Analysis of learning behavior and knowledge changes in inquiry-based learning under mixed learning environments

Jiadong Chen[1,a) Masaya Okada[1,b) Yuta Taniguchi[1,c) Shin'ichi Konomi[1,d)

Abstract: In open-ended learning tasks, the knowledge learned by the learners may be diverse. In order to help learners improve their performance in open-ended learning tasks, it is necessary to understand what behaviors will influence learning outcomes. In this research, we present an experiment based on inquiry-based learning involving multiple types of resources including the web, book summaries and the university library. Learners' usage data and concept map data have been collected to analyze their behaviors and learning outcomes. Through cluster analysis, we have identified 4 types of behavior patterns. By combining learning outcomes, we found that learners who use various types of materials in a balanced way get the best learning results.

Keywords: Educational Data Mining, Learning Analysis, Inquiry-based learning, Concept Map

1. Introduction

Inquiry-based learning has been widely exploited in many educational settings. Different from traditional education, inquiry-based learning enables learners to actively acquire knowledge, apply knowledge, and solve problems, thereby increasing their knowledge level. At the same time, due to the initiative and openness of inquiry-based learning, the different learning behaviors of learners in the learning process often lead to different learning outcomes. Therefore, if we want to help students improve the learning outcomes obtained through inquiry-based learning, it is necessary to investigate the relation between learners' behaviors and learning outcomes. Regarding inquiry-based learning in the field environments, previous study have explored the relation between learners' behavior and knowledge diversity[1]. For the case in the information space or indoor environment, some studies have tried to use search logs[2] or e-book reading logs[3] to identify behavioral patterns that can affect learning outcomes. However, in the above studies, learning processes basically involve a single environment, and the analysis of learners' learning behaviors in mixed environments is relatively rare.

Therefore, this research will focus on the learning behavior of learners in a mixed learning environment, and try to explore the influence of different behavior patterns on learning outcomes. We conducted a user experiment in a university and in the experiment we provide learners with three different inquiry environments at the same time: the internet, the library, and the book summary respectively. The participants behavioral data in different environments will be collected throughout the inquiry-based learning learning process. We perform cluster analysis on the collected data to identify user behavior patterns. As for the evaluation of learning outcomes, concept map-based graphical tools are used to understand what learners have learned. We use, the one-way Anova test Whether the learning outcomes of learners with different behavior patterns are different. We hope that through this research, we can provide basic insights for designing a learning environment that expands learners' knowledge more effectively.

2. Related Work

With the development of science and technology, people's learning methods are also changing. The emergence of electronic search systems has greatly accelerated people's access to information. Because it enables users to quickly and conveniently obtain the information they need, more and more people are also using it for learning. For example, in inquiry-based learning, in addition to discussing with others or reading library materials web search has also become a choice when learners need to find out information about a problem. In some previous studies, researchers have present a perspective on searching as learning which focuses on the learning that occurs during the search and the learning results obtained[4]. The article points out the importance of search for learning and hoped to further explain the relation between search and learning. In the research on search as learning, search systems are usually regarded as learning tools. However, the current search system is designed to help users obtain relevant information, not to help learners improve their learning outcomes[5]. Therefore, in order to help learners obtain better learning results in search, we need to have a deeper understanding of learners'

1 Kyushu University
a) chen.jiadong.450@s.kyushu-u.ac.jp
b) okada@kyoso.kyushu-u.ac.jp
c) taniguchi@ait.kyushu-u.ac.jp
d) konomi.shinichi.848@m.kyushu-u.ac.jp

search behavior and how learners learn in search. Previously, researchers have analysed different behaviours in search to identify factors influencing learning outcomes. Collins-Thompson et al. surveyed the time users' spent per document in the search, they found that searchers who spend more time reading documents will get better learning performance[6]. Carsten Eickhoff et al. used data such as user query words and web page content in the log to point out the influence of searchers' expertise on search behavior[2]. Liu et al. analyze the user's preferences for different sources of information obtained by the search system, and examined its influence on learning outcomes under different types of search learning tasks.[7] However, the search behaviors considered in the above studies are limited to one search system environment.

