IEEE Access
Multidisciplinary Rapid Review Open Access Journal

Received May 10, 2019, accepted June 6, 2019, date of publication June 24, 2019, date of current version July 25, 2019.
Digital Object Identifier 10.1109/ACCESS.2019.2924641

Real-Time Analysis Method and Application of Engagement in Online Independent Learning

SU MU[1], SHAOMING CHAI[2], HONGJIANG WANG[1], YILIN CHEN[1]
[1]School of Information Technology in Education, South China Normal University, China
[2]International Business College, South China Normal University, China
Corresponding author: Shaoming Chai (charmingchai@ibc.scnu.edu.cn)

ABSTRACT Engagement is an important factor influencing the effectiveness and quality of online learning programs and how satisfied online learners are with the online learning experience. Therefore, when developing online learning programs, ways to increase engagement should be one of the top priorities. To determine the parameters that increase engagement of independent online learners, this paper selects three types of quantitative data by which to assess online learning engagement: video playing options, video lecture viewing time, and concurrent learning behaviors. This data, collected from real-time observation and analysis of authentic online learning of 14,000 adult learners, was used to run a real-time growth algorithm that determined a few key parameters that increase independent-learning engagement. A few primary results of this study are: 1) real-time, dynamically calculated data representing general or individual engagement parameters of independent learning can help both teachers and learners be aware of, recognize, and adjust learning status accurately and effectively; 2) the algorithm can identify optimal parameters for online learning by analyzing numerical values of engagement, rules, and characteristics of online learning that are difficult to be observed directly; and 3) it is necessary to dynamically analyze learners' engagement based on their learning processes and behaviors in response to different video lectures to know the effectiveness and feasibility of such materials and to support the design, production, and modification of optimal learning materials.

INDEX TERMS Data analysis, engagement, independent learning, online learning, online learning behavior, video lecture.

I. INTRODUCTION

Online learning has become one of the most important approaches to formal and informal learning. However, online learning poses some challenges for teachers and students who do not know how to effectively engage in such a learning environment [1], [2]. For example, due to the lack of face-to-face interaction between teachers and students in online learning environments, students can find it difficult to concentrate and so do not absorb much of the information and thus turn in low-quality work. As student engagement is a necessary condition for learning, finding ways to engage students who lack direct face-to-face interaction with a teacher is essential for effective online learning. Therefore, it is necessary to analyze and study how students engage in online learning environments in order to help teachers facilitate timely support, help students reflect on their own learning, and promote student engagement during the online learning process.

Some related theories and empirical studies have studied a few of the challenges students experience in online learning environments, such as student satisfaction with learning support services for massive open online courses (MOOC) in China, student acceptance of such courses as legitimate formal learning environments, efficacy of blended learning models, and so on [3], [4]. To study these challenges, more researchers are starting to use online learning management systems to capture or record and analyze data on learners' learning behavior. Performing data analysis of this sort may play an important role in learning development, process management, teaching strategies application, and learning improvement of online learning environments. As noted in the white paper Big Data for Development: Opportunities and Challenges, published by the United Nations Global Pulse, big data has a significant impact in modern society [5]. In the

The associate editor coordinating the review of this manuscript and approving it for publication was Laxmisha Rai.

book of Empowering Online Learning: 100+ Activities for Reading, Reflecting, Displaying, and Doing, Curtis Bonk (2008) construct a comprehensive instructional design model as Read, Reflect, Display, Do(R2D2) [6]. And the book focuses on learning engagement, collaboration and interactivity, as well as learner autonomy, curiosity, product generation, etc., which shows that motivating online learners to engage in online activities is very important. Therefore, this paper analyzes quantitative online learning behavior data to better understand how engagement works in an online learning environment. Focusing on independent online learning, this paper proposes an algorithm to calculate parameters of online learning engagement and summarize the characteristics of online learning, which provides the scientific basis for designing and developing online learning materials and platforms that optimize student learning.

II. LITERATURE REVIEW

A. BEHAVIOR OF ONLINE INDEPENDENT LEARNING

Experience API (xAPI), an e-learning software specification, mainly details how to track and record learning experiences of online learners by capturing and recording personalized learning behaviors in online learning environments, including collaborative learning, mobile learning, and other kinds of learning activities [7]. xAPI has become the standard for describing online learning behavior via activity streams, which primarily include actor, verb, and object. Table 1 lists the 24 frequently-used verb categories that are used often to describe online learning activity behaviors, as stipulated in the xAPI standard [8].

TABLE 1. 24 Frequently Used Verb Categories in the xAPI Standard.

Answered	**Asked**	**Attempted**	Attended
Commented	**Completed**	Exited	Experienced
Failed	Imported	**Initialized**	Interacted
Launched	Mastered	Passed	Preferred
Progressed	Registered	**Responded**	**Resumed**
Scored	Shared	**Suspended**	**Terminated**

In online learning processes, learners usually participate in two types of learning: independent learning and collaborative learning. Independent learning, or self-directed/autonomous learning, is a process and a learning method whereby a learner acquires knowledge by his or her own efforts and develops the ability for inquiry and critical evaluation [9]. In online learning, a learner often controls their own learning pace and time according to a personal learning target they set for themselves. This part of learning includes the activities of watching video lectures, reading digital slides, writing notes, and so on. All these behaviors can be recorded by a learning management system, which captures statistics such as time spent on each activity and content of the notes. According to characteristics of independent learning, this paper used nine action verbs (marked in boldface in Table 1) to capture data revealing independent learning behaviors, and these nine verbs helped establish the parameters necessary for independent learning engagement.

