J., 2014. Engaging with massive online courses. In: Proceedings of the International Conference on World Wide Web. pp. 687–698.
Asli, M.F., Hamzah, M., Ibrahim, A.A.A., Ayub, E., 2020. Problem characterization for visual analytics in MOOC learner's support monitoring: A case of Malaysian MOOC. Heliyon 6 (12), e05733.
Atapattu, T., Falkner, K., Tarmazdi, H., 2016. Topic-wise classification of MOOC discussions: A visual analytics approach. In: International Educational Data Mining Society. ERIC.
Auvinen, T., Hakulinen, L., Malmi, L., 2015. Increasing students' awareness of their behavior in online learning environments with visualizations and achievement badges. IEEE Trans. Learn. Technol. 8 (3), 261–273.
Brinton, C.G., Chiang, M., Jain, S., Lam, H., Liu, Z., Wong, F.M.F., 2014. Learning about social learning in MOOCs: From statistical analysis to generative model. IEEE Trans. Learn. Technol. 7 (4), 346–359.
Buder, J., Schwind, C., Rudat, A., Bodemer, D., 2015. Selective reading of large online forum discussions: The impact of rating visualizations on navigation and learning. Comput. Hum. Behav. 44, 191–201.
Bull, S., Ginon, B., Boscolo, C., Johnson, M., 2016. Introduction of learning visualisations and metacognitive support in a persuadable open learner model. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. pp. 30–39.
Chang, B., Sun, G., Li, T., Huang, H., Liang, R., 2022. MUSE: Visual analysis of musical semantic sequence. IEEE Trans. Vis. Comput. Graphics 1.
Chen, Q., Chen, Y., Liu, D., Shi, C., Wu, Y., Qu, H., 2016a. PeakVizor: Visual analytics of peaks in video clickstreams from massive open online courses. IEEE Trans. Vis. Comput. Graphics 22 (10), 2315–2330.
Chen, Y., Chen, Q., Zhao, M., Boyer, S., Veeramachaneni, K., Qu, H., 2016b. DropoutSeer: Visualizing learning patterns in massive open online courses for dropout reasoning and prediction. In: IEEE Conference on Visual Analytics Science and Technology. pp. 111–120.
Chen, Q., Yue, X., Plantaz, X., Chen, Y., Shi, C., Pong, T.-C., Qu, H., 2020. ViSeq: Visual analytics of learning sequence in massive open online courses. IEEE Trans. Vis. Comput. Graphics 26 (3), 1622–1636.
Citra, K., Wahyuni, F., 2021. Exploring demographic variations of freshmen to online learning anxiety: A data visualization analysis based approach. In: International Research Symposium on Advanced Engineering and Vocational Education. pp. 33–38.
Cobos, R., Palla, V.M., 2017. edX-MAS: Model analyzer system. In: Proceedings of the 5th International Conference on Technological Ecosystems for Enhancing Multiculturality. pp. 1–7.
Coffrin, C., Corrin, L., de Barba, P., Kennedy, G., 2014. Visualizing patterns of student engagement and performance in MOOCs. In: Proceedings of the Fourth International Conference on Learning Analytics and Knowledge. pp. 83–92.
Cui, Y., Song, X., Hu, Q., Li, Y., Shanthini, A., Vadivel, T., 2021. Big data visualization using multimodal feedback in education. Comput. Electr. Eng. 96, 107544.
Denny, P., 2013. The effect of virtual achievements on student engagement. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 763–772.
Dernoncourt, F., Taylor, C., O'Reilly, U.-M., Veeramachaneni, K., Wu, S., Do, C., Halawa, S., 2013. MoocViz: A large scale, open access, collaborative, data analytics platform for MOOCs. In: NIPS Workshop on Data-Driven Education, Lake Tahoe, Nevada, USA.
Dewan, M., Pachon, W.M., Lin, F., 2020. A review on visualization of educational data in online learning. In: Learning Technologies and Systems. pp. 15–24.
El-Assady, M., Sevastjanova, R., Keim, D., Collins, C., 2018. ThreadReconstructor: Modeling reply-chains to untangle conversational text through visual analytics. In: Computer Graphics Forum. Vol. 37. (3), pp. 351–365.
Emmons, S.R., Light, R.P., Börner, K., 2017. MOOC visual analytics: Empowering students, teachers, researchers, and platform developers of massively open online courses. J. Assoc. Inform. Sci. Technol. 68 (10), 2350–2363.
Fu, S., Wang, Y., Yang, Y., Bi, Q., Guo, F., Qu, H., 2018. VisForum: A visual analysis system for exploring user groups in online forums. ACM Trans. Interact. Intell. Syst. 8 (1).
