Kuromiya et al. *Research and Practice in Technology Enhanced Learning* (2024) 19:34
https://doi.org/10.58459/rptel.2024.19034

Research and Practice in Technology Enhanced Learning

RESEARCH Free and Open Access

Learning analytics for student homework activities during a long break: Evidence from K-12 education in Japan

Hiroyuki Kuromiya [1] *, Rwitajit Majumdar [2], Izumi Horikoshi [3] and Hiroaki Ogata [3]

*Correspondence:
khiroyuki1993@gmail.com
Graduate School of Informatics, Kyoto University, Kyoto, Japan
Full list of author information is available at the end of the article

Abstract

Learning Analytics (LA) is an emergent field that aims to better understand students and provide intelligence to learners, teachers, and administrators using learning log data. Although the use of technology in class is increasing in the K-12 sector and tertiary education, cases of effective implementation of LA in secondary schools have rarely been reported. This study offers an example of LA implemented in a junior high Math class during long vacations in Japan. This paper comprises two studies: first, we analyzed 121 students' answer logs and their exam performance after vacation by the K-means clustering method. We found that students' progress patterns were categorized into four types of engagement—early, late, high, and low—and the early and high-engagement groups obtained significantly higher scores than the low-engagement group. In the second study, we implemented a real-time dashboard that visualizes students' progress patterns and gives students insights about their progress during the vacation period. We found that the dashboard significantly increased students' interactions with the assignment, and the questionnaire survey determined that the LA dashboard motivated students to learn during the long vacation period. Considering the previous studies of LA, we estimate that LA-based interventions enhance students' self-regulation skills, which is crucial for learning during long vacation periods. Our study offers a novel approach to implementing LA in K-12 education.

Keywords: Learning analytics, Long vacation period, K-12 education, Learning and Evidence Analytics Framework (LEAF), Evidence-based education

Introduction

Background of the study

It is important to support students' learning during long vacation periods. These extended breaks from school can be a time for students to relax and recharge, but they can also lead to a loss of academic progress. Without the structure and routine of the school day, students may struggle to stay motivated and engaged in their studies (Atteberry & McEachin, 2016; Cooper et al., 1996). By providing tools and resources to support learning during long vacation periods, educators can help ensure that students continue to make progress and are prepared for the next academic term (Cooper, 2003). Such tools can include learning analytics (LA) to track student progress, identify areas where additional support may be needed, and provide access to online resources and educational materials.

In Japan, the summer vacation period is from the end of July to the end of August. During this period, students are often required to complete assignments given by their teachers. Since learning activities during summer vacation are executed in self-regulated learning conditions, learning time and time allocation are completely up to the students. They are expected to proactively work on their assignments while adjusting the time they spend refreshing themselves. However, some students do not study or complete the assignments during this period, which can lead to a loss of academic progress. To prevent this, students must be provided with tools and resources to help them stay motivated and engaged in their studies during the summer vacation period.

To achieve this goal, we adopted an LA approach, which is an emergent field aiming to better understand students and provide intelligence to learners, teachers, and administrators using learning log data (Law & Liang, 2020). Now, the use of learning analytics in K-12 education is still in its early stages compared to its adoption in higher education (Du et al., 2019). This paper offers an example of LA implemented in a junior high Math class during long vacation periods in Japan.

Related works

Analyzing student engagement patterns

Students' engagement patterns have been investigated mainly using psychological questionnaires. Schnitzler et al. (2020) analyzed 397 high school students' profiles using latent profile analysis (LPA) based on three indicators: participation, cognitive engagement, and emotional engagement. Although the first indicator was assessed by the number of hand-raisings in the classroom, the others were measured with survey items. Finally, they discovered five engagement patterns—disengaged, compliant, silent, engaged, and busy—and the significant differences among the learning patterns. A similar approach was taken in a study in the U.S., targeting 1,125 middle school students in a science course to identify engagement profiles and their relationship with science achievement (Bae & DeBusk-Lane, 2019). By applying LPA, the authors discovered five engagement types from the survey of students: moderately engaged or disengaged, behaviorally engaged or disengaged, and disengaged.

Some studies used learning log data to categorize students' engagement patterns. Ebook reading logs were used to categorize students' study patterns at a university in Japan (Akçapinar et al., 2020). They constructed study sequences based on the timestamp they opened the material from the click-stream data and applied hierarchical cluster analysis to the dataset. As a result, they found three different study patterns from the dataset. MOOCs' interaction logs with lectures and assignments were also used to identify learners' study patterns (Boroujeni & Dillenbourg, 2018). The action sequences from learners' log data were extracted and transformed into probability distribution matrices for distance computing. Students were classified into mainly two types based on their learning strategy: fixed approach and changing approach. In the analysis, both hypothesis-driven and data-driven approaches were adopted to assess interventions for motivating students.

We introduced studies that categorized students' engagement patterns in the previous section. However, the goal of LA is not just to analyze learning data, but also to implement effective interventions based on the analysis results (Clow, 2012). Here, we introduce some examples of the interventions based on the analysis results.

In MOOCs, for example, a study conducted in Open University in the UK showed that the students who were predicted to drop out from the course were able to be successfully retained by the email intervention (Rienties et al., 2017). In the study, they predicted the students' probability of dropping out based on the students' interaction logs with the course materials and sent an email to students who were predicted as likely to drop out. As a result, the students who opened the email were more likely to access the course than the students who did not open the email. Another example is a study conducted in real universities (Gutiérrez et al., 2020), in which a dashboard was provided to the teachers, and the ease of use and usefulness of the dashboard was compared with traditional procedures and tools. The study found that the dashboard was more useful and easier for making decisions than traditional tools, especially for young teachers.

Learning analytics implementation at school

Higher education has been using LA to improve services and students' retention rates (Bienkowski et al., 2012). However, adopting LA in school is not easy. It is estimated that it will take two or three years to adopt LA within primary and secondary education (Freeman et al., 2017). There are many barriers to the adoption of LA in the K-12 context: privacy issues are more sensitive (Gunawardena, 2017), resources for supporting analytics implementation are more constrained, and expertise in data analytics is very limited in the K-12 context (Kovanovic et al., 2021). As a result, the number of studies conducted in schools is much fewer than in higher education institutions (Li et al., 2015), with 82.9% of studies focusing on higher education while only 17.1% are in secondary school. As a matter of course, cases from Japanese junior high schools that report LA implementations were very limited in the current state.