B. ONLINE LEARNING ENGAGEMENT

Ralph Tyler (1935) proposed the concept of Time on Task as early as the 1930s. He argued that the more time students spent studying, the more they would benefit from learning. Henceforth, an increasing number of researchers began to conduct quantitative, qualitative, and empirical studies on learning engagement [10]. In the 1980s, Alexander W. Astin (1984) conducted academic studies on student involvement and proposed that student involvement refers to the amount of physical and psychological efforts a student devotes to academic activities [11]. Opinions surrounding this definition of learning engagement are controversial. Scholars, however, generally believe that learning engagement is composed of a three-dimensional index, namely behavior engagement, cognitive engagement, and emotional engagement [12]. Christenson et al. (2012) defined the three types of engagement. Behavior engagement refers to the degree of learners' attention, effort, and persistence. Cognitive engagement refers to the cognitive strategies adopted by learners in dealing with complex situations, such as intensive processing strategies rather than simple memory strategies. Emotional engagement refers to the positive emotions, such as interest, shown by learners in the process of task completion [13].

Throughout the literature, it is found that research on real-time dynamic analysis of online learning behavioral engagement is less frequent than studies focusing on cognitive engagement or emotional engagement, even though nowadays researchers can use log data and operating data captured and recorded by learning platform or APPs to develop algorithms that dynamically describe real-time behavioral engagement. Many researchers investigate possible relationships among learning variables and the three types of student learning engagement (behavioral engagement, emotional engagement and cognitive engagement) in online learning. And there are also many studies on the influencing factors of learning engagement based on data collected from questionnaires, rubric surveys, and interviews. For example, The Online Student Engagement Scale developed by Dixson and M. D. (2010) consists of four dimensions: skill engagement, emotional engagement, participation engagement, and performance engagement [14]. Dixon issued an OESE scale to 186 students from six universities in the Midwest of the United States by email and found that diverse interaction modes and task types could promote the efficient engagement of learners in online learning. In contrast, Sun and Rueda (2012) completed an online survey assessing students' levels of situational interest, computer self-efficacy, self-regulation, and engagement in distance education. They found that online activities and tools increase emotional engagement in online learning, although they do not necessarily increase behavioral or cognitive engagement [15]. Based on the perspective of behavior engagement, this paper analyzes independent learning data in order to provide a basic understanding of rules for online learning and to provide support for studies looking at cognitive engagement.

Throughout the literature on higher education, the term "engagement" in classroom learning or online learning describes a learner's study efforts and patterns, such as how they use time, resources, relationships, and interactions with their teacher, peers, or experts, to gain knowledge and understanding (Kahn, 2014; Trowler, 2010) [16], [17]. Engagement can be measured based on learners' time and efforts given to learning. So, from a behavioral perspective, engagement can be defined as the time and efforts students devote to educationally-purposeful activities (ACER, 2010) [18]. Thus, student engagement is a quantitative, rather than qualitative, measurement of student learning, and it includes several quantitative measures such as the quantification of student behavior, the quantification of cognitive engagement, and the quantification of emotions (Wang, 2017) [19]. At present, research has focused on studying student engagement in the context of theoretical models, explicit behavior statistics, and influencing factors and effect analysis. Yet these methods lack precise measurements of student engagement based on data concerning learning behavior and process.

Research evaluating international higher education has examined mainly the learning quality of students according to changes in a student's authentic learning engagement. Kuh (2001) proposed that learning engagement has two dimensions: (1) student involvement in the process of learning, and (2) student satisfaction with the self-learning status and supportive conditions offered by the school [20]. On this basis, Kuh and his scientific research team jointly designed the American National Survey of Student Engagement, which contained content describing three aspects to student engagement: five benchmarks of learning engagement, a learning gains scale, and student background characteristics. After a decade of development, the American National Survey of Student Engagement has received recognition from more and more scholars and has gained increasing influence around the world. Moreover, several countries, like Australia, New Zealand, and South Africa, also have used the American National Survey of Student Engagement to conduct studies. In addition, the Institute of Education from Tsinghua University in Beijing is introducing the American National Survey of Student Engagement to China to conduct a large-scale investigation [21].

Based on the existing School Engagement Measure (SEM)-MacArthur Network, Li Shuang et al.(2015) have conducted theoretical studies and made specific analyses. According to the characteristics of distance education and training programs, scholars have combined the three student learning dimensions of behavior, emotion, and cognition and then proposed a relatively complete Online Student Engagement Survey, which has since been completely revised. This study provides a reference standard for evaluating learning engagement [22]. Except when using such survey with scale, this paper aims to conduct studies on the methods of processing engagement analysis and to observe learners' independent engagement using behavior data to monitor behavior in real time and to provide an approach to improving online learning.

