Journal of Science Education and Technology (2020) 29:519–535
https://doi.org/10.1007/s10956-020-09833-9

Mining Effective Learning Behaviors in a Web-Based Inquiry Science Environment

Chih-Ming Chen[1] · Wen-Fang Wang[1]

Published online: 12 May 2020

Abstract
Analyzing learners' learning behaviors helps teachers understand how learning behaviors of learners influence learning performance. To determine which learning behaviors influence learners' science-based inquiry learning performance, this work develops an xAPI (Experience Application Programming Interface)-based learning record store module embedded in a Collaborative Web-based Inquiry Science Environment (CWISE) to record detailed data about students' learning processes. This work discusses whether the significant correlation and cause-effect relationship among science inquiry competence, learning time, and learning performance exist, and examines whether remarkable shifts and differences in the learning behaviors of learners with different learning performances and inquiry competences exist by using sequential pattern mining and lag sequential analysis. The results demonstrate that inquire ability, total learning time in the designed inquiry learning course, and learning time in an inquiry buoyancy simulation experiment are positively correlated with learning performance and can predict learning performance, and the learning time in the inquiry buoyancy simulation experiment appears to be the most significant predictor. The results of lag sequential analyses indicate that learners with high learning performance and high inquiry competence re-adjust hypotheses after performing an inquiry buoyancy simulation experiment, while learners with low learning performance and low inquiry competence lack this critical inquiry learning behavior. This study presents a systematic analysis method to insight the effective learning behaviors in a web-based inquiry learning environment based on mining students' learning processes, thus providing potential benefits in guiding learners to adjust their learning behaviors and strategies.

Keywords Learning process analysis · Data mining · xAPI · Sequential pattern mining · Lag sequential analysis · Web-based inquiry learning

Introduction

A trend toward e-learning is evident in modern instruction. Particularly, web-based inquiry learning has gradually been incorporated into science learning activities in recent years (Raes and Schellens 2015; Fang et al. 2016). Inquiry-based learning can help students comprehend the nature of science and reasoning itself, to cultivate positive attitudes toward science (Marx et al. 2004). Abdi (2014) indicated that learners who are engaged in web-based inquiry learning outperform learners who use traditional learning methods. In inquiry-based learning, teachers have to guide students to define questions, analyze problems, and solve problems. Teaching based on the Next Generation Science Standards (NGSS) calls for more student-centered learning that enables students to think on their own problem-solving, communicating, and collaborating abilities—in addition to learning important scientific concepts (National Research Council 2012). Fang et al. (2016) thus designed an inquiry course that aims to cultivate science inquiry competence based on the Collaborative Web-based Inquiry Science Environment (CWISE), which is a modification of WISE developed at the University of Berkeley (Linn and Eylon 2011) that provides powerful inquiry learning material design tools and supports web-based learning mechanisms. In addition to translating the used language of WISE into Chinese, CWISE is planning to expand its function to cater for collaborative learning. Nevertheless, teachers commonly cannot accurately determine learners' learning progress and states in a web-based inquiry learning environment due to without face-to-face instruction and monitoring, and therefore cannot provide proper learning feedback

✉ Chih-Ming Chen
chencm@nccu.edu.tw

1 Graduate Institute of Library, Information and Archival Studies, National Chengchi University, No. 64, Section 2, ZhiNan Road, Wenshan District, Taipei City 116, Taiwan, Republic of China

according to learning states of learners during web-based inquiry instruction. The learning performance of learners is thus negatively affected. Traditional assessment data from post-hoc assessment (e.g., traditional tests, quizzes, survey, etc.) do not allow tracking and supporting student learning in a formative manner because they cannot be done until students complete a course of action (Williams et al. 2014), and excessive use of the traditional assessment methods would disrupt student inquiry processes as they cognitively engaged with scientific simulations for knowledge building. Inquiry-based learning requires complex problem solving involving the completion of numerous sub-tasks, collaboration, and cognitive endeavors, and thus the use of fine-grained behavior data is necessary to support student learning in all stages of inquiry-based learning (Watkins 2016). That is, relying on a single data source (e.g., traditional tests, surveys, etc.) to explore potentially implicit learning behaviors prevents obtaining a comprehensive understanding of student on-going progress in detail for timely supports. Obviously, teachers' understanding of learners' science inquiry behaviors in a real-time manner based on a diagnosis of learning processes or an assessment of learning performance is a critical issue.

In such a situation, tracking and recording learners' learning behaviors and states without affecting them would help teachers adjust their teaching strategies and methods to improve the learning performance of learners. Advanced Distributed Learning (ADL) developed a new-generation learning process recording technology, xAPI, which can overcome the limitations of single system recording. It allows the recording of learning processes from various e-learning environments, and the format of learning process data contains the "actor," "verb," and "object" (Tin Can API 2015). xAPI allows teachers to more accurately understand students' learning processes and progress to facilitate the practical development of learning process-based learning performance assessment and learning diagnosis mechanisms (Lim 2015). Since a huge amount of learning process data is collected in inquiry-based learning environments and they have multiple attributes, data mining technologies such as sequential pattern mining, classification, and clustering can be used to explore implicit and meaningful information and patterns from a huge amount of learning process data. The goals of so doing are to understand students' learning behaviors (Kamber et al. 2012), to predict and evaluate learning performance, and even to identify the key factors that determine learning performance by using automatic data analysis methods so that teachers can adjust their teaching strategies and develop more adaptive teaching modes (Chen and Chen 2009).

In this study, CWISE was used to support science inquiry learning. The xAPI learning record store (LRS) module embedded in CWISE was developed to record students' science inquiry learning process behaviors in detail. To confirm whether the designed science inquiry learning activities on CWISE are valid or not, this study examined whether the difference between learners' pretest and posttest scores significantly existed or not. Moreover, this study also examined whether the correlations and causal relationships among learners' learning performance, inquiry competence, the total learning time in the designed inquiry learning course, and learning time in the inquiry buoyancy simulation experiment significantly existed or not. Finally, to determine which learning behaviors influence learners' science-based inquiry learning performance, sequential pattern mining and lag sequential analysis were used to elucidate remarkable differences in the inquiry learning processes of learners with different learning performances and inquiry competence. The research results can be used as a useful reference for designing web-based inquiry learning courses to improve learning performance.

Literature Review

Science Inquiry Learning

In traditional instruction, teachers and students perform face-to-face teaching and learning activities in classrooms, and students will passively accept teachers' unidirectional instruction if teachers only adopt the basic instruction type, such as lecture method, to teach their students without using higher-level instruction method and interaction with students. In the learning scenario, teacher-student interaction and students' autonomous learning are lacked, and the learning activity is boring. In contrast, science inquiry learning can deepen students' learning with interactive discussions and practical experimentation, which involves the drafting of questions, the planning and execution of experiments, data analysis, and evidence collection (National Research Council 2000). Particularly, many studies indicated the importance of integrating technology and recommended technology-supported inquiry-based learning environments (Edelson et al. 1999; Kim et al. 2007; Hakverdi-Can and Sönmez 2012). The benefits of integrating technology into inquiry learning environments include giving students the opportunity of experiencing scientific modeling, using dynamic simulations, and working with actual scientific data through involvement in a scientific experiment (Hakverdi-Can and Sönmez 2012). Many studies (Backus 2005; Deters 2005; Khan et al. 2011) have indicated that science inquiry learning can facilitate students' understanding of the science concepts, which contributes to the development of scientific process skills and to an increase in their attitudes toward and achievement in science courses, and those studies pointed out that it is the best method in science education. Edelson et al. (1999) identified three benefits of science inquiry learning: acquiring general inquiry competence, acquiring specific investigation skills, and developing an improved understanding of the concepts of science.

Science inquiry learning helps students construct knowledge and cultivate scientific inquiry competence. The cultivation of scientific inquiry competence involves practical experiences, induction, interpretation, and verification, all of which help students understand the inductive process of science problems. Achieving scientific inquiry competence is an essential element of scientific literacy (Arnold et al. 2018). Many studies have discussed the improvement of students' science inquiry competence, indicating that students who learn through science inquiry courses have better scientific knowledge and curiosity about science than those who learn via traditional instruction (Buffler et al. 2001; Cropley and Page 2002). However, relevant research has emphasized the importance of developing science inquiry courses and the evaluation of science inquiry competence (Myers and Burgess 2003; Wu and Hsieh 2006), but few studies have focused on examining the relationship between inquiry learning processes and achievement. Buffler et al. (2001) evaluated students' science competence based on conducting a scientific inquiry experiment. Gobert et al. (2013) found that students' inquiry competence to design a control experiment can be evaluated by using data mining schemes based on log files recording the details of each learner's learning processes. Therefore, analyzing students' science inquiry learning processes in CWISE can help teachers understand students' science inquiry processes, and further to identify the behavior variations among the learning processes of learners with different learning performances and inquiry competence. Accordingly, teachers' teaching contents and steps can be adjusted to improve learning outcomes according to the determined effective learning behaviors, and how teachers should advise their students during performing inquiry learning can also be determined.