Therefore, this paper aims at studying the user user behavior in mixed learning environment. User behavior will be described using Markov model. In previous studies, Markov model has been widely used in the study of user behavior. For example, Adrien Nouvellet, et al. used the mixture of Markov models to model the behavior of using digital library. Analyze various behavior patterns through clustering methods to help managers improve the digital library[8]. Mikhail Ageev et al.,used the Markov model to model whether the user successfully completed the search task by using the user's search behavior. The obtained model is used to predict the success of other users' searches[9].

The evaluation of open-ended learning outcomes is more difficult than evaluating traditional learning outcomes. This is because the knowledge acquired by students through learning is diverse, and it is difficult for teachers to test through test papers with fixed questions. Wilson, M. J. et al. proposed a method for evaluating learning outcomes in open-ended learning[10]. The method is based on Bloom's taxonomy[11], and judges uses the defined evaluation criteria to quantify the handwritten content generated by the learner. In addition, the concept map is used as a tool in the educational environment to organize knowledge, plan learning strategies, and also evaluate[12]. Learners are able to demonstrate their knowledge structure by drawing a concept map which containing concepts, connecting lines, and connecting words. Okada et al. have used concept map to analyze learners' knowledge changes in inquiry-based learning[1]. So, in this study, we will also use concept maps and through quantitative methods we can evaluate learners' learning outcomes.

3. Method

3.1 Participants and Context

The experiment is conducted base on the "interdisciplinary collaborative learning of social issues course" at Kyushu University in the 2020. The course consists of four sessions of two and half hours. Students taking this course will be divided into groups of 4 to 6 and they will study in groups around a given topic. This year's topic is "Design data utilization services for 2030". Our experiment will be conducted in the first two sessions of the course, students are required to engage in 50 minutes of inquiry-based learning in each session. Two inquiry-based learning is represented as L1 and L2. In L1, the teacher will ask students to investigate basic information related to the topic, and in L2 the teacher will encourage students to build on the knowledge gained in L1 and to target their investigations in a more focused way. During both L1 and L2 the students can search the web, read the book summaries, and go to the library to check out materials. At the end of each inquiry-based learning session students are asked to submit a concept map as a basis for assessing the learning outcomes.

Our participants are mainly recruited from the students who take this course. In addition, in order to increase the number of participants, we also recruited students across the Kyushu University. For these participants, they need to take simulation course to ensure consistency with the participants who take the "interdisciplinary collaborative learning of social issues" course. Finally, we have 22 participants, they are mainly freshmen, because the "interdisciplinary collaborative learning of social issues" course is a first-year undergraduate course. The participants majored in various subjects, e.g. engineering, science, information science, pharmacy and economics domains.

3.2 Data Collection

In this experiment, we consider the influence of behaviour in a mixed learning environment on students learning outcomes. Due to the difference of search environment, the different data collection approaches are also required. In the following we will describe each data collection method separately.

Web Search For web search, we choose to collect browser history data to analyze learner behavior.In the process of inquiry-based learning, for easy collection we require participants to use the Google Chrome browser. The chrome browser supports the installation of extensions with various functions. An open source extension named history master[*1] can extract the browser history in TXT format. The extracted files include visit-time, title, visit-count, typed-count, id, and URL information when the user uses the browser to visit the web pages. Before the start of the experiment, we will instruct participants to use the history master extension, and at the end of the experiment, participants will submit the extracted browser history data in TXT format.