Here, we introduce some examples and trends related to adopting LA in school-level education. In Spanish, a research project called PILARES (Smart Learning Analytics Platform to enhance Performance in Secondary Education) was developed for blended learning in secondary school (Sancho et al., 2015) financed by the Spanish government and with the collaboration of the Catalan Ministry of Education. It includes a large Moodle-based LMS called AGORA, used by more than 1,500 schools in Catalonia, which aims at building a LA platform to allow better insight into the learning process through the LMS. In Uruguay, a countrywide LA tool was introduced for secondary education (Macarini et al., 2019). Although they shared several challenges and constraints during its conception and development, they pointed out the feasibility of finding meaningful patterns using the data obtained from the database. They proposed a prototype for tracking the students' scholar trajectory. Although the substantial growth of the LA field itself provided more possibilities to use LA in primary and secondary education, the actual implementations are very limited in the K-12 context (Kovanovic et al., 2021).

Research gap

Other studies on the use of learning analytics in the K-12 context include a study that examined the correlation between log data and performance in digital learning tools in middle school science (Liu et al., 2022), and a study that used smart glasses to visualize student behavior (Holstein et al., 2018). However, to the best of our knowledge, no examples of learning analytics being applied to learning during the summer vacation period were found.

Although the use of technology in class is increasing in the K-12 sector (Staker, 2011), just using technology in classrooms is not enough for what we call LA. According to Clow (2012), LA is defined as a cycle of four phases: learners, data, metrics, and intervention. Closing the loop is crucial for successfully implementing LA (Corrin et al., 2020). In the context of the cycle, this paper explores the effective implementation of LA during long vacation periods at school. We address the following problems in this paper: 1) How can LA be adopted in a Japanese junior high school with the learning log data during the summer vacation period? and 2) How do the LA-based interventions motivate student engagement during the summer vacation? Through these questions, we aim to find a practical approach to implementing LA during long vacation periods in school.

Our approach

This paper consists of two experimental periods. The first is the exploration phase, during which we analyzed 121 students' answer logs and their exam performance after vacation by adopting the K-means clustering method. The exploration phase aims to discover the students' learning patterns during the summer vacation period and the relationship with their academic performance after the vacation. The second is the intervention phase. In this phase, we implemented a real-time dashboard that visualizes students' progress patterns and provides students with insights about their progress during the vacation period. The intervention phase aims to investigate the effect of the dashboard on students' learning while they are out of school.

The paper is structured as follows. In the Methodology section, we first describe our implementation of LA in the target school and the study settings in a junior high Math class in Japan. The results section shows typical students' learning patterns during the summer vacation extracted by an unsupervised clustering analysis and the relationship with their academic performance after the vacation. This part is based on our previous published conference paper (Kuromiya et al., 2021). Then, we introduce the dashboard intervention we implemented in the next winter vacation period and the results of the intervention. Based on the findings, the discussion section proposes developing a system that enables timely intervention during long vacation periods. We also discuss the limitations of our study and future research in this section.

Methods

LEAF platform

Since April 2019, we have been offering an LMS-integrated LA platform called LEAF (Learning Evidence Analytics Framework) to a junior high school in Japan. LEAF offers three online learning tools to teachers - Moodle, BookRoll, and LAViEW (Flanagan & Ogata, 2018).

Moodle is a well-known LMS (Learning Management System) used by more than 100,000 educational institutions worldwide. In this study, we used Moodle as a learning management system. Teachers can set the course and offer students multiple learning materials and tools. Students can access the course and materials from Moodle wherever they are.

BookRoll is an online learning material platform where teachers upload their learning materials for students to read. In BookRoll, teachers can embed quizzes and recommendations (external links) to their learning materials so that students can answer the quizzes and access the external links while reading. BookRoll is accessed from Moodle LMS by LTI (Learning Tools Interoperability) authentication method, so students and

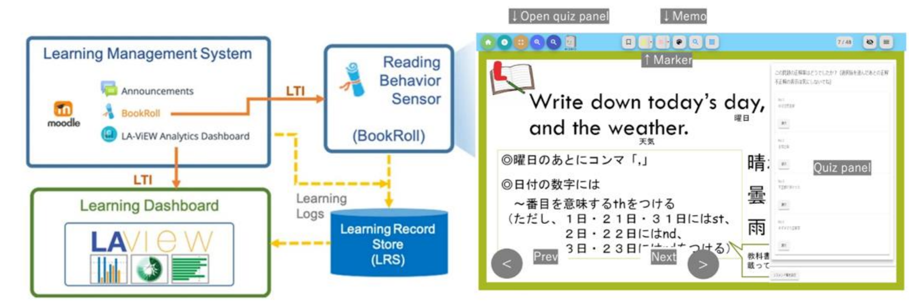

Fig. 1 System components in LEAF

teachers can log in to BookRoll without creating an additional account. It also benefits researchers because we can retrieve students' information by their Moodle IDs.

The last component, LAViEW is a dashboard that visualizes the learner-content interactions in BookRoll. Teachers can see students' highlights, memos, and time spent on each page of the learning materials. LAViEW is also accessed from Moodle by the LTI authentication, and teachers can see all their students' activities in BookRoll. Students can also see their own activities in BookRoll from LAViEW, but they cannot see other students' activities with their names.

Figure 1 shows the workflow of the LEAF platform. First, teachers upload learning materials to BookRoll. Then, students access the materials from Moodle and read them. While reading, students can highlight text, write memos, and answer quizzes. All the interactions are recorded in the database. Finally, teachers can see the students' interactions from LAViEW. This paper focuses on the interactions between students and quizzes in BookRoll.

Experimental context

Monitoring summer vacation assignments with LEAF

In this paper, we focus on a specific use case scenario of the LEAF platform in the summer vacation period in Math class. We targeted three classes containing 121 students in a junior high third-grade Math course. Before the summer vacation, a teacher uploaded an assignment with 49 Math questions to BookRoll (see Figure 2). The assignment consists of one question per page, and a straightforward reflection widget is implemented on each page. The widget has options representing three different understanding levels, perfect, understood, or not well understood, as seen in Figure 2 on the right-hand side. During summer vacation, students must report their understanding levels on BookRoll every time they finish solving a problem and submit a paper that contains the answers and working