Fu, S., Zhao, J., Cui, W., Qu, H., 2017. Visual analysis of MOOC forums with iForum. IEEE Trans. Vis. Comput. Graphics 23 (1), 201–210.
Gibbs, W.J., Olexa, V., Bernas, R.S., 2006. A visualization tool for managing and studying online communications. J. Educ. Technol. Soc. 9 (3), 232–243.
Gómez-Aguilar, D.A., Hernández-García, Á., García-Peñalvo, F.J., Therón, R., 2015. Tap into visual analysis of customization of grouping of activities in elearning. Comput. Hum. Behav. 47, 60–67.
Guo, Y., Guo, S., Jin, Z., Kaul, S., Gotz, D., Cao, N., 2021. A survey on visual analysis of event sequence data. IEEE Trans. Vis. Comput. Graphics 1.
Han, D., Parsad, G., Kim, H., Shim, J., Kwon, O.-S., Son, K.A., Lee, J., Cho, I., Ko, S., 2021. HisVA: a visual analytics system for learning history. IEEE Trans. Vis. Comput. Graphics 1.
Hasnine, M.N., Bui, H.T., Tran, T.T.T., Nguyen, H.T., Akçapınar, G., Ueda, H., 2021. Students' emotion extraction and visualization for engagement detection in online learning. Procedia Comput. Sci. 192, 3423–3431.
He, H., Dong, B., Zheng, Q., Di, D., Lin, Y., 2019a. Visual analysis of the time management of learning multiple courses in online learning environment. In: 2019 IEEE Visualization Conference. pp. 56–60.
He, H., Dong, B., Zheng, Q., Li, G., 2019b. VUC: Visualizing daily video utilization to promote student engagement in online distance education. In: Proceedings of the ACM Conference on Global Computing Education. pp. 99–105.
He, H., Zheng, O., Dong, B., 2018. VUSphere: Visual analysis of video utilization in online distance education. In: IEEE Conference on Visual Analytics Science and Technology. pp. 25–35.
Hsu, H.-H., Huang, N.-F., Chen, S.-C., Lee, C.-A., Tzeng, J.-W., 2017. Misconceptions mining and visualizations for Chinese-based MOOCs forum based on NLP. In: Proceedings of IEEE International Conference on Big Data Analysis. pp. 634–639.
Ilves, K., Leinonen, J., Hellas, A., 2018. Supporting self-regulated learning with visualizations in online learning environments. In: Proceedings of the 49th ACM Technical Symposium on Computer Science Education. pp. 257–262.
Jordão, V.R., Gama, S., Gonçalves, D., 2016. Visualizing sequential educational datamining patterns. Int. J. Creative Interfaces Comput. Graph. 7 (1), 1–18.
Karavirta, V., Shaffer, C.A., 2015. Creating engaging online learning material with the jsav javascript algorithm visualization library. IEEE Trans. Learn. Technol. 9 (2), 171–183.
Kayanda, A.M., Machuve, D., 2020. A web-based data visualization tool regarding school dropouts and user asssesment. Eng. Technol. Appl. Sci. Res. 10 (4), 5967–5973.
Keim, D., Andrienko, G., Fekete, J.-D., Görg, C., Kohlhammer, J., Melançon, G., 2008. Visual analytics: Definition, process, and challenges. In: Information Visualization. pp. 154–175.
Kim, J., Xia, M., 2022. Mobile-friendly content design for MOOCs: Challenges, requirements, and design opportunities.
Kuosa, K., Distante, D., Tervakari, A., Cerulo, L., Fernandez, A., Koro, J., Kailanto, M., 2016. Interactive visualization tools to improve learning and teaching in online learning environments. Int. J. Distance Educ. Technol. 14 (1), 1–21.
Kwon, B.C., Lee, B., 2016. A comparative evaluation on online learning approaches using parallel coordinate visualization. In: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems. pp. 993–997.
Li, H., Wei, H., Wang, Y., Song, Y., Qu, H., 2020. Peer-inspired student performance prediction in interactive online question pools with graph neural network. In: Proceedings of the ACM International Conference on Information Knowledge Management. CIKM '20, pp. 2589–2596.
Li, H., Xu, M., Wang, Y., Wei, H., Qu, H., 2021. A visual analytics approach to facilitate the proctoring of online exams. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. pp. 1–17.
Li, X., Zhang, X., Liu, X., 2015. A visual analytics approach for e-learning education. In: International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing. pp. 34–40.
Liu, M., Calvo, R.A., Pardo, A., Martin, A., 2014. Measuring and visualizing students' behavioral engagement in writing activities. IEEE Trans. Learn. Technol. 8 (2), 215–224.