Book summary SERENDIP[*2] book summary service provides us with 200 book summaries in PDF format. Its contents include summaries of books published between 2018 and 2020, covering 23 categories such as economics, law, skills, science and technology. A PDF viewer system was developed to allow students to read and query these book summaries and collect usage data, as shown in Fig.1. Participants need to use a username and password to log in to the system. And in this system, book summaries are sorted according to content categories, and users can click on the title to read the corresponding book summary. In addition, the system also provides the search function, and users can also search for book summary by using keywords. During the experiment, participants can freely access the system through laptops, mobile phones, tablets and other devices. And the user's usage information of the system will be automatically recorded in the database. The information recorded includes the user's id, the id of book summary, and the timestamp of the access.

*1 https://github.com/jiacai2050/history-master
*2 https://www.serendip.site/

Fig. 1 PDF viewer system

Library During the inquiry-based learning, students are encouraged to visit the library to conduct research. Before the studying begins, we prepare sticky notes for each participant and tell them to record information about the books they check out and what interests them. The sticky notes containing the memos recorded by the participants will be submitted at the end of the study.

Concept Map Compared to drawing concept maps directly on paper, it is more convenient to use special software. Especially for some complex knowledge structures, students can easily modify concept agreement without reducing the readability[13][14]. In this experiment, we use an online collaborative whiteboarding platform called miro[*3] to help students generate their own concept maps. Miro is easy to use, no download, and has share function. Before the experiment begins, participants were trained in the basic operation of miro to ensure they could draw their own concept maps. The concept map generated using miro is shown in Fig.2. After each inquiry-based learning session is completed, participants used miro's sharing feature to generate URL links for their own concept maps. We collect these links, through which we can access and view the concept maps generated by the participants.

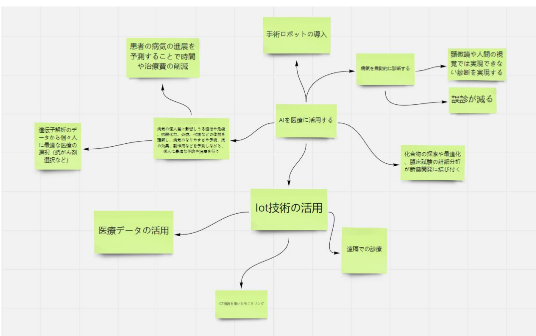

Fig. 2 Concept map generated by miro

Based on the source of the data, data collection can be classified into 3 categories, Questionnaire-based Data Collection (QDC), Manual Data Collection (MDC), Automatic Data Collection (ADC)[15]. Table 1 shows the collection approach and categories of each type of data. Among them, the data belonging to the ADC categories can be conveniently processed by the computer. They are automatically generated by the computer, so they are in a format that the computer can read. As for the data of the MDC category, computers sometimes cannot process it, so manual analysis is required. In the following chapters will introduce the analysis methods for these data.

Table 1 Collection of different types of data

Data Sources	Category	Approach
Web Search	ADC	Browser history record
Book Summary	ADC	Access log
Library	MDC	Handwritten memo
Concept Map	MDC	Image generated by miro

3.3 Representation of learning behaviour

Students' behavior during inquiry-based learning will be modeled using Markov models. First, the behavior state needs to be defined. In this study, we will focus on exploring students' behaviors of using different types of resources in a mixed learning environment. For the resources provided by the foundation, we will define 4 behavior states namely: *Using the ordinary websites*, *Using the search engine websites*, *Using the book summaries*, and *Using the library*. In particular, due to the functionality of search engine websites, it is meaningful to distinguish them from ordinary websites for analysis. Based on the collected data, we can determine the state of the student at each moment, the specific method is as follows:

Using the ordinary website:

Using the URL information recorded in the history record, the host name of the website can be extracted. If the host name does not belong to search engine websites such as www.google.com, www.yahoo.co.jp, etc., it is considered that the student is using an ordinary website.

Using the search engine websites:

Similarly, if the host name of the historical record is a search engine website such as www.google.com, it can be determined that the student is in the state of using the search engine website at that moment.