C. RESEARCH ON EFFECTIVENESS AND ENGAGEMENT OF LEARNING IN OPEN ONLINE COURSES

Over the past few years, with the development and implementation of MOOC, open video courses, and high-quality online courses, researchers have begun to focus on learning engagement in online learning. Online learning engagement has been given a new meaning due to its own characteristics, and it takes on a new tendency: (1) capturing online interactive activities and behavior, and (2) extending to all kinds of online learning experiences [23]. In fact, it is difficult to measure online learning engagement for several following reasons: (1) In the process of online learning, it is hard for teachers to observe students' direct performance because the teachers and students are in separate locations; (2) Students' learning situations are widely varied as independent learning, collaborative learning, offline learning, and learning outcomes make it nearly impossible to describe the learning situation of all students with a single parameter; and (3) In addition to students' online learning behavior, it is difficult to accurately measure students' emotional and cognitive dimensions. In regard to point 3, several researchers have used specific data related to student behavior to measure learning engagement, and they have proposed a learning engagement evaluation rubric. For example, Wise et al. (2014) developed the Learners' Online Listening to Speaking (LOLTS) rubric based on the interactive behavior and performance of learners' online learning discussions [24]. The LOLTS scale mainly reflects the learner's learning engagement from the vertical and horizontal dimensions. The horizontal dimensions are listening pattern and speaking pattern, and vertical dimensions are breadth and depth. Their experiments show that the number of times learners participate in discussions and the amount of time learners spend browsing posts cannot comprehensively evaluate the learning engagement of learners. Learning engagement should be measured more in terms of learning depth. Through interviews and structured questionnaires, Kunhee Ha et al. (2015) conducted experiments using an eye-tracking system to observe how students react to the LAD (Learning Analytics Dashboard), with which students can effectively monitor their online behavior patterns in real-time and utilize such information to change their learning patterns and improve performance [25]. As Kunhee's research relied on an eye-tracking system to collect data and from that deduce learner engagement, it did not suggest a usable way to describe authentic learning engagement. The data in these studies used for evaluating learning engagement includes the duration of time a learner is logged onto the platform, the number of times a learner logs in, the number of posts a learner contributes to the forum, and the number of responses a learner leaves in reply to other learners in the forum. These data points, though, do not reveal insights into actual learning behaviors, as they are only concerned with data that describes material and forum access without directly relating to learning activities. Although these studies simultaneously consider independent and collaborative learning, these studies do not aim to analyze the typical behavior of independent learning and thus fail to identify and present the engagement situation of collaborative learning when only analyzing material and forum access data. Therefore, it is difficult to use such data to learn about and understand learning engagement during the process of online learning.

When using Likert scales to analyze learning engagement, the learner usually completes Likert scales or questionnaires at a single time instead of generating continuous data during the learning process. Thus, using an evaluation scale of this sort is not an effective method for reflecting learners' engagement dynamically and synchronously. In such a case, based on the xAPI standard, this paper selects specific behavior data related to independent learning to establish a commonly-used analytical method of learning engagement to record and reflect learners' engagement orderly and dynamically in the process of online independent learning.

In order to measure students' behavioral engagement in writing activities, Liu et al. (2015) describe a learning analytic system called Tracer, which records the intermediate stages of document development and uses that data to measure learners' behavioral engagement and creates three visualizations of behavioral patterns of students writing on a cloud-based application [26]. Researchers proposed two engagement measurement algorithms to explore different ways to calculate learner engagement and engagement intensity. In addition, the aims of visualizations are that learners can easily check their engagement in the writing activity and change their behaviors in time if they are absent-minded. In currently available online courses, a primary method of online learning is video lectures, which are more vivid and situational than static text and pictures [27]. So, watching video lectures has become a main activity for independent online learners in addition to completing concurrent learning actions while watching these videos. Some studies have indicated that online learning materials, namely video lectures, are but one of the primary factors influencing online learning engagement. In this paper, the source of data that reflects the engagement comes from the operant behavior demonstrated in watching video lectures as well as concurrent learning actions that take place during viewing, such as asking questions, taking notes, and marking content. All the relevant data is automatically recorded in the online learning platform. When students watch a video, their engagement can be analyzed from their behaviors. We supposed that if they take notes, make marks, and/or ask questions, they would devote themselves more to the video learning than without these behaviors. So, we identify these behaviors as the feature of engagement. When they take notes, make marks, and/or ask questions, the learner would definitely pause the video, so it is related to video playing options. The amount of time a learner watches the video lecture and the actual total length of the video lecture, to a certain extent, also reflects the engagement in online independent learning.

III. METHODOLOGY

A. REAL-TIME GROWTH ALGORITHM OF ENGAGEMENT IN INDEPENDENT LEARNING

This paper considers three types of data—video playing options, video lecture viewing time, and concurrent learning actions—to develop and propose a real-time growth algorithm of engagement in independent learning. Video playing options refers to actions a learner does while watching a video, including playing, stopping, pausing, and closing a video. Video lecture viewing time includes the amount of time a learner watches the video lecture, the actual total length of the video lecture, and the viewing length for each time. In our research, concurrent learning actions in dependent learning include a learner asking a question, making a mark, and/or taking notes as they watch the lecture. Learners who are willing to take notes or ask questions demonstrate they are either thinking about the content of the video or are encountering some problems about content in the video lecture. Therefore, if it is possible to regard the online independent learning as an integrated process, data on video playing options, video lecture viewing time, and concurrent learning actions constitute three sub-parameters that can reflect learning engagement. On the basis of these parameters, the following function of learning engagement is formed:

$$\mathrm{E_s = f(Pl, R_t, P_a)\ or\ E_v = f(Pl, R_t, P_a)} \tag{1}$$

Function 1 (The Formula of Learning Engagement): Where E_s represents each learner's engagement and E_v represents learners' engagement while watching video lectures. The f() is a functional relationship model, and Pl is the sub-parameter of video playing option; R_t is the sub-parameter of watching length of video lecture; and P_a is the subparameter of concurrent learning action. The validity of E_s (or E_v) is calculated by weighted summation, that is, $E_s = a^*P_I + b^*R_t/3 + c^*P_a$, where the condition $a + b + c = 1$, and values for a, b, and c are decided by a team teaching online courses or the manager of an online learning program. In the example in this paper, the experts and teaching team give the weight of P_I, R_t, and P_a and they are calculated as $a = 0.4$, $b = 0.4$, and $c = 0.2$ by using the Delphi rule.