Martins, T., Gonçalves, D., Gama, S., 2018. Visualizing historical patterns in large educational datasets. Int. J. Creative Interfaces Comput. Graph. 9 (1), 32–48.

Matin, M.A., Oliullah, S.S.M., Polash, M.M.A., 2018. Implementation of a customizable algorithm visualization tool for E-learning. In: Proceedings of the International Conference on Education and E-Learning. pp. 32–36.

Mazza, R., Botturi, L., 2007. Monitoring an online course with the GISMO tool: A case study. J. Interact. Learn. Res. 18 (2), 251–265.

Mazza, R., Dimitrova, V., 2003. CourseVis: Externalising student information to facilitate instructors in distance learning. In: Proceedings of the International Conference in Artificial Intelligence in Education. pp. 117–129.

Mazza, R., Dimitrova, V., 2007. CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses. Int. J. Hum. Comput. Stud. 65 (2), 125–139.

McGrath, O.G., 2011. Visualizing user activity in open e-learning contexts: challenges and techniques for operational management. In: Proceedings of the Annual ACM SIGUCCS Conference on User Services. pp. 229–234.

Minematsu, T., Shimada, A., Taniguchi, R.-i., 2020. Visualization and analysis for supporting teachers using clickstream data and eye movement data. In: International Conference on Human-Computer Interaction. Springer, pp. 581–592.

Misailidis, E., Charitopoulos, A., Rangoussi, M., 2018. Visualization of educational data mined from the moodle e-learning platform. In: Proceedings of the Pan-Hellenic Conference on Informatics. pp. 82–87.

Miyakita, G., Arima, S., Yasui, M., Okawa, K., 2019. Exploring digital cultural heritage beyond MOOCs: Design, use, and efficiency of generous interfaces. In: IEEE Learning with MOOCS. pp. 42–46.

Moreno-Marcos, P.M., Alario-Hoyos, C., Muñoz-Merino, P.J., Kloos, C.D., 2018. Prediction in MOOCs: A review and future research directions. IEEE Trans. Learn. Technol. 12 (3), 384–401.

Mu, X., Xu, K., Chen, Q., Du, F., Wang, Y., Qu, H., 2019. MOOCad: Visual analysis of anomalous learning activities in massive open online courses. In: EuroVis. Short Papers, pp. 91–95.