Using the book summaries:

Through the access log in the PDF viewer system, we can determine when the student is in the state of using book summery.

Using the library:

When using the library, students will leave their seats to go to the library, so they cannot use computers to access other resources. Therefore, based on the total length of inquiry-base learning, if the time interval between two adjacent data of a student exceeds 15 minutes, the student is considered to be in the

*3 https://miro.com/

state of using the library.

After that, we coded each state, using the ordinary website, using the search engine website, using the book summaries and using the library are denoted as 0, 1, 2 and 3 respectively. And the $S = \{0, 1, 2, 3\}$ will be used as the state space of the Markov model.

Based on the state space S, the participant's behavior data can also be transformed into a sequence of states $[s_0, s_1, \ldots, s_n]$, for $\forall\ i,\ s_i \in S$. The transition probability of the Markov chain can be statistically inferred from the collected data. Therefor, we can express the behavior of each participant through the state transition matrix. we can build a state transition diagram for each participant. An state transition diagram of a participant is shown in Fig.3.

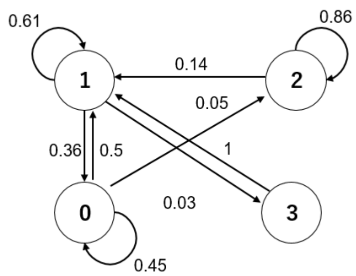

Fig. 3 State transition diagram of a participant

3.4 Evaluating Learning Outcomes

Concept maps are used as graphical tools to show the structured knowledge acquired by the learner. But to use it as an assessment tool for evaluating learning outcomes, we also need a reliable quantitative method for concept maps. Watson et al. compared the three scoring methods proposed in the previous literature[16], including the traditional scoring methods, holistic scoring methods, and categorical scoring methods[17][18][19]. Among them, the holistic scoring method is the most sensitive to detecting differences in the breadth, depth, connection and overall quality of knowledge related to the central topic. The holistic scoring method evaluates the concept map from the three categories of Comprehensiveness, Organization, and Correctness. During the evaluation, the judges use a three-point scale to score the concept map for each category and the total score of the concept map can be obtained by simply adding the three sub-scores. In addition, because subjectivity may cause differences in scoring, the application of the holistic scoring method requires at least two judges.

In this study, we apply the holistic scoring method to score participants' concept maps. For our course topic, we have identified 4 main dimensions as references for evaluating the depth and breadth of knowledge.The following dimensions are determined through discussion with the teacher of this course.

Dimensions of Digital Society Design:

- `Digital technologies`
- `Ethical issues of digital societies`
- `Social context and issues`
- `Domain knowledge`

We have 3 judges, they are master and PhD students from the School of Information Science. After training them to use the holistic scoring method, we selected 7 concept maps generated by the students and asked the judges to score them independently. According to this paper(Assessing Conceptual Knowledge Using Three Concept Map Scoring Methods), Krippendorff's alpha is used to quantify inter-rater reliability. Compared with the widely used Cohen's kappa, Krippendorff's alpha can be used for any number of judges and any type of data. By calculating if the Krippendorff's alpha is greater than 0.8, it is designated as "completely acceptable", and if it is greater than 0.67, it is classified as "acceptable for exploratory research"[20]. For the 7 selected concept maps generated by students, 3 judges gave similar scores. The Krippendorff's alpha of each sub-score in the holistic scoring method are Comprehensiveness: 0.807692, Organization: 0.795918, Correctness: 0.68254 respectively (all over 0.67). After that, the three judges completed the scoring of all the concept maps. For the differences, the judges will discuss and finally get a consistent score. The score obtained for the concept map will be used as a metric to evaluate the learning outcomes. Table 4 shows the rubric of the holistic scoring method. It's modified from [18].