Records of Learning Processes and Application to Instruction

Rapid advances in educational technologies enable teachers to use a large amount of fine-grained usage data that represent student learning processes to analyze the learning behaviors or factors affecting students' learning achievement. Learning processes in a web-based learning environment, which are system's operational behaviors generated by learners during the period of achieving a learning goal, can be purposively, accurately, and authentically recorded (Pan and Hawryszkiewycz 2004). Such records can include details like log traces that represent time on task, interaction, or resource use as well as interactions with other learners, etc. Unlike summative assessment using traditional pencil-and-paper tests, formative assessment implemented by using student learning process data based on data mining and machine learning technologies can comprehensively elucidate the factors that affect learning performance (Chen et al. 2019). Thus, learning progress, learning difficulties, and degree of effort can be thoroughly understood. Clearly, the use of records of learning processes in education is increasing (Sparapani et al. 1996). In recent years, most researchers have used original records of learning processes (Arroyo and Woolf 2005) for learners' behavior analysis, and have focused on single e-learning systems and curricula. Research on cross-platform learning process records suffers from the limitations on the gathering of information about learning processes, but ADL sets a new standard for the recording of learning processes, xAPI (Experience Application Programming Interface), which can record learning processes from different learning platforms at any time and place. Using the JavaScript Object Notation (JSON) format, xAPI records learning processes and stores them in the learning record store (LRS), and each recorded learning process is divided into "actor," "verb," and "object," as in "Mike passed Introduction to REST", for example (Tin Can API 2015).

Manso-Vázquez et al. (2015) indicated that using xAPI to develop a mechanism for recording learners' self-regulation strategies and self-discipline can enhance learners' reflection by providing self-regulated learning conditions and improve self-regulated learning ability. Rabelo et al. (2015) proposed an ontology-based xAPI big data learning process analysis architecture for supporting valuable applications, such as learning performance prediction, learning path analysis, and the extraction of learners' learning behavior patterns for instructors and learners. Lim (2015) presented several applications of xAPI in e-learning, including tracking processes of game-based learning and monitoring students' learning processes of reading course materials to make recommendations for promoting learning performance. xAPI records of learning processes can be shared and communicated with different systems using a simple protocol. xAPI is a cross-platform, highly secure, and highly flexible learning process record standard that can be used in the real world. As a result, this standard was used in this study to record learners' learning processes in CWISE, and sequential pattern mining and lag sequential analysis were used to provide feedback to teachers based on the results of mining the learning processes for improving science inquiry learning performance.

Application of Sequential Pattern Mining to Instruction

Sequential pattern mining, proposed by Agrawal and Srikant (1995), can be used to mine symbolic sequential patterns of elements and events that happened enough frequently. For example, consumers' shopping sequences, webpage click streams, program execution sequences, biological and science engineering, and natural and social developments can all be effectively mined (Kamber et al. 2012). For example, Fortenbacher et al. (2013) collected students' learning processes in different learning environments using the standardized ETL (Extraction–Translation–Load) interface and provided learners' visual learning paths to help teachers evaluate students' learning processes based on sequential pattern mining. Li et al. (2015) used a sequential pattern mining algorithm to find the frequent patterns of students' question-solving behaviors to know how most students perceive a question's level of difficulty. The sequential pattern mining algorithm in the sequential pattern mining framework (SPMF) proposed by Fournier-Viger et al. (2008) was used in this study to mine sequential patterns with time constraints. The method combines multi-dimensional data mining, time intervals, and the *K*-means clustering algorithm to mine flexibly sequential patterns. From learners' learning behaviors, Fournier-Viger et al. (2009) mined the time-weighted sequential patterns of learners' completion of tasks, and organized the results in a database, further enhanced the efficacy of the teaching system. Although the time-weighted sequential pattern mining algorithm has been successfully applied to e-learning, it has still not been applied to investigate inquiry learning behaviors. This study thus tried to identify the differences in the science inquiry learning processes of learners with different learning performances and inquiry competence by using sequential pattern mining with a time weight to enable teachers to adjust their courses or teaching strategies.

Application of Lag Sequential Analysis to Instruction

Lag sequential analysis is a method for investigating how chains of behaviors and events are linked over time (Marono et al. 2018), whereas sequential pattern mining is used to analyze sequences of symbols and find subsequences of symbols that appear frequently in a set of sequences (Agrawal and Srikant 1995). Adopting the two methods to mine learners' learning processes aims to help teachers observe learners' learning behaviors that affect learning performance from different perspectives. In the e-learning field, the lag sequential analysis could be applied to analyze the characteristics of behavior shifts and understand the behavior difference among learners based on coding the time sequence of behavior events gathered from learners (Quera and Bakeman 2000; Moran et al. 1992). Applying the lag sequential analysis to analyze learners' behavior sequences could help deeply understand the behavior process and the factors that may affect learning performance so that e-learning strategies and course design can be effectively improved. Bakeman et al. (2005) proposed seven major steps for lag sequential analysis that includes defining basic questions and ideas, developing behavior coding from ideas to questions, recording behavior codes, presenting original data, confirming the reliability of the observer, describing and filtering the original data, and describing and analyzing the behavior diagram.

Applying the lag sequential analysis to explore e-learning behavior sequences is becoming increasingly popular in recent years. Chen and Lin (2014) discovered that a digital library system with good information organization and design can effectively improve learning performance. This finding was made by performing a lag sequential analysis after the behavior processes of learners were encoded while using the digital library system to perform a learning mission. Chiang, Yang, and Hwang's study (Chiang et al. 2014) guided learners who were engaged in inquiry learning activities based on location-based augmented reality, and performed a lag sequential analysis based on the recorded learning behaviors. They found that location-based inquiry learning allows greater interactivity among learners and facilitates the learners' knowledge construction. Yang et al. (2015) encoded learners' learning behaviors based on their pretest and posttest performances in a two-tier assessment system and performed a lag sequential analysis of learning behaviors. They found that two-tier assessment can help learners develop more effective learning methods than previously used tests, thus improving their learning performance. The application of lag sequential analysis to e-learning can elucidate the material learning sequence, providing a reference for instructors in adjusting the material structure. Therefore, lag sequential analysis of behavior sequence of learners with various learning performances and inquiry competence in the inquiry buoyancy simulation experiment was performed in this study to enable instructors to design and adjust curriculums.

Research Methodology

Research Participants

In consideration of the computer literacy of the research subjects and the difficulty of the designed inquiry learning course, a total of 48 Grade 7 students from a junior high school in Taipei City, Taiwan, were randomly recruited as the research subjects. To examine whether learners with high and low learning performance will lead to different learning behaviors or not, the median split design was used to group the learners into high and low learning performance groups due to that Iacobucci et al. (2015) confirmed that median splits are perfectly acceptable to use when independent variables are uncorrelated. Since this study adopted a random approach to select the research participants, the assumption that the learning performance of the research participants follows a normal distribution can be satisfied. In a normal distribution case, the mean is equal to the median. Moreover, the research participants' learning performance that was divided into high and low groups for the learning behavior analysis is regarded as an independent variable in this study. Obviously, the learners who have high learning performance will not be affected by the learners who have low learning performance, vice versa. Therefore, this case satisfies the condition of the median split. Therefore, based on the learners' mean posttest performance after performing an inquiry learning activity, learners with higher performance than the mean were regarded as the learning group with high learning performance, while the others were regarded as the learning group with low learning performance. Moreover, the hypothesis proposal, experiment process, and data analysis in the inquiry buoyancy simulation experiment were evaluated in this study, and the total score mean of the three parts was taken as the basis to discriminate the learners' inquiry competence. The evaluation method of the learner's inquiry competence is described later in the research methodology section. Tables 1 and 2 present the descriptive statistics of the pretests and posttests of learners with various learning performances and inquiry competence.

Experimental Design and Procedure

Single-group pre-experimental research was adopted in this study and a total of 48 Grade 7 students from a junior high school in Taipei City, Taiwan, were invited to participate in the inquiry instruction experiment. The experimental procedures were first explained to the research participants for 10 min. After that, a pretest was performed for 15 min before performing the inquiry learning activity. The 45-min inquiry learning activity on CWISE was then performed, and a 15-min posttest was then conducted. Finally, to investigate the research participants' learning strategies and methods during the inquiry learning course, the five learners with the best and the five with the poorest learning performances among 48 research subjects were interviewed one-on-one. This is because selecting the research subjects who have extremely high and low learning performance as the interviewees can help this study examine the effects of the designed inquiry learning course on learning perception more distinguishingly. Namely, this study adopted a purposeful sampling for the interview. Incorporating interview data aims to delve into individual student experience in a greater detail than could be only with quantitative data. The learning processes of research participants in the designed inquiry learning course on CWISE were recorded by using the xAPI to determine whether or not significant behavioral differences of learners with various learning performances and inquiry competence existed. Namely, the data of each research participant recorded by using the xAPI was used as the unit of learning process analysis. The purpose of the recording learning process is to identify the learning sequence behaviors that affect the inquiry learning performance.