All of the variables in equation 1 are defined, respectively, as follows:

$$P_l = \{pli, j, i = 1, 2, \ldots, n; j = 1, 2, \ldots, 5\}$$

In this definition, $n = N$ represents the total number of videos or $n = M$ means the total number of learners. In this data matrix, the variable in the column is video playing options, which is defined as the five specific behaviors of playing, pausing, stopping, forwarding, and closing. When it comes to making real-time observations and calculating every learner's learning engagement, the column direction of the matrix will record the times of all learning behaviors in the process of watching certain videos. In terms of total engagement of the learners in watching each video, the row direction of the matrix will record the corresponding sum data of five behaviors for each teaching video. The specific data are shown as below.

$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 1 & 3 & 1 & 0 & 2 \\ 2 & 1 & 1 & 2 & 1 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ Pl_{n,1} & Pl_{n,2} & Pl_{n,3} & Pl_{n,4} & Pl_{n,5} \end{bmatrix}$$

Function 2 (The Matrix of Learner's Engagement in Watching Each Video Lecture): In this two-dimensional data matrix, the determinant of it can stand for the data on video lecture play actions when a specific learner watches all video lectures (namely n = N) or it can stand for the data on video lecture play when a specific video lecture is viewed by many learners (namely n = M). Additionally, the numerical value of each column represents one type of video playing option whose number can update constantly with an increase in the learners' viewing time. Thus, it is possible to realize the purpose of dynamically tracing the learners' video play behavior. By calculating the data with the corresponding formula (2), the Pl of engagement in the behavior data matrix can be calculated. It is convenient for students and teachers to know about the learning situation using this real-time formative parameter's values.

$$\mathrm{P}_l = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left(\frac{Act_{ij}}{\sum_{i=1}^{n}\Sigma_{j=1}^{5}Act_{ij}}\right). \tag{2}$$

Similarly to Pl, R_t = {rt i, j, i = 1, 2, …, n; j = 1, 2}. In this definition, n = N or n = M, which represents the total number of videos or the total number of learners. The parameter R_t is determined by two factors: the actual length of time spent watching the video lecture (namely T_a) and the original full length of the video lecture (namely T_o). As shown in formula (3):

$$\begin{cases} R_t = \frac{1}{n}\sum\nolimits_{i=1}^{n}\frac{T_{ai}}{\tau_o} \\ T_a = \sum\nolimits_{j=1}^{k} t_k. \end{cases} \tag{3}$$

where k represents the number of times a learner watched each video, and t_k represents the viewing duration for each time.

P_a = {pa i, j, i = 1, 2, . . ., n; j = 1, 2}. In this definition, n = N or n = M, which represents the total number of videos or the total number of learners. The parameter P_a is also determined by the quantity of various concurrent learning behaviors and the total number of concurrent behaviors, as shown in formula (4).

$$P_a = \frac{1}{n}\sum\nolimits_{i=1}^{n}\left(\frac{Act_{ij}}{\sum_{i=1}^{n}\sum_{j=1}^{2}Act_{ij}}\right) \tag{4}$$

Many studies of online learning behavioral engagement typically use evidence collected by human observers with scales or questionnaires to measure learning behaviors such as positive body language, consistent focus, and verbal participation [28]. This study attempted to automatically analyze data of online learning processes while a learner watched a video lecture. Considering the video playing options, video lecture viewing times, and concurrent learning behaviors recorded during the process of viewing videos, the normalized numerical values of learning engagement about videos with different lengths of time can be calculated using formulas (1) to (4). During the learning process, E_s and E_v can be calculated automatically during the watching of a video, the occurrence data of concurrent behaviors, and the duration of the viewing time in the platform. The data for learning engagement, which are formed at any time, not only can help learners better understand their learning situation and manage their learning, but it also can provide data for teachers to observe, support, and manage the learning process in online courses.

B. DATA PROCESS AND ANALYSIS OF ENGAGEMENT PARAMETER

In this study, data on 14,000 adult learners who are teachers was collected from an online training project, which aimed to promote teachers' information technology application capability in Guangdong Province, China. The population in the study are the teachers coming from K-12 school in five districts of Zhaoqing City, Guangdong. 58% teachers teach in primary school and other teach in middle school and high school. We only got their learning data captured by learning platform and couldn't get their demographic data and analyze them. On the other hand, based on the research on 3437 young language learners of English, Spanish, German, Italian and French, Suzanne Graham found that there is little relationship with population characteristics and learning performance (2019) [29]. So, Population differences could be ignored here. The learning materials provided by the project mainly included videos of lecture series, teaching cases, and tutorials of application software according to the curriculum standards of training. Learners could play, stop, pause, and close the video lecture through the video player, which was presented in a web browser. The training platform built functions like making notes and asking questions into the video, so these activities could be performed as the video played. At the same time, the training platform could automatically capture users' operations and record data on them, for example, when the teachers played a video in the course, the platform could capture how long the teachers watched the video for each time and how many times they watched it, and so on. Similarly, related concurrent learning behavior could also be recorded by the training platform.

C. METADATA AND DATA CLEANING

The data of online learning behaviors captured by the training platform totaled more than 1 million records, involving 14,000 adult learners. Before watching and learning a new video lecture, every learner had to complete and master the

TABLE 2. Platform data sheet.

Index	Learner_ID	Video_ID	Action	Length	Addtime

content in the present video. However, they were free to choose the order in which they watched videos from the list. In the learning process, the online learning platform automatically records data on teachers' online learning behaviors, and each record includes the Index, Learner_ ID, Video_ ID, Action, Length (namely the time length of a video) and Addtime (as shown in Table 2).

In the process of learning on the platform, the video will play automatically as long as the learners click a topic in the list and open the video player. At this point, the system will automatically record the playing time and this action as 0. Once a learner pauses the video lecture after it plays, the training platform will record the time and this action as 1. Thus, every time the learner pauses a video lecture, the system then will add 1 to Action and time in Addtime. When the learner continues to play the video, the system will automatically record a playing time and then set Action to 0. After a user pauses the video and does not resume playing for more than 30 minutes, the system will automatically stop playing the video. At that point, the system will not record data of playing time and will add 2 to Action.