4. Result

4.1 Data Processing

Using the data collection approaches introduced earlier, we collected data from 22 participants in inquiry-based learning. Among these data, there are 1436 records from the browser history, 89 from the PDF-viewer system, and 44 concept maps (two for each participant). In particular, although we asked participants to use sticky notes to record the information they found in the library of interest, no participants were able to submit effective content. Afterwards, all data from different sources was integrated, and the data of each participant will be sorted in order of timestamp. The data in the sequence will be coded according to the previously defined behavior to obtain the participant's behavior state sequence. For all 44 concept maps, the judges scored according to three categories: comprehensiveness, organization, and correctness. Finally, 132 score data (three for each concept map) in the range of 1-3 are obtained. All data will be differentiated according to whether they are from the first inquiry-based learning (L1) or from the second (L2), and the two sets of data will be analysed separately.

4.2 Influence of related knowledge

In our experiment, participants engaged in two inquiry-based learning sessions, denoted L1 and L2, which differed in that L2 took place after L1 and L2 exploration built on the knowledge gained in L1. That is, participants at L2 are more familiar with the topic to be explored than at L1, and participants have more knowledge about the topic. Therefore, in order to check whether the difference of related knowledge will affect the learner's behavior, We performed student t-tests on the number of times participants were in each state. The significance level was taken as 0.05, and results of the test are presented in Table 2.

Fig. 4 Concept Map Scoring Rubric

	Score = 1	Score = 2	Score = 3
Comprehensiveness	Map lacks subject definition. Knowledge is very simple or/and limited. Limited breadth of concepts. The map barely covers some of the qualities of the subject area.(i.e., Only one major dimension was mentioned in the concept map, or/and the details provided were insufficient.)	Map has adequate subject definition, but knowledge is limited in some areas. Map suggests a somewhat narrow understanding of the subject matter.(i.e., No less than two major dimensions are included in the concept map, but some dimensions are not sufficiently detailed.)	Map completely defines the subject area, with content lacking in no more than one major dimension. Content demonstrates extensive understanding of the subject matter. (i.e.,At least three major dimensions are included in the concept map and all of them are provided in sufficient detail.)
Organization	Map arranged with concepts mostly linearly connected. Few (or no) connections within/between the branches. Concepts are not well integrated.	The map has adequate organization with some connections within or between branches. Some, but not complete, integration of branches is apparent. A few feedback loops may exist.	Map is well organized with concept integration and the use of feedback loops. Sophisticated branch structure and connectivity.
Correctness	Map is naive and contains inaccurate knowledge and/or misconceptions about the subject area; inappropriate words or terms are used. The map documents an inaccurate understanding of certain subject matter.	Map has few subject matter inaccuracies. Most links are correct.	Map integrates concepts properly and reflects an accurate understanding of subject matter. Few or no inaccuracies and/or misconceptions

Major dimension : Digital technologies, Ethical issues of digital societies, Social context and issues, Domain knowledge

Table 2 Student's t test of Behavior

	Mean (1st/2nd)	STD (1st/2nd)	Statistic	P value
OW	13.77/21.55	18.87/27.25	-1.075	0.2885
SEW	15.0/14.95	14.72/10.21	0.0116	0.9907
BS	2.41/1.64	4.00/2.55	0.7469	0.4592
LIB	0.23/0.18	0.42/0.39	0.3657	0.7164

OW : Using the ordinary website SEW : Using the Search Engine Website
BS : Using the Book Summary LIB : Using the Library

Through the results we found that there was no significant difference in the behaviour of the participants.

Further, we tested the learning outcomes between the L1 and L2. The results are shown in Table 3. There was a significant difference between L1 and L2 in the comprehensive category scores of the concept map ($p = 0.0152 < 0.05$, L1 mean = 1.64, L2 mean = 2.09), which demonstrates that the breadth and depth of knowledge gained by students through inquiry-based learning was better in L2. That means the learning outcome of inquiry-based learning is better when the learners have more relevant knowledge about the topic.