Research Tool

Collaborative Web-Based Inquiry Science Environment with xAPI Learning Process Monitoring Module

Collaborative Web-Based Inquiry Science Environment (CWISE) allows teachers to design inquiry learning activities that are aimed at promoting students' science concepts. It offers several course activity management and editing functions to support teachers to design inquiry learning courses. CWISE can support collaborative learning through a brainstorming discussion board, but the designed inquiry instruction experiment of this study only allowed learners to perform individual learning without using the function of a brainstorming discussion board. Figure 1 shows a webpage that accompanies a CWISE course. In CWISE, science inquiry learning activities are designed to cultivate students' understanding of basic scientific concepts, to increase their curiosity in science, and to increase course interactivity by applying multi-media elements, such as films and animation. The inquiry curriculum design can help students explore and understand scientific questions during the learning process, to obtain a complete grasp of scientific concepts with their teachers' assistance, and to apply these concepts in everyday life.

Figure 2 shows the developed system architecture of CWISE with the xAPI learning process monitoring module for recording learners' learning processes and the sequential pattern mining module for mining learners' learning behaviors. First, the students logged into the platform from the course learning system interface to start the course learning. After they clicked on the course through the course learning system interface, CWISE would inquire about the course database and display the corresponding learning course. At the same time, the learning process data are recorded and transmitted through the xAPI learning process monitoring module to the LRS for data storage. The sequential pattern mining

Table 1 Descriptive statistic results of pretest/posttest results of learners with high/low learning performance

Group	Number of learners	Pretest		Posttest	
		Mean	Standard deviation	Mean	Standard deviation
High learning performance	23	4.70	1.69	5.82	1.07
Low learning performance	25	2.57	1.18	2.94	1.20

Table 2 Descriptive statistics results of pretest/posttest results of learners with high/low inquiry competence

Group	Number of learners	Pretest		Posttest	
		Mean	Standard deviation	Mean	Standard deviation
High inquiry competence	20	4.42	1.58	5.18	1.68
Low inquiry competence	28	3.00	1.72	3.70	1.73

module extracts the learning process data from the LRS to perform sequential pattern mining and lag sequential analysis. The major difference between the learning process database and the LRS process database is that the data in the former are the general learning process data that were originally designed in CWISE, and not the sequential behavior pattern data that comprise "actor," "verb," and "object," which are required for the sequential pattern mining module.

Web-Based Inquiry Learning Course on CWISE

The inquiry course called "Floating Future" was used in this study. Its curriculum design was based on the science and technology standards in the Grade 1–9 curriculum guidelines of Taiwan, and it incorporates unit modules that include explanatory components. "Floating Future" includes three activities, which are called, "Floating Future," "Sinking and Floating Phenomena," and "Buoyancy." Table 3 presents the course node coding of the floating future course for sequential pattern mining.

At the beginning of the course, environmental problems such as rising sea levels, and the idea of Dutch floating homes, are introduced. A video lecture about floating homes is included to increase the variety of course materials. Most of the population of Taiwan resides in cities along the coast, so learners are guided to think of factors that affect the floating of objects. When learners notice sinking and floating phenomena in everyday life, they then are more likely to think about related factors affecting buoyancy. Figure 3 shows the coding scheme of the inquiry buoyancy simulation that is used in sequential pattern mining and lag sequential analysis. In Fig. 3, the hypotheses concerning the factors that affect buoyancy were respectively coded as H1, H2, and H3, the determination of the "size" of the duck was coded as D1, the selection of the "material" of the duck was coded as D2, the selection of the "liquid" of the buoyancy experiment was coded as D3, the "executing" of the buoyancy simulation experiment was coded as D4, and the conclusions concerning the factors that affect buoyancy were respectively coded as E1, E2, E3, E4, E5, and E6.

This inquiry buoyancy simulation experiment is a complete inquiry activity that allows learners to propose hypotheses, perform simulation experiments, and obtain analysis results. The curriculum design incorporates numerous interactive elements that increase the absorption of knowledge in the learning process. In the simulation experiment, which involves a duck, the students can change four variables, which are (1) the material of which the duck is made (brick, wood, ice, Styrofoam, or stainless steel), (2) the size of the duck (large, medium, or small), (3) the structure of the duck (solid or hollow), and (4) the liquid (water, brine, petrol, or mercury). The learners can reason to concepts about buoyancy from the results of the experimental simulation. The learners' inquiry learning competence is

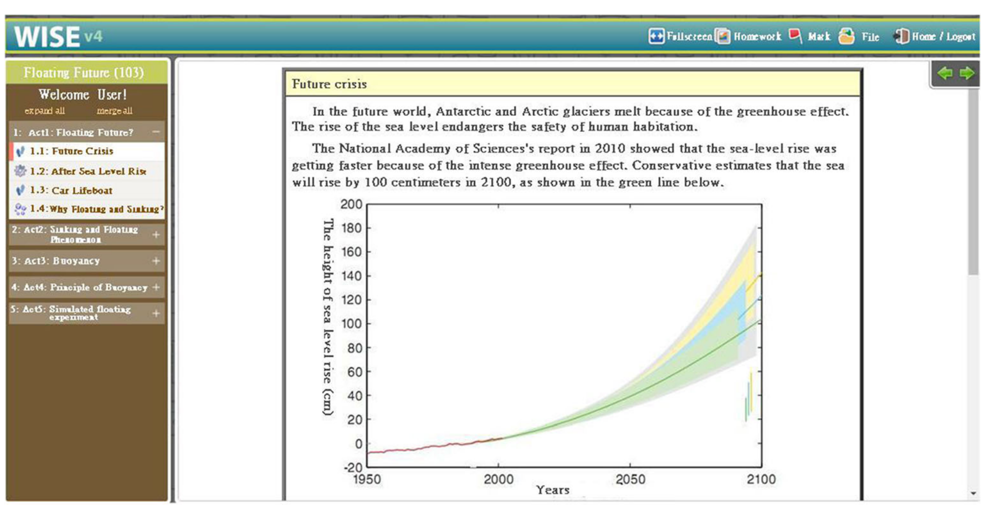


Fig. 1 The webpage of a CWISE course

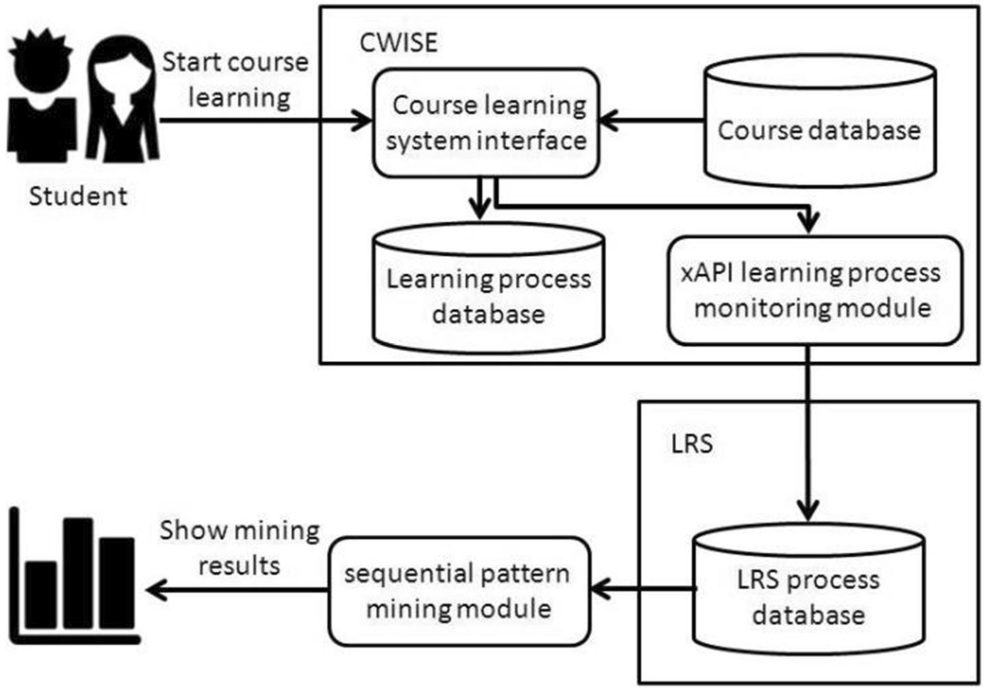


Fig. 2 System architecture of CWISE with the xAPI learning process monitoring module for recording learners' learning processes and the sequential pattern mining module for mining learners' learning behaviors

evaluated by performing the inquiry buoyancy simulation experiment, which comprises three phases; these are the generation of a hypothesis, the execution of the experiment, and the analysis of results. Table 4 presents the standards against which learners' inquiry learning competence is evaluated.