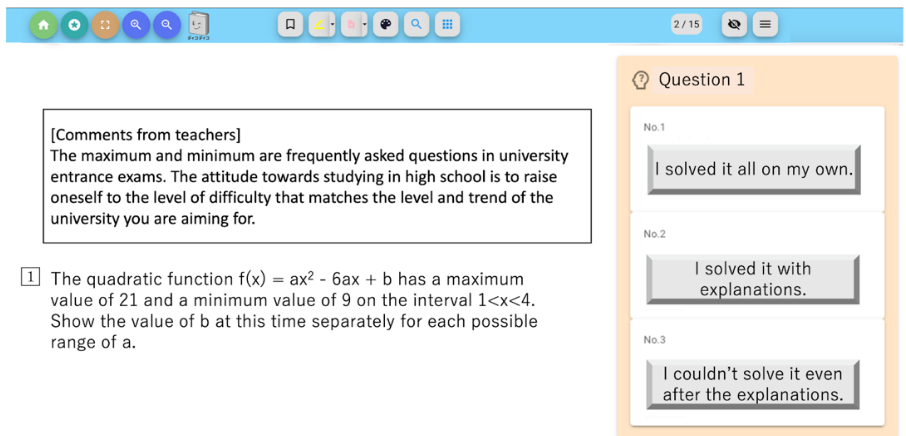

Fig. 2 Summer vacation assignment in BookRoll

formulas to all the questions. The summer vacation was from July 22 to August 21, 2019. To measure students' performance, the examination was conducted on August 23 after the summer vacation. We used this score to investigate the relationship between students' behavior and performance.

Preparing the dashboard for long vacation period

Based on the results of the first experiment, we implemented a real-time dashboard that visualizes students' progress patterns and gives students insights into their progress during the vacation period. In the winter vacation period next year, we offered a dashboard that visualizes students' progress compared to their peers. The dashboard was implemented in LAViEW and students can access it from Moodle LMS.

We developed this dashboard to automate the analysis of comprehension check data and provide more frequent feedback to students and teachers. With a one-time analysis, providing real-time feedback to students is impossible. Especially during long periods when students do not come to school, the presence or absence of feedback from the system is crucial. By transforming data analysis into a dashboard, even during extended breaks at home, students are aware of the response status of their peers. This is particularly useful for learning during extended absences from school.

The interface of the dashboard is shown in Figure 3. It consists of three different components: (A) content selector, (B) page-wise analysis panel, and (C) student-wise analysis panel. In the content selector (labeled "A"), students are required to select the assignment content. We implemented this component because students usually have multiple assignments during the long vacation period. In this period, we focus on one assignment, but students can also visualize their progress on other assignments.

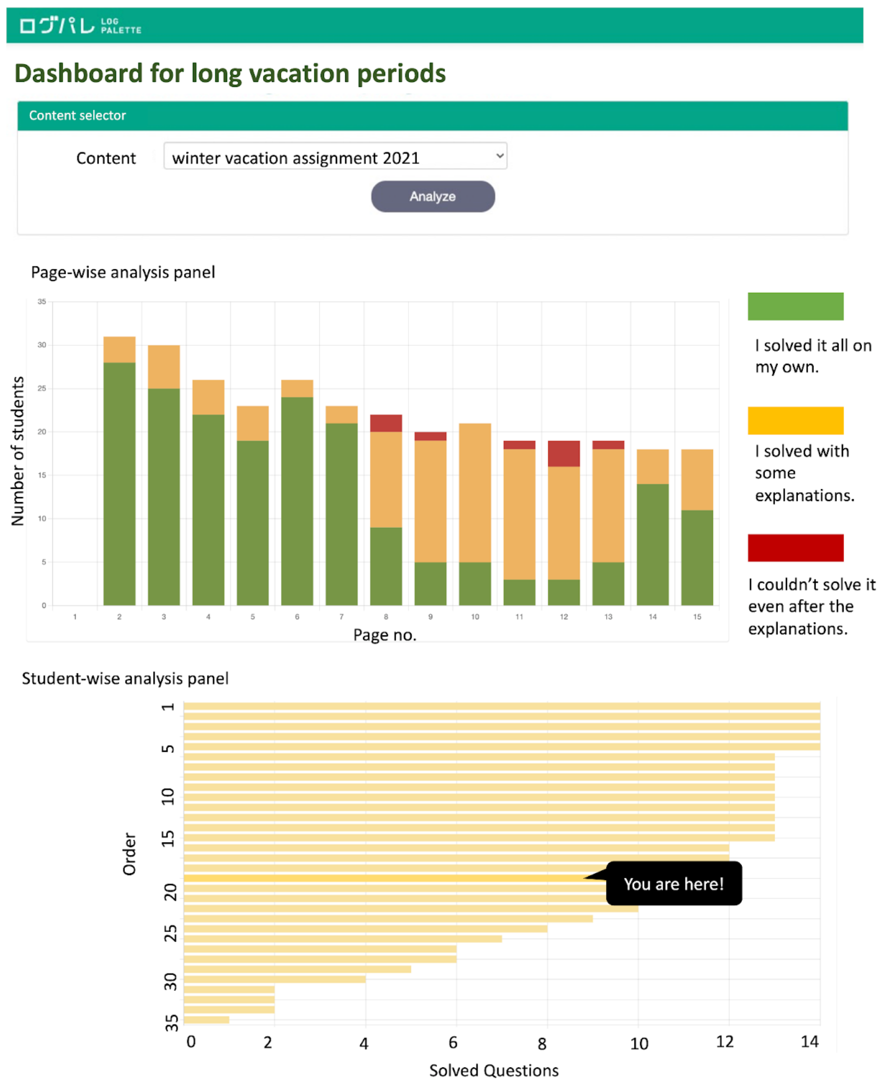

Fig. 3 Dashboard for long vacation periods

Once a user selects the content, two charts will be available: a page-wise analysis panel and a student-wise analysis panel. In the page-wise analysis panel (labeled "B"), students can see how many students tried to solve that question and how students responded to the comprehensive check survey for each page. In the student-wise analysis panel (labeled "C"), students can check their progress on the assignments relative to the whole class. To prevent students from comparing their progress with other students, we did not show the names of other students in the dashboard. Instead, we showed the number of students who solved the question before the target student.

Data collection and analysis

Exploration phase

To measure students' engagement during the summer vacation period, we used the students' reflection logs on BookRoll. Learners used BookRoll to solve problems and record their reflections on their understanding, while we analyzed their learning logs. By analyzing the reflection logs, we can determine which students solved which question on what day. We targeted ebook logs from July 17 to August 22, 2020 for the log analysis. Overall, we extracted 6,250 answers from the students. We excluded 1,172 answers from page 1 where the content was just a description of the assignment. Then, we excluded 132 answers after the exam. As a result, we analyzed 4,936 student answers. Sixteen students were excluded throughout this process because they had no reflection data. Three other students were excluded because they had no score data. We also excluded one student who sent too many answers in only a few minutes. In total, we obtained 101 students for the analysis.