In the learning process, concurrent learning actions mainly refers to taking notes and asking questions. Learners who are willing to take notes or ask questions demonstrate they are either thinking about the content of the video or are encountering some problems while watching the video lecture. When these two actions occur, the system will mark the attribute value, respectively, as 7 or 8 to Action. The data of time corresponding to action and videos also will be recorded automatically by the platform. Seen from the data recorded by the platform, taking notes and asking questions both occur after the learners finish watching the video lectures.

According to online learning activities based on the xAPI standard, video playing options include three types of activities: playing, pausing, and stopping. The system will mark the attribute values as 0, 1, and 2, respectively, and will automatically record the time when they occur. In addition, length of video lecture viewing time can be obtained from the calculation of data tracking playing, pausing, and stopping times. In other words, the operator selects all of the records of the same learner_ ID or same video_ ID and then uses the Addtime value, which corresponds to an action whose value is 2. After suspending the video for more than 30 minutes, the platform will end play and mark the action as 1. Then, the platform will subtract the Addtime of the action whose value is 0 to complete the equation, which includes the watching time length of a learner, namely as Length $=$ Addtime $|$Action $= 2$ or Action $= 1$—Addtime$|$Action $= 0$.

Before calculating and analyzing the data, first it is necessary to preprocess more than 1 million data records in order to remove abnormal data and recorders, which are collected from the online learning platform. For example, due to slow Internet connection, videos' suspension can occur frequently while watching, which includes automatic and manual suspension from the users. When the video suspends, the system automatically records a data of action, and namely the action is 0. Additionally, without any video playing options, the system can directly record the concurrent learning action, such as taking notes and asking questions, but this data is not related to a certain video and will not be used to analyze the learning engagement of a video lecture. After data cleaning, 466,000 usable data entries were obtained, and these entries recorded the viewing behavior for 134,000 videos. Second, as the time length of video lectures varied from 1 minute to 99 minutes, data about those videos within 36 minutes were selected as the test data for convenience. Finally, data matrix of online learning behavior was calculated and analyzed, centering on the three quantitative aspects, i.e., video playing options, video lecture viewing time, and concurrent learning actions.

IV. DATE PROCESSING AND ANALYSIS

A. GENERAL CHARACTERISTICS OF ENGAGEMENT IN ONLINE INDEPENDENT LEARNING

Considering the video playing options, video lecture viewing times, and concurrent learning behaviors recorded during the process of viewing videos, the normalized numerical values of learning engagement about videos with different lengths of time can be calculated using formulas (1) to (4). Detailed data is given in Table 3, and Figure 1 is a visual diagram showing the early changes.

As shown in Figure 2, online learners' engagement in independent learning generally is at the medium level, which ranges from 0.5 to 0.7, distributed around 0.6. The data for learning engagement indicated that if learners watched the shorter teaching videos, such as those lasting from 1 to 3 minutes or 4 to 6 minutes, the independent learning's engagement is nearly or more than 0.7 and is higher than engagement for longer videos. Learners' learning engagement gradually lessens when the video lasts longer. When learners watched videos that are 16 to 18 minutes long,

TABLE 3. Detailed data of normalized learning engagement data of different videos with different lengths of time.

Length of video(min)	1-3	4-6	7-9	10-12	13-15	16-18	19-21	22-24	25-27	28-30	31-34	34-36
Learning engagement	0.72	0.69	0.64	0.62	0.54	0.66	0.65	0.57	0.55	0.52	0.62	0.63

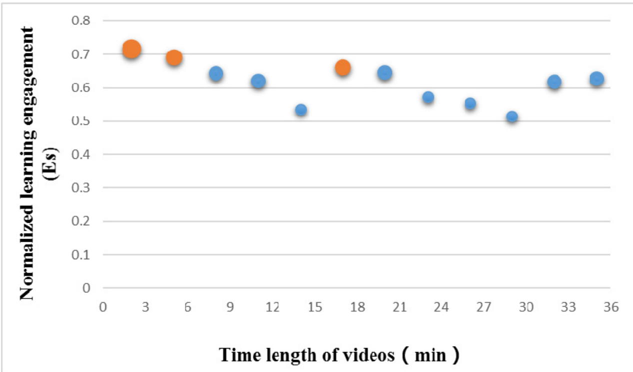

FIGURE 1. Normalized learning engagement of videos with different time lengths.

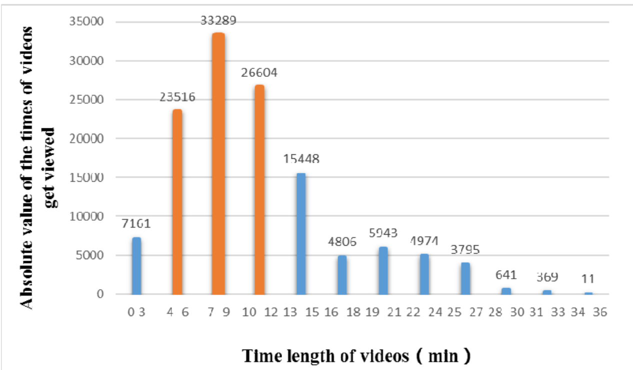

FIGURE 2. Statistical distribution of viewing times for videos of varying lengths.

learning engagement rises again and is more than 0.65. This tendency to increase fails to continue later, and if the video lasted about 30 minutes, a third fluctuation occurs. What this data shows is that most learners can concentrate very well on watching video lectures that are within 6 minutes long. Conversely, learners cannot completely focus on a video's content if the video lectures are longer than 6 minutes. To deal with the distraction and to improve learning engagement, learners can adjust their learning by suspending the videos, replaying the videos, and taking notes. Therefore, when teaching teams design and produce video lectures, it is necessary to put the key points and main content in the first 6 minutes of the video. If a video lecture is longer than 20 minutes, taking full advantage of the segment during 16 to 18 minutes will enhance learning engagement. However, in this segment, video cannot show off a teacher's presentation, explanation, and interpretation, but it can encourage learners' thoughts and promote their interaction with the video contents by asking them questions, using text labels or special effects to emphasize the key points, and frequently changing scenes to hold learners' attention.