Mubarak, A.A., Cao, H., Zhang, W., Zhang, W., 2021. Visual analytics of video-clickstream data and prediction of learners' performance using deep learning models in MOOCs' courses. Comput. Appl. Eng. Educ. 29 (4), 710–732.

Muñoz-Merino, P.J., Ruipérez-Valiente, J.A., Alario-Hoyos, C., Pérez-Sanagustín, M., Kloos, C.D., 2015. Precise effectiveness strategy for analyzing the effectiveness of students with educational resources and activities in MOOCs. Comput. Hum. Behav. 47, 108–118.

Nakayama, M., Mutsuura, K., Yamamoto, H., 2012. Visualization analysis of student's notes taken in a fully online learning environment. In: International Conference on Information Visualisation. pp. 434–439.

Nickels, S., Stöckel, D., Mueller, S.C., Lenhof, H.-P., Hildebrandt, A., Dehof, A.K., 2013. PresentaBALL—A powerful package for presentations and lessons in structural biology. In: 2013 IEEE Symposium on Biological Data Visualization. pp. 33–40.

Oliveira, A.P., Mealha, Ó., Santos, C., 2010. Visualisation of web based e-learning activity. In: International Conference Information Visualisation. pp. 219–224.

Paiva, R., Bittencourt, I.I., Lemos, W., Vinicius, A., Dermeval, D., 2018. Visualizing learning analytics and educational data mining outputs. In: International Conference on Artificial Intelligence in Education. pp. 251–256.

Pérez-Álvarez, R., Maldonado-Mahauad, J., Pérez-Sanagustín, M., 2018. Design of a tool to support self-regulated learning strategies in MOOCs. J. UCS 24 (8), 1090–1109.

Qu, H., Chen, Q., 2015. Visual analytics for MOOC data. IEEE Comput. Graph. Appl. 35 (6), 69–75.

Ruiz, S., Charleer, S., Urretavizcaya, M., Klerkx, J., Fernández-Castro, I., Duval, E., 2016. Supporting learning by considering emotions: tracking and visualization a case study. In: Proceedings of the Sixth International Conference on Learning Analytics & Knowledge. pp. 254–263.

Schubert, M., Durruty, D., Joyner, D.A., 2018. Measuring learner tone and sentiment at scale via text analysis of forum posts. In: Proceedings of the 8th Edition of the International Workshop on Personalization Approaches in Learning Environments.

Schwab, M., Strobelt, H., Tompkin, J., Fredericks, C., Huff, C., Higgins, D., Strezhnev, A., Komisarchik, M., King, G., Pfister, H., 2017. booc.io: An education system with hierarchical concept maps and dynamic non-linear learning plans. IEEE Trans. Vis. Comput. Graphics 23 (1), 571–580.

Shah, D., Patel, D., Adesara, J., Hingu, P., Shah, M., 2021. Integrating machine learning and blockchain to develop a system to veto the forgeries and provide efficient results in education sector. Vis. Comput. Ind. Biomed. Art 4 (1), 1–13.

Shi, C., Fu, S., Chen, Q., Qu, H., 2015. VisMOOC: Visualizing video clickstream data from massive open online courses. In: IEEE Pacific Visualization Symposium. pp. 159–166.

Shi, H., Li, Y., Hong, D., 2021. Characterizing academic help-seeking moods for enrollment performance of institutional online student. Procedia Comput. Sci. 192, 3885–3894.

Shi, Y., Liu, Y., Tong, H., He, J., Yan, G., Cao, N., 2020. Visual analytics of anomalous user behaviors: A survey. IEEE Trans. Big Data.

Shneiderman, B., 2003. The eyes have it: A task by data type taxonomy for information visualizations. In: The Craft of Information Visualization. pp. 364–371.

Sun, D., Feng, Z., Chen, Y., Wang, Y., Zeng, J., Yuan, M., Pong, T.-C., Qu, H., 2020. Dfseer: A visual analytics approach to facilitate model selection for demand forecasting. In: Proceedings of the CHI Conference on Human Factors in Computing Systems. pp. 1–13.