As the summer vacation period was over in 37 days, we separated the answers into two periods: logs before August 4 were from the first half and those from after that date were from the second half. Duplicate answers to the same question in the same period were excluded. Finally, standard K-means was conducted based on each student's interactions during the first-half and second-half periods. The Elbow plot determined the number of clusters. Here, we decided four was the optimal number of clusters because it was the lowest BIC value in the plot.

After the categorization step, we compared each cluster's average exam score after the vacation period. ANOVA and post hoc testing were adopted using statistical testing software, JASP (Love et al., 2019).

Intervention phase

Based on the results from the exploration phase, we offered the dashboard to the students in junior high first grade (N = 114) during the winter vacation period from 24 December 2021 to 10 January 2022. Before entering winter vacation, teachers taught students the basic usage of the dashboard. Whether or not students use the dashboard during winter vacation is left up to the students themselves. About half of the students used the dashboard during the winter vacation period (N = 54), while the others did not (N = 60). We divided the students into two groups: the dashboard-use group and the nonuse group. The dashboard-use group comprised the students who accessed the dashboard at least once during the winter vacation period. The non-use group comprised the students who did not access the dashboard at all during the winter vacation period. We compared the engagement and the post-exam performance between the two groups.

Table 1 Survey items and response options

	Question	Options
1	How useful was the page-wise analysis for studying during vacation?	1-5 Likert scale
2	How useful was the student-wise analysis for studying during vacation?	1-5 Likert scale
3	Please write anything you want to say.	Free writing

After the winter vacation, we conducted a questionnaire survey to the students who used the dashboard. The questionnaire consisted of five items and two free-response questions. The items were about the usability of the dashboard, and the free-response questions were about the usefulness of the dashboard and the difficulty of understanding the dashboard. The questionnaire was conducted in Japanese, and the students answered in Japanese. Table 1 showed the descriptions of the items and available response options in the survey. We gathered responses from dashboard users and summarized how many students answered in what way for each item in histogram format. For free-response questions, we translated the entire text of the respondent's response into English and posted it.

Results

Exploring students' engagement patterns

Figure 4 (left) shows the scatter plot of the clustering results. The horizontal axis represents the number of answers in the first-half period, and the vertical axis represents the number of answers in the second-half period. The scatter plot was divided into four clusters. We labeled each cluster as follows: early engagement group (N = 19), high-engagement group (N = 14), late engagement group (N = 37), and low-engagement group (N = 32). Table 2 shows the descriptive statistics of each cluster. Although the number of completed quizzes

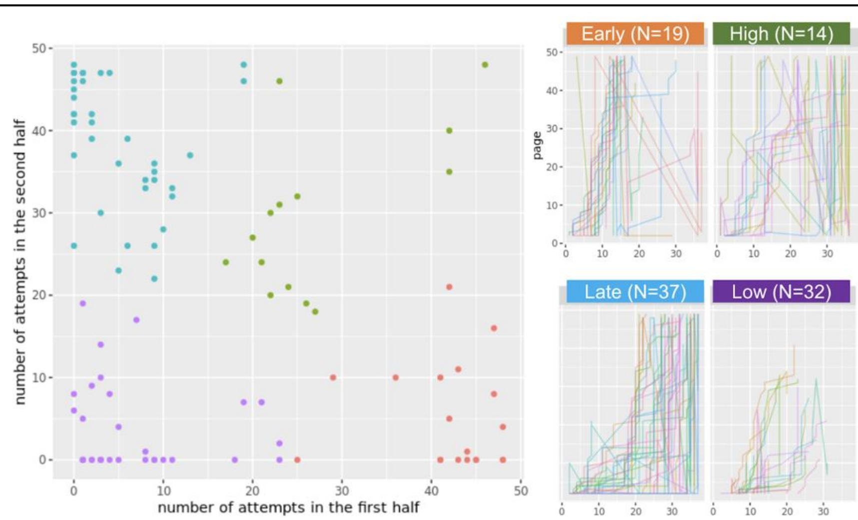

Fig. 4 Clustering by student learning patterns

Table 2 Characteristics of each cluster

Cluster		# attempts	Reviewed quizzes	Active week	Active day
High	(N = 14)	58.3 (21.1)	11.8 (13.0)	4.2 (0.7)	10.0 (3.0)
Early	(N = 19)	48.3 (9.9)	3.9 (6.6)	2.8 (1.0)	7.5 (3.8)
Late	(N = 37)	48.0 (19.0)	5.0 (10.9)	3.2 (1.4)	6.4 (3.2)
Low	(N = 32)	10.7 (8.8)	0.5 (1.6)	1.9 (0.9)	3.3 (2.1)

Note. It shows the means of each indicator. The numbers in the parentheses are standard deviations.

were almost the same among the early, high and late engagement groups, there were more reviewed quizzes in the high-engagement group than in the other groups. Conversely, there were fewer completed and reviewed quizzes in the low-engagement group than in the other groups.

Figure 4 (right) shows actual students' progress patterns for each cluster. The horizontal axis represents the day of the vacation, and the vertical axis represents the page of the question they answered. The students in the early engagement group finished the assignments within the first-half period. In contrast, students in the late engagement group made little progress by the second half of the vacation. Students in the high-engagement group can be categorized as two subgroups: some students finished the assignment once half the vacation was over and solved the questions again at the end of the vacation. In contrast, others continuously worked throughout the entire vacation. Conversely, students in the low-engagement group could not finish the assignment.

Additionally, we investigated the relationship between their solving patterns and performance. Figure 5 shows the descriptive plot of students' exam scores among four

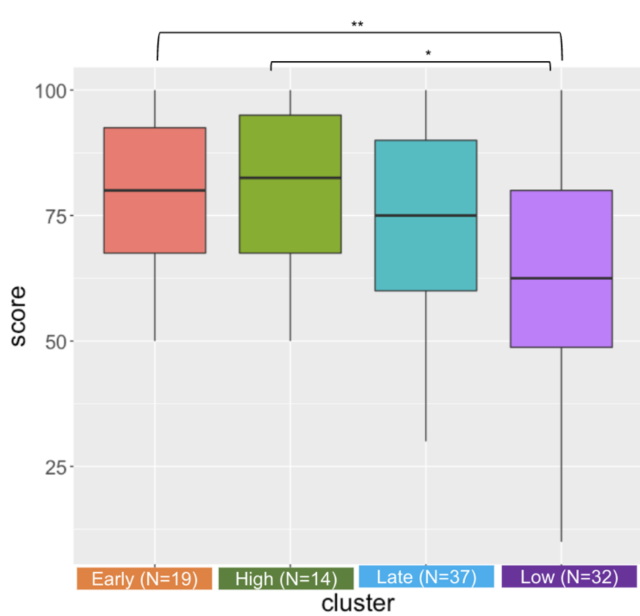

Fig. 5 Exam scores after vacation for each cluster

Table 3 Post hoc comparison results between every cluster combination

		Difference	SE	T	P
Early	High	0.075	6.618	0.011	1.000
	Late	7.411	5.303	1.398	0.504
	Low	18.914	5.442	3.476	0.004**
High	Late	7.336	5.896	1.244	0.600
	Low	18.839	6.021	3.129	0.012*
	Low	11.503	4.536	2.536	0.061