B. LEARNING ENGAGEMENT REFLECTED IN VIDEO PLAYING OPTIONS AND CHARACTERISTIC ANALYSIS

This test involved 1,630 video lectures of varying lengths, ranging from 1 to 36 minutes, as shown in Figure 2. As learners were free to choose a fixed number of video lectures to watch, theoretically, the number of all videos watched is random. The data on video lecture viewing times was calculated and classified based on time length. The statistical graph showing the number of times videos of different lengths were viewed can be seen in Figure 3. Data shows most of the video lectures that were watched are videos that last 7 to 9 minutes long. Two main reasons account for this situation. First, there are many 7–9 minute-long videos in the training courses. Second, learners appear to be more willing to watch videos shorter than 15 minutes. Thus, the longer the video is, the less learners wanted to watch it. The video lectures longer than 27 minutes were hardly watched.

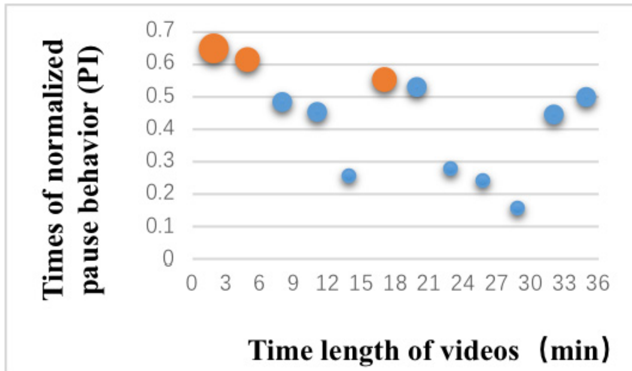

FIGURE 3. Statistical distribution for frequency of normalized video pause behavior.

In online independent learning, learners click the button to pause videos, which means that learners need either to have a break or a chance to perform a certain concurrent behavior, such as thinking, asking questions, or doing online exercises. Here, frequent pauses that are caused by network problems are not considered in this study. Figure 4 shows the distribution of the normalized number of pauses that occur in video viewing during online independent learning. Data in Figure 3 indicates that many pauses occur during video lectures that are 1 to 3 minutes long. Because these videos are concise and clear, learners can understand and grasp the contents rapidly, and thus they may manually pause or stop the videos without watching the entire video. While viewing these kinds of videos, learning engagement is relatively high, as shown in Figure 2. The normalized data of video pausing decreases in the first 15 minutes and learners normally watched the whole video in a single go. Furthermore, similar to general learning engagement, the subparameter of engagement concerning pausing video, namely Pl, increases for videos between 16 to 19 minutes and 31 to 35 minutes. This finding means that learners will pause a video if it is longer and try to adjust their learning situation. Although microlectures have obvious advantages in adult learning, due to adults' learning experiences and better self-regulation capabilities, they can watch the videos in segments based on their own needs and also may adjust their learning strategies to improve their engagement.

C. LEARNING ENGAGEMENT REFLECTED IN WATCHING TIME LENGTH OF VIDEO LECTURE

This paper uses the variable of valid time length to represent learning engagement, which is reflected by the length of time it took to watch various video materials. Figure 4 provides the valid time parameter diagram of the videos with different lengths of time. As shown in Figure 5, videos that last 1 to 3 minutes have high valid time. Compared with other videos, videos lasting 1 to 3 minutes long have longer valid watching time length and learners engage more when they watched these videos. They watched videos multiple times or paused the videos, which shows that learners put more effort into learning the content of these videos. The data shows valid times of watching videos that last 28 to 30 minutes long rise again to more than 2.85 and is the same as videos lasting 9 to 12 minutes. This trend illustrates that learners try to help themselves concentrate on their learning by interacting with the learning content, showing they are aware of their learning situation and know how to adjust learning pace, how to exercise self-control, and how to use learning strategies to focus on learning. Therefore, learning engagement is relatively high.

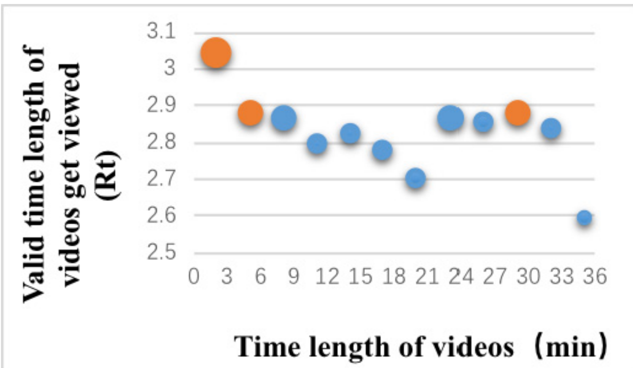

FIGURE 4. Valid time parameter diagram of videos with different time length.