Table 3 Student's t test of Score

	Mean (1st/2nd)	STD (1st/2nd)	Statistic	P value
COM	1.64/2.09	0.48/0.67	-2.53030*	0.0152
ORG	1.91/1.82	0.67/0.72	0.4255	0.6727
COR	2.27/2.09	0.62/0.51	1.0377	0.3053
TOTAL	5.81/6.0	1.19/1.41	-0.4504	0.0.6547

COM : Comprehensiveness ORG : Organization
COR : Correctness TOTAL : Total Score

4.3 Cluster analysis

Based on the collected data, the behavior of each participant can be represented by a transition matrix. To further explore the behavioural patterns of the participants, we conducted a cluster analysis of these behaviours. We have chosen to use hierarchical clustering and the distances will be calculated according to the Ward method[21]. Clustering will be done based on the probability distribution of the participants' behaviour, and given the nature of Markov chains, the transfer matrix can be replaced with its stationary distribution when clustering. The data from L1 and L2 are clustered separately, and their hierarchical dendrogram is shown in Fig. 5.

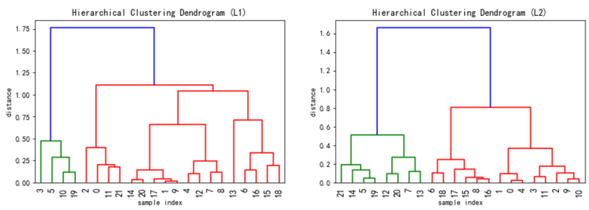

Fig. 5 Hierarchical Clustering Dendrogram

For L1, we chose 0.75 as the maximum distance, and participants were divided into four clusters, the number of people in each cluster is (C1:4, C2:4 C3:9, C4:5). For L2 we chose 0.4 as the maximum distance and participants were also divided into 4 clusters the number of people in each group is (C1:4, C2:4, C3:6, C4:8). The probability distribution of participants in each state in the cluster is shown in Fig. 6. Based on the observations, we can group the individual clusters into four behavioural patterns:

Ordinary websites use preference type
Participants of this type will choose to use ordinary web content for exploratory learning with a high probability. In contrast, the probability of using other types of resources is far less than using ordinary web pages. (e.g, C1 in L1, C1,C2 in L2)

Search engine websites use preference type
Participants of this type will use search engine websites more often, but not often visit the searched web pages or book summery. (e.g, C2 in L1, C4 in L2)

Websites use preference type
Participants in this category will use web content (both search engine websites and ordinary websites) more balanced, but rarely use the book summary or library. (e.g, C3 in L1, C3 in L2)

Websites and book summary use preference type

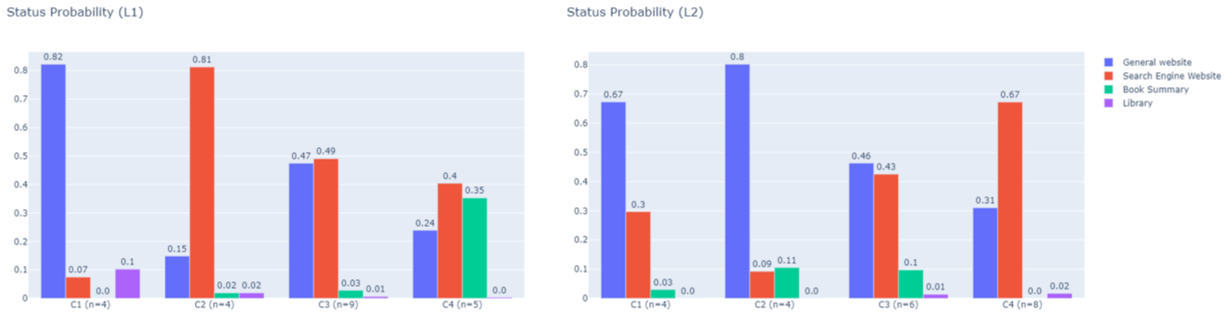

Fig. 6 Probability of each state

Participants of this type will use web content and book summary in a balanced manner, but still rarely use the library. (e.g, C4 in L1)

4.4 Relationships between clusters and learning outcomes

In order to understand the relationship between clusters and learning outcomes, we performed a one-way ANOVA on the clustering result. The significance level was set to 0.05, and the results are shown in Table 4.