Note. p-values adjusted for comparing a family of four *p < .05, **p < .01.

clusters. The error bar stands for the standard error of each cluster. Students in the high-engagement group obtained the highest score (M: 80.8, SD: 14.5), while students in the early engagement group obtained the second highest score (M: 80.7, SD: 15.0). Students in the late engagement group were the third ranked (M: 73.4, SD: 19.9) and students in the low-engagement group received the lowest score (M: 61.9, SD: 21.0). Finally, we conducted an ANOVA to test the difference in exam scores by clusters. Before adopting ANOVA, we checked the homogeneity of the data, and the result was not significant ($p = .18$). The result of ANOVA was significant ($p = .001$), so we conducted a post hoc comparison between clusters (see Table 3). As a result, we obtained significant differences between clusters one and four and between clusters two and four. The p-values were adjusted for multiple comparisons.

Results of the dashboard intervention

In the second phase experiment, we expected that the dashboard would motivate students to increase their engagement during the winter vacation period. In the first experiment, we separated the answers into two periods: logs before January 2 were from the first half, and after that, they were from the second half. Duplicate answers to the same question in the same period were excluded in this process.

Figure 6 shows the results of the dashboard intervention. We compared the number of completed and reviewed quizzes and post exam scores between the dashboard-use group and nonuse group. The dashboard-use group completed and reviewed more quizzes than the nonuse group. However, there was no difference in the exam scores between the two groups. We conducted statistical tests to investigate the difference between the two groups (see Table 4). For the engagement part, we found significant differences between the two groups in the first and second half periods ($p < .01$). However, for the exam score part, we found no significant difference between the two groups ($p = .27$). Looking at the effect size (Cohen's d), the effect size of the engagement part was large ($d = 0.8$) but the effect size of the exam score part was small ($d = -0.2$).

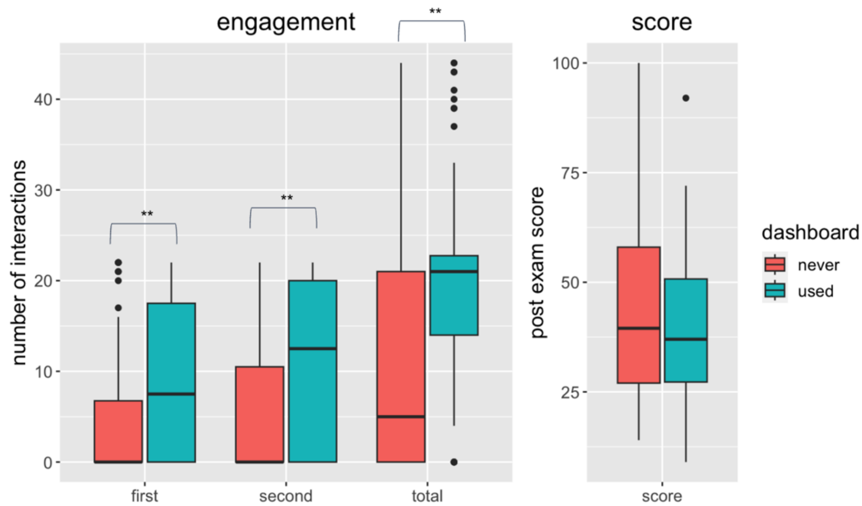

Fig. 6 Effect of the dashboard on student engagement and performance

Table 4 Difference in engagement and performance between dashboard use and nonuse groups

	T	DF	P	Cohen's d
Engagement in the first half	2.869	103.780	0.005**	0.541
Engagement in the second half	3.794	108.020	<0.001**	0.713
Engagement in total	4.764	107.880	<0.001**	0.896
Post exam score	-1.099	111.757	0.274	−0.205

Note. Welch's t-test *p < .05, **p < .01.

To investigate the effect of the dashboard from the students' perspective, we conducted a questionnaire survey with the students who used the dashboard. The response rate was 74% (N = 40). Figure 7 shows the result of the questionnaire about the dashboard. The horizontal axis represents the items in the questionnaire and the vertical axis represents the number of students who answered each option.

Most of the students indicated that the dashboard was useful for their learning. The mean score of the item was 4.2 (SD: 0.6) for the page-wise analysis panel, and 4.0 (SD: 0.9) for the student-wise analysis panel. The difference between the two panels was not significant (p = .25). However, the difference from a neutral score (3.0) was significant for both panels (p < .01). In addition, we included a free-response question about the usefulness of the dashboard. Table 5 shows the result of the free-response question in the questionnaire. There were 6 positive comments and 2 negative comments. The positive comments were about the usefulness of the dashboard for student learning, versus the negative comments that were about the difficulty of using the dashboard to understand their progress.