D. LEARNING ENGAGEMENT REFLECTED IN CONCURRENT LEARNING ACTION

Concurrent learning actions refer to the subsequent actions, such as asking questions and taking notes, that occur while watching videos. In watching the video lectures or following the viewing, if learners can ask questions or take notes about the learning content in online learning platform, the learners' engagement will be higher than that of learners who only watch the videos without opinions, questions, nor thoughts. On the other hand, video learning materials that can trigger concurrent learning behaviors also can better facilitate learners' deep thinking. Figure 5 shows the normalized parameter of concurrent actions performed while watching videos of varying lengths. As data in this figure shows, after normalization, the concurrent actions for videos lasting between 22 to 30 minutes are more than those of other video lengths. At the same time, the component parameter value, namely P_a,

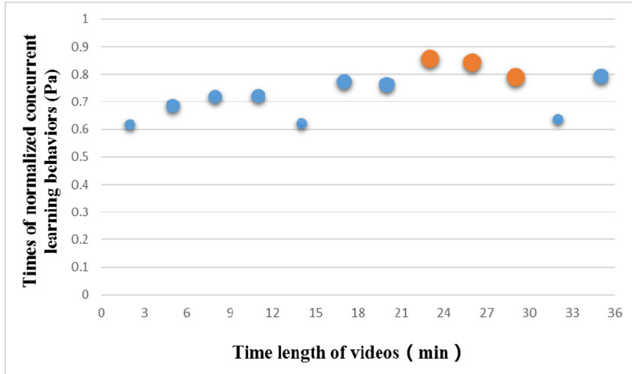

FIGURE 5. Normalized parameter diagram of concurrent behaviors with different video lengths.

is comparatively high. Obviously, concurrent learning actions are more likely to occur while watching the longer teaching videos. Instead, concurrent learning action occurs much less frequently if learners watch the shorter teaching videos. Whether or not the characteristics that are identified by component P_a are related to the opinion that fragmented learning cannot easily trigger learners' deep thoughts, higher order thinking, and in-depth learning, it is necessary to conduct additional studies based on more learning behavior data. In regard to shorter video lectures, however, due to their content being short and brief, learners can quickly understand the video content after watching. Thus, certain concurrent learning actions, like taking notes and asking questions, occur noticeably less frequently in the process of watching and, therefore, they cannot help learners arouse deep thinking because of the concise contents. Evidently, to online learning, video lectures of varying lengths have their own advantages and each video can also provide better learning engagement, not merely videos with short length or concise contents. Consequently, it needs further analyses and studies performed to discuss what kind of teaching content is better for microlectures and what kind of content can be presented more clearly with the aid of a 20-minute video.

V. CONCLUSION AND DISCUSSION

Based on the above research on suggesting and applying real-time arithmetic to explore the engagement in independent online learning, this paper reached the following conclusions:

1. By analyzing learning engagement based on online learning data, making good use of common learning data can help support an effective understanding and monitoring of online learning engagement.

When this analysis method is applied to an online learning platform, it can clearly promote a precise description of the dynamic state of individual engagement in independent learning. This method also can help learners better understand and adjust their learning accurately, and can help teachers effectively monitor and provide learning support to their students. In addition, learning engagement data in different stages can reflect both universal laws of learners' independent learning and that of the individual. The analysis and application of the data of learning engagement in different stages will be a focus in a follow-up study.

2. Analysis of the test results of more than 1 million records shows that the rules and characteristics of online independent learning that are difficult to observe directly can be identified through the parameters of engagement. With the help of these rules and characteristics, learners can better understand and adjust their learning. In the meanwhile, teaching teams can refine curriculum content and rearrange learning requirements, and course developers can revise the design and production of learning materials.

As revealed in the engagement data, learners demonstrate high learning engagement for 7 to 9 minute-long video lectures during online independent learning. After that point, learning engagement then decreases and sometimes fluctuates. Therefore, more emphasis should be placed on designing and developing learning content that fits within a 7 to 9 minute video lecture. Meanwhile, the actual amount of time which the learner spends on watching shorter videos is longer, and these videos may be watched multiple times. In such a case, learners' learning engagement is high. As a result, short videos may be a better way to learn key and difficult points. Adult learners, by contrast, are able to divide learning into several parts while watching longer video lectures, and at the same time, certain concurrent actions, like taking notes and asking questions, will occur.

3. The analysis method of independent learning engagement not only can be used to dynamically analyze learning engagement, but it also can be used to calculate and analyze learning engagement in different stages of video learning materials, which can be useful for identifying the effectiveness and feasibility of learning these materials.

It is helpful to provide data support and basis for designing, developing, adjusting, and improving video learning materials so as to avoid the blindness and subjectivity in design and development of learning materials.

REFERENCES

[1] D. Gao, "A coolheaded response to hot MOOCs: Reflection on the six problems of MOOCs," *J. Distance Educ.*, vol. 2, pp. 39–47, Feb. 2014.

[2] X. L. Dai, K. Guo, and L. Liu, "Empirical research on factors affecting the MOOC learners satisfaction-based structural equation Anal. MOOC learners universities China," *J. Mod. Distance Educ.*, pp. 17–23, Feb. 2017.

[3] X. Fang, X. P. Cui, and G. X. Yang, "Research on the satisfaction of MOOC learning support service based on structural equation model," *J. Open Edu. Res.*, pp. 76–85, May 2016.

[4] Q. H. Zhen, C. Yu, and L. Chen, "A survey of MOOCs instructional interaction in China on MOOC students perspective," *J. China Educ. Technol.*, pp. 77–85, Jun. 2016.

[5] U. G. Pulse. (2015). *Big Data For Development: Opportunities & Challenges*. [online]. Available: http://www.unglobalpulse.org/projects/BigDataforDevelopment

[6] C. J. Bonk, K. Zhang, *Empowering Online Learning: 100+ Activities for Reading, Reflecting, Displaying, and Doing*. San Francisco, CA, USA: Jossey-Bass Publishers, 2008.