Table 4 ANOVA on Clustering Result

	COM	ORG	COR	TOTAL
L1	0.364	1.662	2.130	3.405*
L2	1.140	2.774	0.867	2.597

$^*p < 0.05$

In L1, we found that the total score of the concept map is significantly different in each cluster ($p = 0.04 < 0.05$). Observing the mean value of each cluster (C1: 5.0, C2: 4.75, C3: 6.33, C4: 6.4), we can find that C4 has the highest average total score. In L2, we did not find a statistically significant difference.

5. Discussion

As a result of the analysis of the experimental data, we have gained some basic insights into the learning behaviour of learners in a mixed learning environment. Firstly, we compared learners' behaviour and learning outcomes in two inquiry-based learning sessions. It was found that after the first inquiry-based learning session, the second learning outcome is better in terms of the comprehensiveness of knowledge. This suggests that learners are more able to investigate topics in depth when they are more familiar with them. Conversely, if learners are helped to become familiar with the topic quickly, their learning outcomes in inquiry-based learning will also be improved. We then carried out a cluster analysis of the learner's behaviour. The results obtained from the observations revealed four patterns of student behaviour in the mixed learning environment. These were: Ordinary websites use preference type, Search engine websites use preference type, Websites use preference type, Websites and book summaries use preference type.It is worth noting that in all types the participants' probability of using the library was very low. And although we encouraged participants to use sticky notes to record information of interest obtained in the library before the experiment began, no participants were able to submit valid notes. The time factor may be part of the reason for this, as obtaining the information needed through the library would take far more time and effort than other means, and therefore participants would prefer to use other means of accessing information in the limited time available. Finally, we have analysed the relationship between various behavioural patterns and learning outcomes. The results show that the participants of the websites and book summaries use preference type achieved the best learning outcomes in the inquiry-based learning and were statistically significant. This suggests that using a combination of different types of materials in learning process can help learners to achieve better results than if only one type of material is used.

6. Conclusion

In this paper, we explored the behaviors and learning outcomes of learners in inquiry-based learning by conducting experiment. Different from the previous research, we provide learners with a mixed learning environment including web, book summaries and library. Through the analysis of the collected data and the concept map of the learners, we have identified four behavioral patterns in mixed learning environments. Among them, learners who use various resources in a balanced manner have obtained the best learning outcomes in the first inquiry-based learning sessions. In addition, our results suggest that learners' familiarity with the survey topic influences the depth and breadth of their exploration of this topic. Through this research, we hope that our results can provide insights for the development of support systems that enhance the learning outcomes of inquiry-based learning.

References

[1] Okada, M., Tada, M. et al.: Extracting Spatial Characteristics of a Real-world Learning Field by Behavior Sensing and Knowledge Externalization, *Journal of Information Processing*, Vol. 53, No. 4, pp. 1433–1447 (2012).

[2] Eickhoff, C., Teevan, J., White, R. and Dumais, S.: Lessons from the journey: a query log analysis of within-session learning, *Proceedings of the 7th ACM international conference on Web search and data mining*, pp. 223–232 (2014).

[3] Yin, C., Yamada, M., Oi, M., Shimada, A., Okubo, F., Kojima, K. and Ogata, H.: Exploring the relationships between reading behavior patterns and learning outcomes based on log data from e-books: A human Factor Approach, *International Journal of Human–Computer Interaction*, Vol. 35, No. 4-5, pp. 313–322 (2019).