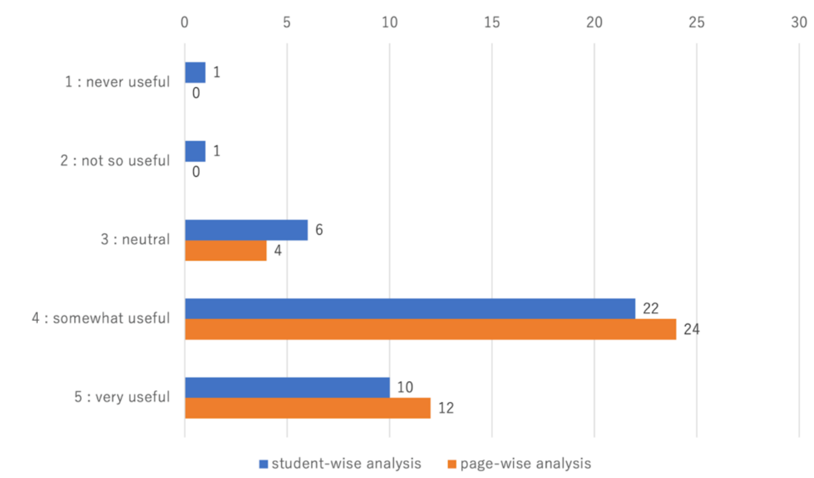

Fig. 7 Usability survey results

Table 5 Student responses to free-text items

Student No.	Comment	Polarity
S1	I was able to check everyone's progress during winter break in class and it motivated me	Positive
S2	I think it's convenient to be able to compare myself with others.	Positive
S3	I was able to solve as many problems as possible in a short amount of time.	Positive
S4	I want a mark on the graph of the progress by page to show which level I was at.	Positive
S5	It was helpful when I didn't know how far I had gone, the graph was very useful.	Positive
S6	Actually, I used it during winter break and it was very convenient to know how much others were doing and how difficult they perceived the problems to be. I was able to make effective use of it. So, I hope you will continue to do it in the future.	Positive
S7	There were times when I solved problems but they weren't reflected, so I want that to be improved.	Negative
S8	Since the graph in the survey is different from the actual graph, I can't evaluate my study patterns, so there's nothing I can do about it.	Negative

Discussions

Principal findings

Until here, we explored the students' progress patterns during the summer vacation period and the effective implementation of the dashboard during long vacation periods in school. Throughout the two experiments, we found that students' progress patterns were able to be categorized into four types: early engagement, late engagement, high engagement, and low engagement. Plus, the early and high-engagement group got significantly higher scores than the low-engagement group, which means that motivating students' engagement during the long vacation period is important for their learning and academic performance.

In the second study, we implemented a dashboard that visualizes students' progress status in real time. We expected that the dashboard would motivate students to increase their engagement during the long vacation period. As a result, we found that the dashboard significantly increased students' engagement with the assignment, but did not affect their exam performance. The result of the questionnaire supported the fact that the dashboard supported students' learning during the long vacation period. This study gives us a clue that the dashboard, which visualizes students' progress status in real time, is effective for motivating students' engagement during long vacations. However, there is room for improvement of the dashboard.

Comparison with previous studies

Students engagement patterns and their performance

Online engagement during an assignment is widely considered to affect students' academic performance. For instance, a study investigated predictors of students' weekly achievement and found that time spent on homework and labs was more strongly related to their performance than time spent on discussion boards or books (DeBoer & Breslow, 2014). Another research insisted that assignment features, such as average submission lead time and total quiz submission, play important roles in the dropout prediction model in MOOCs (Gardner & Brooks, 2018). These studies indicated that engagement with the assignment is important for students' performance. However, unlike our study, the learning process was not often considered in the performance prediction before. As far as we know, only one study investigated the relationship between students' engagement patterns and their performance in MOOCs (Kizilcec et al., 2017), and they found that early engagement was a good predictor of students' performance. Our study is consistent with their findings and extends the results to the secondary education context.

Based on the previous studies in the similar context, we hypothesize that early engagement reflects the high self-regulation skills of students (Yang et al., 2022, 2023). Self-regulation skills are the ability to control one's behavior, emotions, and thoughts (Zimmerman, 2002). We can say that students who finished the assignments early have high self-regulation skills because they can set a goal and deadline and make a plan to get there in time. It would prove the student's high self-regulation skills, which lead to a high examination score after the vacation.

Until now, many studies indicated that students' self-regulation skills contributed to high performance in MOOCs. A path analysis between self-regulated learning and learning achievement from a large dataset from Korean Cyber University revealed that self-regulated learning was positively correlated with participation, which was positively correlated to learning achievement (Im & Kang, 2019). There, the self-direction skill indirectly contributes to the learning performance. Another study is a case study of a MOOC course in which the students' self-regulation skills were measured by a survey investigating the relationship between self-regulation skills and learning persistence (Joo et al., 2014). They found that self-regulation skills were positively correlated to the learning persistence. These studies would support our hypothesis that self-regulation skill is a key factor for students' learning during the long vacation period.

Effective interventions during long vacation periods

As we showed in this study, the dashboard that visualizes students' progress status in real time is effective for motivating students' engagement during long vacations. This phenomenon is known as "social comparison." A previous study (Joshi et al., 2023) showed that students have different preferences for social comparison and the effect of social comparison depends on the type of social comparison. In our study, we visualized students' progress status in real time, which was effective for motivating students to remain engaged during the long vacation period. However, preferences may vary across students; understanding this variation would be a good future direction for our research.

Also, effective usage of the dashboard was reported in the previous studies of LA. For example, a large-scale study showed that complementing the dashboard with personalized feedback increased student performance (Pardo et al., 2019). The authors claimed that sending a personalized email to students who were predicted to drop out would be effective for retaining students' performance (Pardo et al., 2018). Another conducted in a similar context, a high school Math course in Japan, showed that personalized recommender systems with explanations improved student performance (Dai et al., 2024; Majumdar et al., 2023). These studies provide us clues for improving the dashboard. In our study, we implemented a dashboard that visualizes students' progress status in real time. However, we did not send any personalized messages to students. Sending personalized messages to students who are predicted to have a low level of engagement in the second half of the vacation period would be effective for motivating students to stay involved. This is another good future direction for our research.

Limitations

One limitation of this study lies in the generalizability of the results. We conducted the experiment in a single school in Japan. Although we found the dashboard to be effective in the target school, it is not clear if the dashboard is effective in other schools. To solve this problem, we plan to conduct the experiment in other schools. As we have a large-scale research network including other countries that use the LEAF platform (Ogata et al., 2022), using the dashboard in other schools would increase the generalizability of the results.

Another limitation lies in the study's clustering features. In this study, we used the number of answers in the first and second half of the vacation as the clustering features. These features give us advantages when we visualize the results of the clustering. However, in that context, we were not able to consider the rich time-series information for pattern clustering because they are aggregated features of daily engagement of students. Advanced clustering methods such as time-series clustering (Aghabozorgi et al., 2015) may enable us to treat the rich time-series information of students' answer data. In a study in which time-series clustering was applied to the learning log data in MOOCs (Hung et al., 2015), the time-series features were extracted from the log data and time-series clustering was applied to the dataset. By examining the detailed interactions between learners and the LEAF system, we may be able to obtain clearer results on what kind of behavior affects learning performance.