Suntiwichaya, S., Khanti, P., Chunwijitra, S., Tummarattananont, P., Wutiwiwatchai, C., 2018. Improving data analytics visualization with advancing information for MOOCs: A case study on ThaiMOOC platform. In: 2018 15th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology. pp. 98–101.

Tervakari, A.-M., Silius, K., Koro, J., Paukkeri, J., Pirttilä, O., 2014. Usefulness of information visualizations based on educational data. In: Proceedings of IEEE Global Engineering Education Conference. pp. 142–151.

Trimm, D., Rheingans, P., et al., 2012. Visualizing student histories using clustering and composition. IEEE Trans. Vis. Comput. Graphics 18 (12), 2809–2818.

Tsung, S., Wei, H., Li, H., Wang, Y., Xia, M., Qu, H., 2022. BlockLens: Visual analytics of student coding behaviors in block-based programming environments. In: Proceedings of the Ninth ACM Conference on Learning @ Scale. pp. 299–303.

Tubman, P., Oztok, M., Benachour, P., 2019. New platform affordances for encouraging social interaction in MOOCs: The" comment discovery tool" interactive visualisation plugin. In: Proceedings of International Conference on Advanced Learning Technologies. Vol. 2161. pp. 34–36.

Venkatarayalu, N., 2018. Interactive visualization-based E-learning aids for vector calculus. In: IEEE International Conference on Teaching, Assessment, and Learning for Engineering. pp. 725–729.

Vidakis, N., Barianos, A.K., Trampas, A.M., Papadakis, S., Kalogiannakis, M., Vassilakis, K., 2019. In-game raw data collection and visualization in the context of the ''ThimelEdu'' educational game. In: International Conference on Computer Supported Education. Springer, pp. 629–646.

Vieira, C., Parsons, P., Byrd, V., 2018. Visual learning analytics of educational data: A systematic literature review and research agenda. Comput. Educ. 122, 119–135.

Vivian, R., Tarmazdi, H., Falkner, K., Falkner, N., Szabo, C., 2015. The development of a dashboard tool for visualising online teamwork discussions. In: IEEE International Conference on Software Engineering. Vol. 2. pp. 380–388.

Wang, Y., Chen, Z., Li, Q., Ma, X., Luo, Q., Qu, H., 2016. Animated narrative visualization for video clickstream data. In: SIGGRAPH Asia Symposium on Visualization. pp. 1–8.

Wei, H., Li, H., Xia, M., Wang, Y., Qu, H., 2020. Predicting student performance in interactive online question pools using mouse interaction features. In: Proceedings of the International Conference on Learning Analytics Knowledge. LAK '20, pp. 645–654.

Weiand, A., Manssour, I.H., 2015. Towards visual analysis techniques for monitoring students of distance education courses. In: Proceedings of the Workshop on Visual Analytics, Information Visualization and Scientific Visualization.

Williams, F.P., Conlan, O., 2007. Visualizing narrative structures and learning style information in personalized e-learning systems. In: IEEE International Conference on Advanced Learning Technologies. pp. 872–876.

Wong, G.K., Li, S.Y., 2016. Academic performance prediction using chance discovery from online discussion forums. In: IEEE Annual Computer Software and Applications Conference. Vol. 1. pp. 706–711.

Wong, J.-S., Pursel, B., Divinsky, A., Jansen, B.J., 2015. An analysis of MOOC discussion forum interactions from the most active users. In: International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction. Springer, pp. 452–457.

Wong, J.-S., et al., 2018. MessageLens: a visual analytics system to support multifaceted exploration of MOOC forum discussions. Vis. Inform. 2 (1), 37–49.

Wortman, D., Rheingans, P., 2007. Visualizing trends in student performance across computer science courses. In: Proceedings of the SIGCSE Technical Symposium on Computer Science Education. pp. 430–434.