[4] Rieh, S. Y., Collins-Thompson, K., Hansen, P. and Lee, H.-J.: Towards searching as a learning process: A review of current perspectives and future directions, *Journal of Information Science*, Vol. 42, No. 1, pp. 19–34 (2016).

[5] Rieh, S. Y., Gwizdka, J., Freund, L. and Collins-Thompson, K.: Searching as learning: Novel measures for information interaction research, *Proceedings of the American Society for Information Science and Technology*, Vol. 51, No. 1, pp. 1–4 (2014).

[6] Collins-Thompson, K., Rieh, S. Y., Haynes, C. C. and Syed, R.: Assessing learning outcomes in web search: A comparison of tasks and query strategies, *Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval*, pp. 163–172 (2016).

[7] Liu, C. and Song, X.: How do Information Source Selection Strategies Influence Users' Learning Outcomes', *Proceedings of the 2018 Conference on Human Information Interaction & Retrieval*, pp. 257–260 (2018).

[8] Nouvellet, A., Beaudouin, V., d'Alché Buc, F., Prieur, C. and Roueff, F.: Discovery of usage patterns in digital library web logs using Markov modeling (2019).

[9] Ageev, M., Guo, Q., Lagun, D. and Agichtein, E.: Find it if you can: a game for modeling different types of web search success using interaction data, *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information Retrieval*, pp. 345–354 (2011).

[10] Wilson, M. J. and Wilson, M. L.: A comparison of techniques for measuring sensemaking and learning within participant-generated summaries, *Journal of the American Society for Information Science and Technology*, Vol. 64, No. 2, pp. 291–306 (2013).

[11] Bloom, B. S.: Taxonomy of educational objectives: The classification of educational goals, *Cognitive domain* (1956).

[12] McClure, J. R., Sonak, B. and Suen, H. K.: Concept map assessment of classroom learning: Reliability, validity, and logistical practicality, *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, Vol. 36, No. 4, pp. 475–492 (1999).

[13] Cañas, A. J., Hill, G., Carff, R., Suri, N., Lott, J., Gómez, G., Eskridge, T. C., Arroyo, M. and Carvajal, R.: CmapTools: A knowledge modeling and sharing environment (2004).

[14] Ruiz-Primo, M. A. and Shavelson, R. J.: Problems and issues in the use of concept maps in science assessment, *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, Vol. 33, No. 6, pp. 569–600 (1996).

[15] Yin, C., Okubo, F., Shimada, A., Oi, M., Hirokawa, S. and Ogata, H.: Identifying and analyzing the learning behaviors of students using e-books, *Proceedings of the 23rd International Conference on Computers in Education*, Asia-Pacific Society for Computers in Education Hangzhou, China, pp. 118–120 (2015).

[16] Watson, M. K., Pelkey, J., Noyes, C. R. and Rodgers, M. O.: Assessing conceptual knowledge using three concept map scoring methods, *Journal of engineering education*, Vol. 105, No. 1, pp. 118–146 (2016).

[17] Novak, J. D., Gowin, D. B. and Bob, G. D.: *Learning how to learn*, cambridge University press (1984).

[18] Besterfield-Sacre, M., Gerchak, J., Lyons, M. R., Shuman, L. J. and Wolfe, H.: Scoring concept maps: An integrated rubric for assessing engineering education, *Journal of Engineering Education*, Vol. 93, No. 2, pp. 105–115 (2004).

[19] Segalàs, J., Ferrer-Balas, D. and Mulder, K. F.: What do engineering students learn in sustainability courses? The effect of the pedagogical approach, *Journal of Cleaner Production*, Vol. 18, No. 3, pp. 275–284 (2010).

[20] Krippendorff, K.: *Content analysis: An introduction to its methodology*, Sage publications (2018).

[21] Ward Jr, J. H.: Hierarchical grouping to optimize an objective function, *Journal of the American statistical association*, Vol. 58, No. 301, pp. 236–244 (1963).