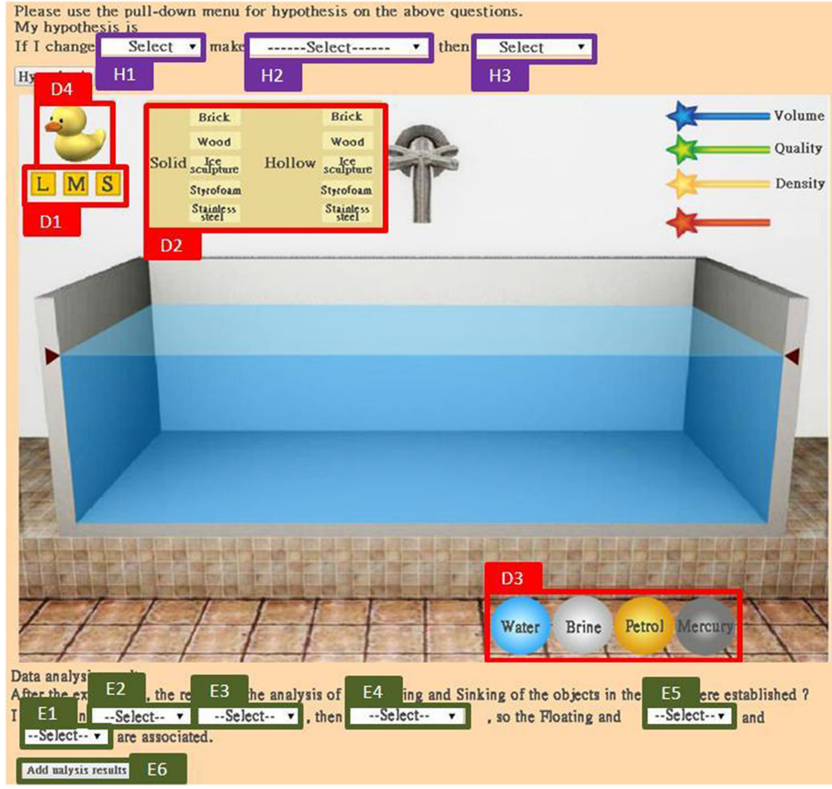
The Gathered Log Data of the Learning Process by xAPI

xAPI is a new generation learning process record standard. "Actor," "verb," and "object" with a corresponding time stamp are the record formats generated from the learners' learning processes recorded through xAPI. Namely,

Table 3 The course node coding of the floating future course

Course node	Topic
1	Floating future?
1.1	Future crisis
1.2	After sea level rise
1.3	Car lifeboat
1.4	Why floating and sinking?
2	The sinking and floating phenomenon
2.1	Prediction: factors in the floating and sinking phenomenon
2.2	Page: factors in the floating and sinking phenomenon
2.3	Simulation: factors in the floating and sinking phenomenon
2.4	Challenge problem: floating and sinking
2.5	Challenge mission
3	Buoyancy
3.1	Page: what is buoyancy?
3.2	Your idea
3.3	Simulation: buoyancy of buoys
3.4	Explanation: what happens?
3.5	Simulation: buoyancy of sinks
3.6	Explanation: what happens?
3.7	Evaluation of buoyancy explanation

Fig. 3 The coding scheme of the buoyancy inquiry simulation experiment



the granularity of the log data of the learning process recorded based on xAPI is composed of time stamp, "actor," "verb," and "object" called a statement. The actor is the executor of behaviors and could be an individual or a group, for example, a learner. The verb is the type of action performed by the actor, for example, answered or logged in. The object interacts with the actor, with activity as the unit, and could be a combination of instruction, experience, or performance with descriptive meanings and figures or tangible objects, such as course. xAPI provides flexible design architecture to allow us to self-define the verb names for the recorded learning process. The learning process data could then be transmitted to the learning record store (LRS) through the data transfer protocol for data storage. The data stored in the LRS could be used for the learning process analysis, and the LRS will be developed to provide the functions of retrieval, visual presentation, and analysis in the future (ADL Net 2015). For example, the format of a statement recorded by xAPI can be {2019/7/25 morning 10:13:28, David, proposed, hypothesis 1}, including the time stamp, "actor," "verb," and "object."

Sequential Pattern Mining Tool

Sequential pattern mining is aimed to mine symbolic sequential patterns of elements and events happened enough frequently (Agrawal and Srikant 1995). The sequential pattern mining framework (SPMF) which is the software package programmed with Java which offers more than 55 data mining algorithms proposed by Fournier-Viger et al. (2008) was used as the research tool for mining sequential patterns with time constraints to analyze inquiry learning processes in this study (Fournier-Viger et al. 2014). The algorithm could perform sequential pattern mining by matching with time weight (SPMF 2015). Table 5 shows the input data format of the used sequential pattern mining algorithm. In the example of the input learning process sequence data of the student with ID S1, time is 0, behavior 1 shows the time weight of 2, and behavior 2 does not have a time weight value; behavior 3 appears when time is 1; behavior 6 appears when time is 2; and behavior 5 shows a time weight of 1 when time is 3; and so on. ID could be coded with the student's class and seat number, which are further transformed into the ID number for distinguishing IDs in the sequential pattern mining. Regarding the learning process sequences, a "verb" recorded with the xAPI format is used after being coded. The italicized values in the brackets represent the weights of the corresponding behaviors, which would be adjusted with different mining demands. During sequential pattern mining, the minimum support has to be set to mine the maximum sequential patterns that are higher than the minimum support.

Lag Sequential Analysis Calculator

Lag sequential analysis is a method for investigating how chains of behaviors and events are linked over time (Marono et al. 2018). To analyze the behavioral transfer of the research

Table 4 Evaluation standards for the inquiry simulation experiment

Phase	Score	Description
Experimental hypothesis	3	Complete and correct hypothesis, accepting hypotheses. - When the size of the duck so that its density is greater than the liquid density, it would sink. Accepting hypothesis. Complete but wrong hypothesis, refusing hypothesis. - When the liquid mass is changed so that it is greater than the mass of the duck, the duck would float. Refusing hypothesis.
	2	Complete and correct hypothesis, refusing hypotheses. - When the size of the duck is changed so that its density is greater than the liquid density, it would sink. Refusing hypothesis. Complete but wrong hypothesis, accepting hypotheses. - When the liquid mass is changed to be greater than the mass of the duck, the duck would float. Accepting hypothesis.
	1	Incomplete hypothesis.
	0	No hypothesis proposed.
Experiment proceeding	3	Experiment operation effectively controls variables and corresponds to the hypothesis. - Assuming that the density of the duck is greater than the liquid density, the duck would sink. The duck material is changed, and the size and liquid are controlled in the experiment.
	2	Experiment operation controls some variables but does not correspond to the hypothesis. - Assuming that the density of the duck is greater than the liquid density, the duck would sink. The duck material and size are changed, and the liquid is controlled in the experiment.
	1	Experiment operation does not correspond to the hypothesis.
	0	No experiment operation.
Result analysis	3	Analysis results are completely correct and correspond to the hypothesis. - Assuming the density of the duck is greater than the liquid density, the duck would sink. The analysis results show that the duck sinks because its density is greater than that of the liquid. The floating and sinking of an object are therefore related to the object density and liquid density.
	2	Some analysis results are correct and correspond to the hypotheses. - Assuming that the density of the duck is greater than the liquid density, the duck would sink. The analysis results show that the duck sinks due to its density being greater than that of the liquid. The floating and sinking phenomenon are related to the object mass and the liquid density. Completely correct analysis results, but not corresponding to the hypothesis. - Assuming that the mass of the duck is greater than the liquid mass, the duck would sink. The analysis results reveal that the duck sinks because its density is greater than that of the liquid. Accordingly, the floating and sinking of an object are related to object density and liquid density.
	1	Some analysis results are correct but not correspond to the hypotheses. - Assuming that the mass of the liquid is greater than the duck mass, the duck would sink. The analysis results reveal that the duck would float because the density of the liquid is greater than that of the duck. Accordingly, the floating and sinking of an object are related to object mass and liquid mass. Analysis results are incorrect but correspond to the hypotheses. - Assuming that the mass of the liquid is greater than the duck mass, the duck would float. The analysis results reveal that the duck would float because the mass of the liquid is greater than that of the duck. Accordingly, the floating and sinking of an object are related to object mass and liquid mass.
	0	Analysis results are incorrect but not correspond to the hypotheses. - Assuming that the density of the duck is greater than the liquid mass, the duck would sink. The analysis results reveal that the duck would sink because the mass of the liquid is greater than that of the duck. Accordingly, the floating and sinking of an object are related to object mass and liquid mass. No any Analysis results are proposed