Conclusion

In this paper, we offered an example of a LA platform implemented in an actual junior high Math class in Japan. In particular, we introduce a case of a summer vacation assignment. The summer vacation period is a temporal remote learning period in face-to-face classrooms; thus, the use of technology is easier than in normal face-to-face learning period. The two experiments were conducted in the summer and winter vacation periods. In the first experiment, we analyzed the students' reflection logs on BookRoll and categorized students' progress patterns into four types of engagement: early, late, high, and low. We found that the early and high-engagement group obtained significantly higher scores than the low-engagement group. In the second experiment, we implemented a dashboard that visualizes students' progress status in real time. We found that the dashboard significantly increased students' interactions with the assignment. However, it did not affect students' exam performance. The result of the questionnaire indicated that the dashboard supported students' learning during the long vacation period. Our case is the first model case of how to implement LA in secondary school in Japan. We are confident that the use of LA will increase in the secondary education sector as well as in higher education institutions.

Abbreviations

LA: Learning Analytics; LEAF: Learning and Evidence Analytics Framework; LPA: Latent Profile Analysis; MOOCs: Massive Open Online Courses; LMS: Learning Management System; LTI: Learning Tools Interoperability.

Acknowledgements

We extend our heartfelt gratitude to the school teachers whose participation and insights were crucial in informing our research, highlighting the invaluable connection between academic theories and classroom realities. Similarly, we thank our technical staff for their expertise and meticulous attention to detail, which have been indispensable to our project's success.

Authors' contributions
HK developed the system, performed data analysis and drafted the initial manuscript. RM co-conducted the experiment and edited the manuscript. IH provided the support to the data analysis and edited the manuscript. HO was responsible for funding, experiment, and the whole research process. All authors read and approved the final manuscript.

Authors' information
Hiroyuki Kuromiya is a doctoral student in the Graduate School of Informatics, Kyoto University. His scholarly contributions focus on leveraging learning analytics to foster evidence-based education, understanding learner engagement and perception through ebook-based activities, especially noted during the COVID-19 pandemic.

Rwitajit Majumdar is an associate professor at the Research and Educational Institute for Semiconductors and Informatics, and the Graduate School of Social and Cultural Sciences, Division of Instructional System Studies at Kumamoto University. His research focuses on learning analytics and human-data interactions in educational platforms.

Izumi Horikoshi is an assistant professor at the Academic Center for Computing and Media Studies and the Graduate School of Informatics, Kyoto University. Her research interests include learning analytics and classroom visualization for formative assessment and reflection.

Hiroaki Ogata is a full professor at the Academic Center for Computing and Media Studies, Kyoto University. His academic expertise and interests span a variety of areas within educational technology, including educational data science, learning analytics, mobile and ubiquitous learning, computer-supported collaborative learning (CSCL), and knowledge awareness.

Funding
This work is partially funded by the following research grants.

Dr. Hiroyuki Kuromiya
JST ACT-X JPM- JAX20AA
JSPS KAKENHI Grant-in-Aid for JSPS Fellows Number 21J14514

Dr. Rwitajit Majumdar
JSPS KAKENHI Grant-in-Aid for Early-Career Scientists 20K20131
JSPS KAKENHI Grant-in-Aid for Scientific Research (B) 20H01722

Prof. Hiroaki Ogata
JSPS KAKENHI Grant-in-Aid for Scientific Research (S) 16H06304
NEDO Special Innovation Program on AI and Big Data JPNP18013
NEDO JPNP20006

Availability of data and materials
Not applicable.

Declarations

Competing interests
The authors declare that they have no competing interests.

Author details
[1] Graduate School of Social Informatics, Kyoto University, Japan
[2] Research and Educational Institute for Semiconductors and Informatics, Kumamoto University, Japan
[3] Academic Center for Computing and Media Studies, Kyoto University, Japan

Received: 20 October 2023 Accepted: 8 March 2024
Published online: 20 March 2024

References

Aghabozorgi, S., Shirkhorshidi, A. S., & Wah, T. Y. (2015). Time-series clustering – A decade review. *Information Systems*, *53*, 16–38. https://doi.org/10.1016/j.is.2015.04.007

Akçapinar, G., Hasnine, N. M., Majumdar, R., Chen, A. M.-R., Flaganan, B., & Ogata, H. (2020). Exploring temporal study patterns in ebook-based learning. In H. J. So et al. (Eds.), *Proceedings of the 28th International Conference on Computers in Education* (pp. 342–247). Asia-Pacific Society for Computers in Education.

Atteberry, A., & McEachin, A. (2016). School's out: Summer learning loss across grade levels and school contexts in the United States today. In K. Alexander, S. Pitcock & M. Boulay (Eds.), *Summer learning and summer learning loss* (pp. 35–54). Teachers College Press.

Bae, C. L., & DeBusk-Lane, M. (2019). Middle school engagement profiles: Implications for motivation and achievement in science. *Learning and Individual Differences*, *74*, 101753. https://doi.org/10.1016/j.lindif.2019.101753

Bienkowski, M., Feng, M., & Means, B. (2012). *Enhancing teaching and learning through educational data mining and learning analytics: An issue brief*. Office of Educational Technology, US Department of Education.

Boroujeni, M. S., & Dillenbourg, P. (2018). Discovery and temporal analysis of latent study patterns in MOOC interaction sequences. In A. Pardo, K. Bartimote-Aufflick & G. Lynch (Eds.), *Proceedings of the 8th International Conference on Learning Analytics and Knowledge* (pp. 206–215). Association for Computing Machinery. https://doi.org/10.1145/3170358.3170388

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. In S. Buckingham Shum, D. Gasevic & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134–138). Association for Computing Machinery. https://doi.org/10.1145/2330601.2330636

Cooper, H. (2003). *Summer learning loss: The problem and some solutions*. ERIC Digest.

Cooper, H., Nye, B., Charlton, K., Lindsay, J., & Greathouse, S. (1996). The effects of summer vacation on achievement test scores: A narrative and meta-analytic review. *Review of Educational Research*, *66*(3), 227–268.