Wu, M., Dewan, M.A.A., Lin, F., Murshed, M., 2019. Visualization of course discussion forums: A short review from online learning perspective. In: 2019 IEEE Canadian Conference of Electrical and Computer Engineering. CCECE, pp. 1–4.

Xia, M., Asano, Y., Williams, J.J., Qu, H., Ma, X., 2020a. Using information visualization to promote students' reflection on" gaming the system" in online learning. In: Proceedings of the Seventh ACM Conference on Learning@ Scale. pp. 37–49.

Xia, M., Sun, M., Wei, H., Chen, Q., Wang, Y., Shi, L., Qu, H., Ma, X., 2019a. PeerLens: Peer-inspired interactive learning path planning in online question pool. In: Proceedings of the Conference on Human Factors in Computing Systems. CHI '19, pp. 1–12.

Xia, M., Velumani, R.P., Wang, Y., Qu, H., Ma, X., 2021. QLens: Visual analytics of multi-step problem-solving behaviors for improving question design. IEEE Trans. Vis. Comput. Graphics 27 (2), 870–880.

Xia, M., Wei, H., Xu, M., Lo, L.Y.H., Wang, Y., Zhang, R., Qu, H., 2019b. Visual analytics of student learning behaviors on k-12 mathematics e-learning platforms. arXiv preprint arXiv:1909.04749.

Xia, J., Wilson, D.C., 2018. Instructor perspectives on comparative heatmap visualizations of student engagement with lecture video. In: Proceedings of the 49th ACM Technical Symposium on Computer Science Education. pp. 251–256.

Xia, M., Xu, M., Lin, C.-e., Cheng, T.Y., Qu, H., Ma, X., 2020b. SeqDynamics: Visual analytics for evaluating online -solving dynamics. In: Computer Graphics Forum. Vol. 39. (3), Wiley Online Library, pp. 511–522.

Xia, M., Zhao, Y., Hong, J., Erol, M.H., Kim, T., Kim, J., 2022. RLens: A computer-aided visualization system for supporting reflection on language learning under distributed tutorship. arXiv preprint arXiv:2204.08033.

Xiaohuan, W., Guodong, Y., Huan, W., Wei, H., 2013. Visual exploration for time series data using multivariate analysis method. In: International Conference on Computer Science & Education. pp. 1189–1193.

Xiaoya, G., Kan, L., Ping, L., 2009. Visual analysis of college students' scores in english test. In: International Conference on Computer Science & Education. pp. 1816–1819.

Xu, H., Qu, J., Ma, X., Ling, Y., 2021. Prediction and visualization of academic procrastination in online learning. In: International Conference on Distance Education and Learning. pp. 133–139.

Yuan, J., Chen, C., Yang, W., Liu, M., Xia, J., Liu, S., 2021. A survey of visual analytics techniques for machine learning. Comput. Vis. Media 7 (1), 3–36.

Zarra, T., Chiheb, R., Faizi, R., El Afia, A., 2018. Student interactions in online discussion forums: Visual analysis with LDA topic models. In: Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications. pp. 1–5.

Zhang, Y., Sun, X., 2008. The research on ontology-based knowledge visualization in E-learning resources management. In: 2008 Fourth International Conference on Semantics, Knowledge and Grid. pp. 495–496.

Zhao, J., Bhatt, C., Cooper, M., Shamma, D.A., 2018. Flexible learning with semantic visual exploration and sequence-based recommendation of MOOC videos. In: CHI '18, pp. 1–13.

Zhao, Y., Shi, J., Liu, J., Zhao, J., Zhou, F., Zhang, W., Chen, K., Zhao, X., Zhu, C., Chen, W., 2021. Evaluating effects of background stories on graph perception. IEEE Trans. Vis. Comput. Graphics 1.

Zou, M., Wang, T., Xu, H., Li, X., Wu, X., 2020. Using process visualization and early warning based on learning analytics to enhance teaching and learning. In: International Conference on Artificial Intelligence and Security. Springer, pp. 175–183.