Table 5 Input data format of sequential pattern mining

ID	Sequences
S1	(*0*, 1(*2*) 2), (*1*, 3), (*2*, 6), (*3*, 5(*1*))
S2	(*0*, 1(*2*) 2), (*1*, 4(*8*)), (*2*, 3), (*3*, 5(*2*) 6 7)
S3	(*0*, 1(*3*) 2), (*1*, 4(*7*)), (*2*, 3), (*3*, 5(*4*))
S4	(*0*, 1(*3*) 2), (*1*, 4(*6*)), (*2*, 5(*5*)), (*3*, 6 7)

The italicized values in the brackets represent the time weight of the corresponding behavior

participants who used the web-based inquiry learning course, the system's operation behaviors of the research participants were encoded as a series of behavior sequence samples with time stamp, according to the course's nodes and the operations of inquiry buoyancy simulation experiment, for lag sequential analysis. To perform the lag sequential analysis, this study employed the lag sequential analysis calculator which is available at https://pulipulichen.github.io/HTML-Lag-Sequential-Analysis/ as the research tool. According to the tool, the calculation is suitable for behavior sequence samples with a non-normal distribution when their probabilities are equal to each other. The number of samples in sequential analyses was calculated by the frequency of the neighboring pairs of events. The zero-order model proposed by Bakeman (1986) was used to calculate the Z score. A Z score above 1.96 indicates that the sequence presents remarkable coding transfer and the research participants with an obvious behavioral transfer in the system's operation could be observed, and a high Z score indicates a larger behavioral transfer compared to a low Z score.

Experimental Results and Analysis

Analysis of the Learning Performance of All Learners

Since the assumption of normality for the t test was satisfied, the paired samples t test thus was used to analyze the difference between pretest and posttest scores. Table 6 presents the results, which indicate a remarkable difference between the learners' pretest and posttest scores ($t = -3.567$, $p = .001 < .05$). The pretest and posttest means were 3.59 and 4.32, respectively. The higher posttest mean reveals that the designed science inquiry learning course "Floating Future" significantly improved learners' learning performance.

Correlation and Regression Analysis of Learning Performance and Several Considered Learning Factors

Correlation analysis of learning performance and inquiry competence, the total learning time in the designed inquiry learning course, and learning time in the inquiry buoyancy simulation experiment were performed. Table 7 shows the results. The analytical results present significantly positive correlations between learning performance and inquiry competence ($r = .351$*), the total learning time in the designed inquiry learning course ($r = .297$*), and learning time in the inquiry buoyancy simulation experiment ($r = .397$**). Therefore, a higher inquiry competence, a longer total learning time in the designed inquiry learning course, and a longer learning time in the inquiry buoyancy simulation experiment are correlated with better learning performance. Regression analysis was performed on the independent variables including inquiry competence, the total learning time in the designed inquiry learning course, and learning time in the inquiry buoyancy simulation experiment with the dependent variable learning performance. Table 8 shows the results, indicating that inquiry competence can predict learning performance (β=0.267, $p = 0.014 < .05$) and the explained variance is 12.3%, total learning time in the designed inquiry learning course can predict learning performance (β=0.001, $p = 0.040 < .05$) and the explained variance is 8.8%, and learning time in the inquiry buoyancy simulation experiment can predict learning performance (β=0.004, $p = 0.005 < .01$) and the explained variance is 15.8%. The results indicate that learning time in the inquiry buoyancy simulation experiment is the factor that most strongly affects learning performance among the three considered learning factors.

Table 7 Correlation analysis of learning performance and several considered learning factors

Comparison item	Inquiry competence	Total learning time in the designed inquiry learning course	Learning time in the inquiry buoyancy simulation experiment
Learning performance	.351*	.297*	.397**
	.014	.040	.005

*indicates p<.05; **indicates p<.01

Sequential Pattern Mining of Science Inquiry Behavior

To understand the sequence of learners' science inquiry behaviors during the "Floating Future" course, sequential patterns with time constraints were mined, as proposed by Fournier-Viger et al. (2008), to analyze the variation among behavior sequences of learners with various learning performances and inquiry competence. However, this study found that conducting sequential pattern mining for learners with different inquiry competence cannot find any frequent sequential patterns under considering large enough minimum support, implying that no meaningful sequential patterns for learners with different (high and low) levels of inquiry competence could be found in this study. Therefore, this study only provides the results of sequential pattern mining analysis of learners with different learning performances.

Table 6 Paired samples t test of pretest and posttest

	Number of learners	Mean	Standard deviation	t	Significance
Pretest	48	3.59	1.79	− 3.6**	.001
Posttest	48	4.32	1.84		

**Indicates $p < .01$

Table 8 Linear regression analysis results between inquiry competence, total learning time, inquiry simulation experiment, and activity learning time and learning performance

Model summary			ANOVA		Unstandardized coefficients		
Selected variable	*R*	R^2	*F*	Sig.	*β*	*t*	Sig.
Inquiry competence	0.351	0.123	6.467	0.014	0.267	2.543	0.014*
Total learning time in the designed inquiry learning course	0.297	0.088	4.461	0.040	0.001	2.112	0.040*
Learning time in the inquiry buoyancy simulation experiment	0.397	0.158	8.604	0.005	0.004	2.933	0.005**

*indicates p<.05; **indicates p<.01

Overall Course Node Sequential Pattern Mining Analysis of Learners with Different Learning Performances

Based on the posttest mean, the learners were divided into the high learning performance group and the low learning performance group. The maximum sequential patterns mining of the course node behaviors was thus performed. Tables 9 and 10 present the results. This study found that although the learners who participated in the inquiry learning experiment could do inquiry tasks in a different order, the analytical results based on sequential pattern mining with the pre-determined minimum support which is a parameter to determine the least threshold of mining sequential patterns indicate that both groups of learners significantly intended to follow the designed learning sequence. Interestingly, learners with low learning performance exhibited higher course node sequence continuity than those with high learning performance. This finding will be discussed in the discussion section.

Table 9 Sequential pattern mining results of course nodes for learners with different learning performances (1)

High learning performance group		Low learning performance group	
Minimum support: 0.789		Minimum support: 0.689	
Sequence	Action	Sequence	Action
<1>	1.1	<1>	1.1
<2>	1.2	<2>	1.2
<3>	1.3	<3>	1.3
<4>	1.4	<4>	1.4
		<5>	2.1

"Minimum support" means the least threshold applied to find all frequent learning behaviors

"Sequence" means the discovered learning behavior sequence under the pre-determined minimum support

"Action" means the learning behavior of the corresponding course node

The number inside < > means the order of the encoding learning behavior sequence

Sequential Pattern Mining Analysis of Results of Inquiry Buoyancy Simulation Experiment for Learners with Different Learning Performances

Sequential pattern mining of the behaviors of learners with different learning performances in the inquiry buoyancy simulation experiment was performed. Tables 11 and 12 show the results. The analytical results indicate that the two groups are similar. From Table 11, almost 50% of learners in both groups clicked on "size (D1)," then "material (D2)," and then "executing (D4)." In Table 12, almost 50% of learners clicked on "size (D1)," then "size (D1)," and then "material (D2)." Learners with high learning performance acted continually in operating the "size," "material," and "liquid" of the water pool and the "executing" of the buoyancy simulation experiment, while those with low learning performance did not. This finding will also be discussed in the "Discussion" section.