[7] Y. W. Tang, T. Zhao, and W. Wang, "XAPI-A new generation of learning technical specification to lead a new standard for smart education," *J. Mod. Educ. Technol.*, pp. 107–113, Jan. 2015.

[8] H. G. Fang, Y. Y. Hu, and Z. Du, "Research on LMS network system architecture based on xAPI learning record," *J. China Educ. Technol.*, pp. 65–69, Feb. 2015.

[9] P. C. Candy, *Self-Direction for Lifelong Learning: A Comprehensive Guide to Theory and Practice*, San Francisco, CA, USA: Jossey-Bass, 1991.

[10] R. W. Tyler, "Evaluation: A challenge and opportunity to progressive education," *Educ. Rec.*, pp. 31–121, Jan. 1935.

[11] A. W. Astin, "Student involvement: A developmental theory for higher education," *J. College Student Develop.*, vol. 40, no. 5, pp. 297–308, Apr. 1984.

[12] J. A. Fredricks, P. C. Blumenfeld, and A. H. Paris, "School engagement: Potential of the concept, state of the evidence," *Rev. Educ. Res.*, vol. 74, no. 1, pp. 59–109, Jan. 2004.

[13] S. L. Christenson, A. L. Reschly, and C. Wylie, *Handbook of Research on Student Engagement*. New York, NY, USA: Springer, 2012, pp. 36–42.

[14] M. D. Dixson, "Creating effective student engagement in online courses: What do students find engaging?" *J. Scholarship Teach. Learn.*, vol. 10, no. 2, pp. 1–13, 2010.

[15] J. Chih-Yuan Sun and R. Rueda, "Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education," *Brit. J. Educ. Technol.*, vol. 43, no. 2, pp. 191–204, 2012.

[16] P. E. Kahn, "Theorising student engagement in higher education," *Brit. Educ. Res. J.*, vol. 40, no. 6, pp. 1005–1018, 2014.

[17] V. Trowler, *Student Engagement, Literature Review*. York, U.K.: Higher Education Academy, 2010.

[18] *Doing More for Learning: Enhancing Engagement and Outcomes*, Australas. Student Engagement Rep., Austral. Council Educ. Res., Camberwell, U.K., 2010.

[19] P. Wang, "Research on model of student engagement in online learning," *EURASIA J. Math. Sci. Technol. Educ.*, vol. 13, pp. 2869–2882, Jul. 2017.

[20] G. D. Kuh, "Assessing what really matters to student learning inside the national survey of student engagement," *Change Mag. Higher Learn.*, vol. 33, pp. 10–17, Mar. 2001.

[21] W. Shu, "The impact on student learning of student engagement in research universities—Based on 'NSSE-China' 2009 data analysis," *Tsinghua J. Educ.*, vol. 4, pp. 24–32, Apr. 2011.

[22] S. Li and C. Yu, "Development and implementation of distance student engagement scale," *J. Open Edu. Res.*, pp. 62–70, Jun. 2015.

[23] R. Yin and H. Y. Xu, "The progress and prospects of online learning engagement research in foreign countries," *J. Open Educ. Res.*, vol. 22, pp. 89–97, Mar. 2016.

[24] A. F. Wise, S. N. Hausknecht, and Y. Zhao, "Attending to others' posts in asynchronous discussions: Learners' online 'listening' and its relationship to speaking," *Int. J. Comput.-Supported Collaborative Learn.*, vol. 9, no. 2, pp. 185–209, 2014.

[25] K. Ha, I.-H. Jo, and S. Lim, "Tracking students' eye-movements on visual dashboard presenting their online learning behavior patterns," in *Emerging Issues in Smart Learning*. Berlin, Germany: Springer, 2015, pp. 371–376.

[26] M. Liu, R. A. Calvo, A. Pardo, and A. Martin, "Measuring and visualizing students' behavioral engagement in writing activities," *IEEE Trans. Learn. Technol.*, vol. 8, no. 2, pp. 215–224, Apr./Jun. 2015.

[27] J. Wang, Y. H. Hao, and J. L. Lu, "Experimental investigation on how instructional videos present affect the result of independent learning," *J. Edu. Res.*, pp. 93–99, Mar. 2014.

[28] XXXXX.

[29] L. Courtney and S. Graham, "'It's like having a test but in a fun way': Young learners' perceptions of a digital game-based assessment of early language learning," in *Language Teaching for Young Learners*. 2019.

SU MU received the B.S., M.S., and Ph.D. degrees in educational technology from Central China Normal University, South China Normal University, in 1994, 1997, and 2007, respectively. She is currently a Professor with South China Normal University, and a Doctoral Supervisor with the School of Information Technology in Education. She is currently serving as a Reviewer for some journals, such as *Open Education Research*, *E-education Research*, and *Modern Distance Education Research*. Her current research focuses on distance education, learning analysis, and learning and online learning.

SHAOMING CHAI received the Ph.D. degree in educational technology from South China Normal University, in 2009. He is currently an Associate Professor with South China Normal University. His research focuses on computer in education, computer supported collaborative learning, online learning, and knowledge building. He worked with Carl Bereiter and Marlene Scardamalia, the founders of Knowledge Building, University of Toronto, as a Visiting Scholar, from 2016 to 2017. He is currently serving as a Reviewer for some journals, such as, *Journal of Distance Education*.

HONGJIANG WANG received the Ph.D. degree in communication and information systems from the South China University of Technology. He is currently an Associate Researcher with the School of Information Technology in Education, South China Normal University. He has published some national invention patents and journal papers, two papers on SCI. His research interests include educational big data and online learning theory.

YILIN CHEN is the graduate student of School of Information Technology in Education, South China Normal University. She majors in modern educational technology. Her research mainly focuses on the strategy of learner-centered English teaching supported by the information technology.