Table 10 Sequential pattern mining results of course nodes for learners with different learning performances (2)

High learning performance group		Low learning performance group	
Minimum support: 0.6		Minimum support: 0.6	
Sequence	Action	Sequence	Action
<1>	2.5	<1>	2.5
<2>	3.1	<2>	3.1
<3>	3.2	<3>	3.2
<4>	3.3	<4>	3.3
<5>	3.4	<5>	3.4
<6>	3.5	<6>	3.5
		<7>	3.6

"Minimum support" means the least threshold applied to find all frequent learning behaviors

"Sequence" means the discovered learning behavior sequence under the pre-determined minimum support

"Action" means the learning behavior of the corresponding course node

The number inside < > means the order of the encoding learning behavior sequence

Table 11 Same sequential pattern mining result of the inquiry simulation experiment activity for learners with different learning performances (1)

High learning performance group		Low learning performance group	
Minimum support: 0.473		Minimum support: 0.518	
Sequence	Action	Sequence	Action
<1>	D1	<1>	D1
<2>	D2	<2>	D2
<3>	D4	<3>	D4

"Minimum support" means the least threshold applied to find all frequent learning behaviors

"Sequence" means the discovered learning behavior sequence under the pre-determined minimum support

"Action" means the learning behavior of the corresponding course node

The number inside < > means the order of the encoding learning behavior sequence

Lag Sequential Analysis

Lag Sequential Analysis of Behaviors of Learners with Different Learning Performances in Inquiry Buoyancy Simulation Experiment

A lag sequential analysis of behaviors in the inquiry simulation experiment for learners with different learning performances was performed. Figure 4 shows the results. The significant behavioral shifts of learners with high learning performance were "propose hypothesis—perform experiment" ($Z = 6.09$), "perform experiment—propose hypothesis" ($Z = 4.17$), and "perform experiment—analyze results" ($Z = 4.55$). The remarkable behavioral shifts of learners with low learning performance were "enter course—propose hypothesis" ($Z = 3.80$), "propose hypothesis—perform experiment" ($Z = 2.74$), "perform experiment—analyze results" ($Z = 2.21$), and "propose hypothesis—analyze results" ($Z = 2.74$). Apparently, learners with high learning performance modified their hypotheses after the experiment, while learners with low learning performance directly analyzed the results after proposing the hypotheses, without performing the inquiry buoyancy simulation experiment. Therefore, performing experimental simulation and hypothesis modification are key inquiry competence. Learners must be guided to modify hypotheses and perform the inquiry buoyancy simulation experiment to improve their inquiry learning performance.

Lag Sequential Analysis of Behaviors of Learners with Different Inquiry Competence in Inquiry Buoyancy Simulation Experiment

In the inquiry buoyancy simulation experiment, hypotheses were proposed by learners to confirm what factors affect buoyancy through experimental operations. Three phases of the inquiry buoyancy simulation experiment were evaluated separately in this study. The performances in the three phases were summed, and learners' performances that were higher and lower than the mean were divided into the learning groups with high and low inquiry competence, respectively. A lag sequential analysis of the learning behaviors of both groups in the inquiry buoyancy simulation experiment was performed. Figure 5 shows the results, which indicate that learners with a high inquiry competence exhibited the remarkable behavioral shifts of "propose hypothesis—perform experiment" ($Z = 8.25$), "perform experiment—propose hypothesis" ($Z = 5.66$), and "perform experiment—analyze results" ($Z = 3.43$), while those with a low inquiry competence exhibited the remarkable behavioral shifts of "enter course—propose hypothesis" ($Z = 3.43$), "propose hypothesis—analyze results" ($Z = 2.79$), and "perform experiment—analyze results" ($Z = 3.43$). In this case, learners with high inquiry competence modified their hypotheses and analyzed results after the experiments; learners with low learning performance, in contrast, immediately proposed hypotheses after entering the course and directly analyzed the results without

Table 12 Same sequential pattern mining result of the inquiry simulation experiment activity for learners with different learning performances (2)

High learning performance group		Low learning performance group			
Minimum support: 0.526		Minimum support: 0.518		Minimum support: 0.555	
Sequence	Action	Sequence	Action	Sequence	Action
<1>	D1	<1>	D1	<1>	D1
<2>	D1	<2>	D1	<4>	D1
<3>	D2	<5>	D2	<8>	D2

"Minimum support" means the least threshold applied to find all frequent learning behaviors

"Sequence" means the discovered learning behavior sequence under the pre-determined minimum support

"Action" means the learning behavior of the corresponding course node

The number inside < > means the order of the encoding learning behavior sequence

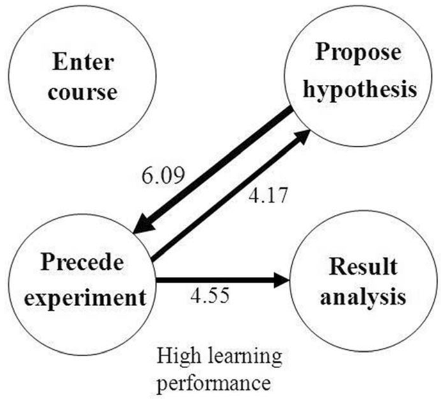


Fig. 4 Behavior sequence shift difference analysis of learners with distinct learning performances

performing the inquiry buoyancy simulation experiment. Accordingly, performing experimental simulation and hypothesis modification are critical steps in inquiry learning. Learners must be guided to modify hypotheses and perform an experimental simulation to improve their inquiry competence.

Interviews

To understand the learners' thoughts and perceptions of using the designed inquiry learning course on the topic of "Floating Future," five learners with good and five with poor average learning performance and inquiry competence were invited to take part in one-on-one interviews. Most of the interviewees expressed that they were impressed by the inquiry buoyancy simulation experiment and favored this inquiry learning mode because they could perform the simulation experiment to verify their hypotheses, thus more clearly understanding the key factors that affect buoyancy than does the traditional learning method. Furthermore, interviewees with better learning performance and inquiry competence expressed that they would induce the experiment results and adjust their hypotheses in the inquiry learning process, while interviewees with poorer learning performance and inquiry competence expressed that they did not know how to operate the experimental variables to induce the results. Thus, the concepts were still unclear to them even they have already operated the inquiry buoyancy simulation experiment. The interview results are consistent with the results of behavioral analysis by using lag sequential analysis. Most interviewees expressed that the designed inquiry learning course on CWISE should display the standard answers and detailed solutions after they had answered the test questions to improve their understanding of the ideas related to the course.

Discussion

The above experimental results verified that the "Floating Future" inquiry learning course designed on CWISE improved learners' learning performance. This result is consistent with many studies of inquiry learning (Fang et al. 2016; Taylor and Bilbrey 2012; Wu et al. 2015; Spronken-Smith 2012). Fang et al. (2016) designed six science inquiry learning units on CWISE and found that their designed courses reinforced students' conceptual knowledge and inquiry competence. Taylor and Bilbrey (2012) indicated that teachers who used the inquiry learning method to teach mathematics and science remarkably improved learners' learning performances. Wu et al. (2015) found that inquiry learning improved learners' learning performances and learning efficiency, while Spronken-Smith (2012) proposed that inquiry

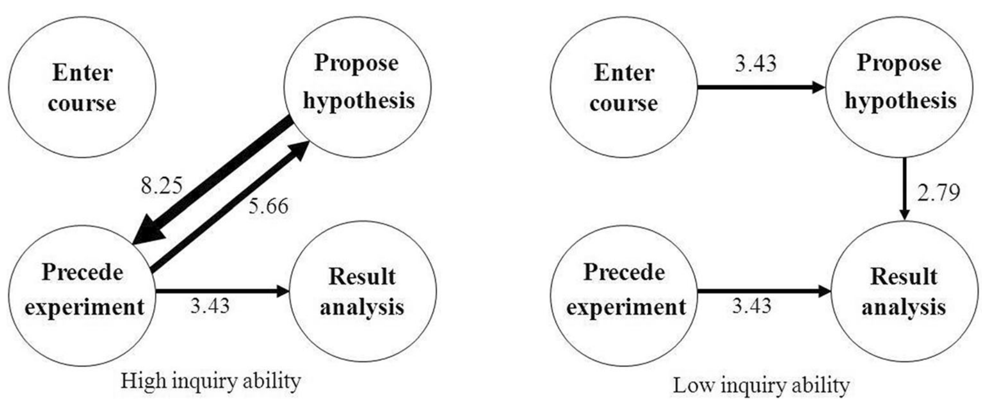


Fig. 5 Behavior sequence shift difference analysis of learners with different inquiry competences

learning promoted students' participation and their interaction with teachers, further enhancing their learning performances.

Sequential pattern mining revealed no significant difference in the viewing sequences of learners with different learning performances in the designed inquiry learning course; they all followed the sequence in the curriculum design. This result might have followed from the fact that the top-down inquiry learning course design on CWISE sequentially guided learners' learning so that most learners did not go back over concepts that have learned. Interestingly, learners with low learning performance exhibited higher course node sequence continuity than those with high learning performance in the designed inquiry learning course. In an inquiry-based learning activity, learners are encouraged to perform inquiry learning tasks in a different order according to their autonomous learning needs and knowledge acquisition statuses. The learners with low learning performance who followed up more course node sequences designed by the teacher than the learners with high learning performance may imply that they have lower self-determined learning abilities. Also, the learners with high learning performance acted continually in operating the "size," "material," and "liquid" of the water pool and the "executing" of the buoyancy simulation experiment, while those with low learning performance did not. The result shows that learners with high learning performance made more logical operation behaviors than those of low learning performance in the inquiry buoyancy simulation experiment.

Lag sequential analysis indicated that learners with high learning performance and high inquiry competence exhibited a significant behavioral shift of "perform experiment—propose hypothesis" in the inquiry buoyancy simulation experiment, while those with low learning performance and low inquiry competence exhibited a notable behavioral shift of "propose hypothesis—analyze results." Apparently, learners with high inquiry competence and high learning performance verified, judged, and modified their preset hypotheses after their experiments, while those with low inquiry competence and low learning performance would immediately analyze the experimental results after proposing their hypotheses. The theoretical foundations of inquiry-based learning are based on the constructive theory of learning, which states that knowledge is constructed by learners (Piaget 2013; Khalaf and Zin 2018). The key tenet is that an individual learner must actively construct knowledge and skills through their experience and interaction with the environment and using the correct ways (Bruner 1990). Besdies, Klahr and Dunbar (1988) regarded scientific inquiry as a dual space search and considered that learners should set hypotheses based on knowledge and experience when performing inquiry learning. Chen and She (2014) also indicated that teaching learners to verify and modify hypotheses experimentally can improve their science inquiry competence and learning performance. The above effective behavior patterns explored from learners with high learning performance and high inquiry competence echo the theoretical foundations of inquiry-based learning in the constructive theory of learning, dual space search of scientific inquiry, and empirical studies. More importantly, the results can provide useful insights to teachers so that they can advise students to avoid deviate effective learning behaviors during performing an inquiry learning activity.

Learners' inquiry competence, the total learning time in the designed inquiry learning course, and learning time in the inquiry buoyancy simulation experiment are significantly positively correlated with learning performance, with which they have causal relationships. Corlu and Corlu (2012) found that independently measured scientific inquiry levels are highly correlated with student grades in practical courses. In particular, linear regression analysis indicated that learning time in the inquiry buoyancy simulation experiment most strongly affects learning performance. As a result, science inquiry competence is determined to be one of the key factors that affect science inquiry learning performance, and the inquiry buoyancy simulation experiment importantly affects inquiry-based learning performance. Moreover, the interviews revealed that most of the interviewees were impressed with the learning activity in the buoyancy inquiry simulation experiment, and the lag sequence analysis of their learning processes revealed that learners with high inquiry competence and high learning performance adjusted their hypotheses after the simulation experiment. Thus, teachers should advise students to take much more time to conduct the inquiry buoyancy simulation experiment for getting the correct knowledge about the principles of floating and should direct students to avoid the ineffective learning behaviors of operating this simulation experiment. These results are similar to those of Rutten et al. (2012), which indicated that computer simulations can enhance science learning, based on a review of quasi-experimental research conducted over the past decade.

Conclusions and Future Work

According to the results of analyzing and mining the learning processes generated from research subjects who participated in the designed web-based inquiry learning activity on CWISE, this study found that several remarkable positive correlations among several considered learning behaviors, such as existing remarkable positive correlations between the learning time in the inquiry buoyancy simulation experiment and learning performance. The result is very helpful in determining what kind of learning behavior will affect learning performance so that teachers can guide their students to pay much more attention to engaging in this kind of learning activity. Moreover, this study also further found that several considered learning behaviors can predict learning performance, such as the existing causal relationship between the total learning time in the designed inquiry learning course and learning performance to some degree. The result provides useful information in knowing what kind of learning behavior will cause good learning performance, thus encouraging students to take much more time in this kind of learning behavior. Also, sequential pattern mining can explore the difference between learners with high and low learning performance in learning behavior sequences over time in a web-based inquiry learning activity. Especially, this study found that learners with high learning performance exhibited continual operation behaviors followed by the sequence of the designed inquiry learning materials in their inquiry learning processes, while those with low learning performance did not. The result can help teachers to diagnose students' incorrect or non-logical learning behaviors during learning processes. Finally, lag sequence analysis can also explore the difference between learners with high and low learning performance in two neighbor learning behavior sequences over time in a web-based inquiry learning activity. Importantly, this study found that learners with high learning performance and high inquiry competence exhibited a significant bidirectional behavior shift of "perform experiment—propose hypothesis" in the inquiry buoyancy simulation experiment, while those with low learning performance and low inquiry competence did not. Based on these results, teachers can know that learners must be guided to verify and modify their hypotheses throughout their experiments to improve their inquiry competences and learning performances. In conclusion, this study presents a series of systematic analysis methods that include inference statistics, sequential pattern mining, and lag sequence analysis to insight the effective learning behaviors in a web-based inquiry learning environment based on mining students' learning processes, thus providing potentially useful information in assisting teachers to know how to guide students' learning directions and suggesting students to adjust their learning behaviors and strategies.

Additional studies are warranted. First, only learners' learning processes on CWISE were discussed in this study. Various types of e-learning environments for cultivating students' learning literacy could be integrated using the xAPI technique to collect a broader range of learning process data and to favor learning process mining analyses. The results would provide an effective reference for instructional design and students' reflections on their learning. Second, analytical results of sequential pattern mining and lag sequential analysis demonstrate that learners with different inquiry competence and learning performances exhibited distinct learning behavior sequences and learning behavioral shift patterns. Based on learning behaviors, future work should improve the functions of CWISE with xAPI to provide timely reminders of learning states to guide learners when detecting their behaviors that are not associated with high inquiry competence and learning performance. Finally, the learning processes were analyzed offline in this study so that real-time feedback or reminders from the system could not be offered to the learners during the inquiry-based learning course. Therefore, future work should develop a high-speed cloud-based computing environment that can support real-time analyses of learners' learning processes and immediately display the results for teachers and learners. If so, important reference signals can be provided immediately for the real-time adjustment of teachers' instructional strategies and students' learning strategies.

Compliance with Ethical Standards

To consider the research ethics of the designed experiment that involves recording the learning behaviors of the research subjects by using xAPI technologies, written informed consent was obtained from the research subjects following a full explanation of the experiment. The informed consent letter contains the specific nature of the research, including the data that collect from them, are only for the research, their name will never appear on any data collected and that instead, we will provide a unique identification number on their data and that this information will remain secure such that only the principal investigator of this study will have access to it, the collected data that is no longer needed will be destroyed, and how participation will make a contribution to our study's goals. Moreover, all procedures performed in this study involving human participants were by the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards. Finally, we certify that there is no conflict of interest in this paper.

References

Abdi, A. (2014). The effect of inquiry-based learning method on students' academic achievement in science course. *Universal Journal of Educational Research, 2*(1), 37–41.

ADL Net. (2015). xAPI-Dashboard. Retrieved December 8, 2018, from https://github.com/adlnet/xAPI-Dashboard

Agrawal, R., & Srikant, R. (1995). Mining sequential patterns. In: *Proceedings of the Eleventh International Conference on Data Engineering* (pp. 3–14), Washington, DC, USA: IEEE Computer Society.

Arnold, J. C., Boone, W. J., Kremer, K., & Mayer, J. (2018). Assessment of competencies in scientific inquiry through the application of Rasch measurement techniques. *Education in Science, 8*(4), 184–204.

Arroyo, I., & Woolf, B. P. (2005). Inferring learning and attitudes from a Bayesian network of log file data. Proceedings of the 12th international conference on artificial intelligence in education, 33–40.

Backus, L. (2005). A year without procedures. *The Science Teacher, 72*(7), 54–58.

Bakeman, R. (1986). *Observing interaction: an introduction to sequential analysis*. New York: Cambridge University Press.

Bakeman, R., Deckner, D. F., & Quera, V. (2005). Analysis of behavioral streams. In D. M. Teti (Ed.), *Handbook of research methods in developmental science* (pp. 394–420). Oxford: Blackwell Publishers.

Bruner, J. (1990). *Acts of meaning*. Cambridge: Harvard University Press.

Buffler, A., Allie, S., & Lubben, F. (2001). The development of first year physics students' ideas about measurement in terms of point and set paradigms. *International Journal of Science Education, 23*(11), 1137–1156.

Chen, C. M., Chen, L. C., & Horng, W. J. (2019). A collaborative reading annotation system with formative assessment and feedback mechanisms to promote digital reading performance. *Interactive Learning Environments*, 1–18 https://doi.org/10.1080/10494820.2019.1636091.

Chen, C. M., & Chen, M. C. (2009). Mobile formative assessment tool based on data mining techniques for supporting web-based learning. *Computers & Education, 52*(1), 256–273.

Chen, C. M., & Lin, S. T. (2014). Assessing effects of information architecture of digital libraries on supporting e-learning: a case study on the digital library of nature & culture. *Computers & Education, 75*(1), 92–102.

Chen, C. T., & She, H. C. (2014). The effectiveness of scientific inquiry with/without integration of scientific reasoning. *International Journal of Science and Mathematics Education, 13*(1), 1–20.

Chiang, T. H. C., Yang, S. J. H., & Hwang, G. J. (2014). Students' online interactive patterns in augmented reality-based inquiry activities. *Computers & Education, 78*, 97–108.

Corlu, M. A., & Corlu, M. S. (2012). Scientific inquiry based professional development models in teacher education. *Educational Sciences: Theory & Practice, 12*(1), 514–521.

Cropley, A. J., & Page, K. (2002). Creativity in education and learning- a guide for teachers and educators. *Long Range Planning, 35*, 199–200.

Deters, K. M. (2005). Student opinions regarding inquiry-based labs. *Journal of Chemical Education, 82*(8), 1178–1180.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *The Journal of the Learning Sciences, 8*(3&4), 391–450.

Fang, S. C., Hsu, Y. S., Chang, H. Y., Chang, W. H., Wu, H. K., & Chen, C. M. (2016). Investigating the effects of structured and guided inquiry on students' development of conceptual knowledge and inquiry abilities: a case study in Taiwan. *International Journal of Science Education, 38*(2), 945–1971.

Fortenbacher, A., Beuster, L., Elkina, M., Kappe, L., Merceron, A., Pursian, A., & Wenzlaff, B. (2013). LeMo: a learning analytics application focussing on user path analysis and interactive visualization. In: *IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS), 2*, 748–753.

Fournier-Viger, P., Nkambou, R., & Nguifo, E. M. (2008). A knowledge discovery framework for learning task models from user interactions in intelligent tutoring systems. In: *Proceedings of the 7th Mexican International Conference on Artificial Intelligence: Advances in Artificial Intelligence*, MICAI '08, pages 765–778, Berlin, Heidelberg, Springer-Verlag.

Fournier-Viger, P., Nkambou, R., Nguifo, E. M., & Faghihi, U. (2009). Building agents that learn by observing other agents performing a task: a sequential pattern mining approach. In B. C. Chien & T. P. Hong (Eds.), *Opportunities and challenges for next-generation applied intelligence. Studies in computational intelligence, 214*. Berlin, Heidelberg: Springer.

Fournier-Viger, P., Gomariz, A., Gueniche, T., Soltani, A., Wu, C. W., & Tseng, V. S. (2014). SPMF: a Java open-source pattern mining library. *Journal of Machine Learning Research, Res, 15*, 3389–3393.

Gobert, J. D., Pedro, M. S., Raziuddin, J., & Baker, R. S. (2013). From log files to assessment metrics: measuring students' science inquiry skills using educational data mining. *The Journal of the Learning Sciences, 22*(4), 521–563.

Hakverdi-Can, M., & Sönmez, D. (2012). Learning how to design a technology supported inquiry-based learning environment. *Science Education International, 23*(4), 338–352.

Iacobucci, D., Posavac, S., Kardes, F. R., Schneider, M. J., & Popovich, D. L. (2015). The median split: robust, refined, and revived. *Journal of Consumer Psychology, 25*(4), 690–704.

Kamber, M., Han, J., & Pei, J. (2012). *Data mining: concepts and techniques* (3rd ed.). Boston: Morgan Kaufmann.

Khalaf, B. K., & Zin, Z. B. M. (2018). Traditional and inquiry-based learning pedagogy: a systematic critical review. *International Journal of Instruction, 11*(4), 545–564.

Khan, M., Shaukat, H., Riasat, A., Majoka, M. I., & Ramzan, M. (2011). Effect of inquiry method on achievement of students in chemistry at secondary level. *International Journal of Academic Research, 3*(1), 955–959.

Kim, M. C., Hannafin, M. J., & Bryan, L. A. (2007). Technology-enhanced inquiry tools in science education: an emerging pedagogical framework for classroom practice. *Science Education, 91*(6), 1010–1030.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*(1), 1–48.

Li, B., Kuo, R., Chang, M., & Garn, K. (2015). Reward points calculation based on sequential pattern analysis in an educational mobile app. In: *Proceedings of 21st International Conference on Distributed Multimedia Systems* (pp. 186-190), Vancouver, Canada.

Lim, K. C. (2015). Case studies of xAPI applications to e-learning. The twelfth international conference on eLearning for knowledge-based society, 3.1–3.12.

Linn, M. C., & Eylon, B. S. (2011). *Science learning and instruction: taking advantage of technology to promote knowledge integration*. New York: Routledge.

Manso-Vázquez, M., Caeiro-Rodríguez, M., & Llamas-Nistal, M. (2015). xAPI-SRL: Uses of an application profile for self-regulated learning based on the analysis of learning strategies. In: *Proceedings of the 2015 IEEE Frontiers in education conference*. IEEE Computer Society, Washington, DC, 1–8.

Marono, A., Clarke, D., Navarro, J., & Keatley, D. (2018). A sequence analysis of nonverbal behaviour and deception. *Journal of Police and Criminal Psychology, 33*(2), 109–117.

Marx, R. W., Blumenfeld, P. C., Krajcik, J. S., Fishman, B., Soloway, E., Geier, R., & Tal, R. T. (2004). Inquiry-based science in the middle grades: assessment of learning in urban systemic reform. *Journal of Research in Science Teaching, 41*(10), 1063–1080.

Moran, G., Dumas, J. E., & Symons, D. K. (1992). Approaches to sequential analysis and the description of contingency in behavioral interaction. *Behavioral Assessment, 14*, 65–92.

Myers, M. J., & Burgess, A. B. (2003). Inquiry-based laboratory course improves students' ability to design experiments and interpret data. *Advances in Physiology Education, 27*(1–4), 26–33.

National Research Council. (2000). *Inquiry and the National Science Education Standards: a guide for teaching and learning*. Washington, DC: National Research Council.

National Research Council. (2012). *A framework for K-12 science education: practices, crosscutting concepts, and core ideas*. Washington, DC: National Academies Press.

Pan, W., & Hawryszkiewycz, I. (2004). A method of defining learning processes. In R. Atkinson, C. McBeath, D. Jonas-Dwyer, & R. Phillips (Eds.), *Beyond the comfort zone: Proceedings of the 21st ASC1LnE Conference* (pp. 734–742). Perth: 5-8 December.

Piaget, J. (2013). *Principles of genetic epistemology: Selected works* (Vol. 7). Routledge.

Quera, V., & Bakeman, R. (2000). Quantification strategies in behavioral observation research. In T. Thompson, D. Felce, & F. J. Symons (Eds.), *Behavioral observation* (1st ed., pp. 297–315).

Rabelo, T., Lama, M., Amorim, R. R., & Juan C. Vidal, J. C. (2015). SmartLAK: A big data architecture for supporting learning analytics services. *Proceedings of 2015* Frontiers in education conference (FIE); El Paso, Texas, pp. 1–5.

Raes, A., & Schellens, T. (2015). Unraveling the motivational effects and challenges of web-based collaborative inquiry learning across different groups of learners. *Educational Technology Research and Development, 63*(3), 405–430.

Rutten, N., van Joolingen, W. R., & van der Veen, J. T. (2012). The learning effects of computer simulations in science education. *Computers & Education, 58*(1), 136–153.

Sparapani, E. F., Abel, F. J., Easton, S. E., Edwards, P., & Herbster, D. L. (1996). Portfolio assessment: a way to authentically monitor progress and evaluate teacher preparation. In: *Annual Meeting of the Association of Teacher Educators' 76th Annual Meeting* (pp. 1–20), St. Louis, Missouri.

SPMF. (2015). Documentation SPMF. Retrieved October 19, 2018, from http://www.philippe-fournier-viger.com/spmf/index.php?link=documentation.php#example13 .

Spronken-Smith, R. (2012). Experiencing the process of knowledge creation: the nature and use of inquiry-based learning in higher education. *The Journal of Geography in Higher Education, 2*, 183–201.

Taylor, J., & Bilbrey, J. (2012). Effectiveness of inquiry based and teacher directed instruction in an Alabama elementary school. *Journal of Instructional Pedagogies, 8*, 17.

Tin Can API. (2015). *What is the Tin Can API?* Retrieved October 17, 2018, from https://tincanapi.com/overview/

Watkins, J. (2016). *Effects of self-monitoring during inquiry based learning on the behavior and academic performance of at-risk middle school students.* (Electronic Thesis or Dissertation). Retrieved from https://etd.ohiolink.edu/

Williams, P., Wray, J., Farrall, H., & Aspland, J. (2014). Fit for purpose: traditional assessment is failing undergraduates with learning difficulties. Might eAssessment help? *International Journal of Inclusive Education, 18*(6), 614–625.

Wu, H. K., & Hsieh, C. E. (2006). Developing sixth graders' inquiry skills to construct explanations in inquiry-based learning environments. *International Journal of Science Education, 28*(11), 1289–1313.

Wu, J. W., Tseng, J. C. R., & Hwang, G. J. (2015). Development of an inquiry-based learning support system based on an intelligent knowledge exploration approach. *Educational Technology & Society, 18*(3), 282–300.

Yang, T. C., Chen, Y., & Hwang, G. J. (2015). The influences of a two-tier test strategy on student learning: a lag sequential analysis approach. *Computers & Education, 82*, 366–377.