Corrin, L., Scheffel, M., & Gašević, D. (2020). Learning analytics: Pathways to impact. *Australasian Journal of Educational Technology*, *36*(6), 1–6. https://doi.org/10.14742/ajet.6853

Dai, Y., Takami, K., Flanagan, B., & Ogata, H. (2024). Beyond recommendation acceptance: Explanation's learning effects in a math recommender system. *Research and Practice in Technology Enhanced Learning*, *19*, 020. https://doi.org/10.58459/rptel.2024.19020

DeBoer, J., & Breslow, L. (2014). Tracking progress: Predictors of students' weekly achievement during a circuits and electronics MOOC. In M. Sahami (Ed.), *Proceedings of the First ACM Conference on Learning@ Scale Conference* (pp. 169–170). Association for Computing Machinery. https://doi.org/10.1145/2556325.2567863

Du, X., Yang, J., Shelton, B. E., Hung, J.-L., & Zhang, M. (2019). A systematic meta-review and analysis of learning analytics research. *Behaviour & Information Technology*, *40*(1), 49–62. https://doi.org/10.1080/0144929X.2019.1669712

Flanagan, B., & Ogata, H. (2018). Learning analytics platform in higher education in Japan. *Knowledge Management & E-Learning: An International Journal*, *10*(4), 469–484. https://doi.org/10.34105/j.kmel.2018.10.029

Freeman, A., Adams Becker, S., Cummins, M., Davis, A., & Hall Giesinger, C. (2017). *NMC/CoSN horizon report: 2017 K-12 Edition*. The New Media Consortium.

Gardner, J., & Brooks, C. (2018). Student success prediction in MOOCs. *User Modeling and User-Adapted Interaction*, *28*, 127–203. https://doi.org/10.1007/s11257-018-9203-z

Gunawardena, A. (2017). Brief survey of analytics in K12 and higher education. *International Journal on Innovations in Online Education*, *1*(1).

Gutiérrez, F., Seipp, K., Ochoa, X., Chiluiza, K., De Laet, T., & Verbert, K. (2020). LADA: A learning analytics dashboard for academic advising. *Computers in Human Behavior*, *107*, 105826. https://doi.org/10.1016/j.chb.2018.12.004

Holstein, K., McLaren, B. M., & Aleven, V. (2018). Student learning benefits of a mixed-reality teacher awareness tool in AI-enhanced classrooms. In C. Penstein Rosé et al. (Eds.), *Artificial Intelligence in Education. AIED 2018. Lecture Notes in Computer Science, vol 10947* (pp. 154–168). Springer, Cham. https://doi.org/10.1007/978-3-319-93843-1_12

Hung, J.-L., Wang, M. C., Wang, S., Abdelrasoul, M., Li, Y., & He, W. (2015). Identifying at-risk students for early interventions-A time-series clustering approach. *IEEE Transactions on Emerging Topics in Computing*, *5*(1), 45–55. https://doi.org/10.1109/TETC.2015.2504239

Im, T., & Kang, M. (2019). Structural relationships of factors which impact on learner achievement in online learning environment. *International Review of Research in Open and Distributed Learning*, *20*(1). https://doi.org/10.19173/irrodl.v20i1.4012

Joo, Y. J., Joung, S., & Kim, J. (2014). Structural relationships among self-regulated learning, learning flow, satisfaction, and learning persistence in cyber universities. *Interactive Learning Environments*, *22*(6), 752–770. https://doi.org/10.1080/10494820.2012.745421

Joshi, A., Molenkamp, B., & Sosnovsky, S. (2023). Student perception of social comparison in technology enhanced learning. In O. Viberg, I. Jivet, P. Muñoz-Merino, M. Perifanou & T. Papathoma (Eds.), *Responsive and sustainable educational futures* (pp. 118–132). Springer, Cham. https://doi.org/10.1007/978-3-031-42682-7_9

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in massive open online courses. *Computers & Education*, *104*, 18–33. https://doi.org/10.1016/j.compedu.2016.10.001

Kovanovic, V., Mazziotti, C., & Lodge, J. (2021). Learning analytics for primary and secondary schools. *Journal of Learning Analytics*, *8*(2), 1–5. https://doi.org/10.18608/jla.2021.7543

Kuromiya, H., Majumdar, R., & Ogata, H. (2021). Mining students' engagement pattern in summer vacation assignment. In M. M. T. Rodrigo et al. (Eds.), *Proceedings of the 29th International Conference on Computers in Education, Vol. 1* (pp. 559–568). Asia-Pacific Society for Computers in Education.

Law, N., & Liang, L. (2020). A multilevel framework and method for learning analytics integrated learning design. *Journal of Learning Analytics*, *7*(3), 98–117. https://doi.org/10.18608/jla.2020.73.8

Li, K. C., Lam, H. K., & Lam, S. S. (2015). A review of learning analytics in educational research. In J. Lam, K. Ng, S. Cheung, T. Wong, K. Li & F. Wang (Eds.), *Technology in Education. Technology-Mediated Proactive Learning. ICTE*

2015. Communications in Computer and Information Science, vol 559 (pp. 173–184). Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-662-48978-9_17

Liu, M., Cai, Y., Han, S., & Shao, P. (2022). Understanding student navigation patterns in game-based learning. *Journal of Learning Analytics*, *9*(3), 50–74. https://doi.org/10.18608/jla.2022.7637

Love, J., Selker, R., Marsman, M., Jamil, T., Dropmann, D., Verhagen, J., Ly, A., Gronau, Q. F., Smíšek, M., Epskamp, S., Matzke, D., Wild, A., Knight, P., Rouder, J. N., Morey, R. D., & Wagenmakers, E.-J. (2019). JASP graphical statistical software for common statistical designs. *Journal of Statistical Software*, *88*(2), 1–17. https://doi.org/10.18637/jss.v088.i02

Macarini, L. A., Cechinel, C., Santos, H. L. d., Ochoa, X., Rod, V., Alonso, G. E., Casas, A. P., & Díaz, P. (2019). Challenges on implementing learning analytics over countrywide K-12 data. In S. Hsiao, J. Cunningham, K. McCarthy & G. Lynch (Eds.), *Proceedings of the 9th International Conference on Learning Analytics & Knowledge* (pp. 441–445). Association for Computing Machinery. https://doi.org/10.1145/3303772.3303819

Majumdar, R., Kyosuke, T., & Ogata, H. (2023). Learning with explainable ai-recommendations at school: Extracting patterns of self-directed learning from learning logs. In M. Chang, N.-S. Chen, R. Kuo, G. Rudolph, D. G. Sampson & A. Tlili (Eds.), *Proceedings of the 23rd IEEE International Conference on Advanced Learning Technologies* (pp. 245–249). IEEE. https://doi.org/10.1109/ICALT58122.2023.00078

Ogata, H., Majumdar, R., Yang, S. J., & Warriem, J. M. (2022). Learning and evidence analytics framework (LEAF): Research and practice in international collaboration. *Information and Technology in Education and Learning*, *2*(1), 001. https://doi.org/10.12937/itel.2.1.Inv.p001

Pardo, A., Bartimote, K., Buckingham Shum, S., Dawson, S., Gao, J., Gašević, D., Leichtweis, S., Liu, D., Martínez-Maldonado, R., Mirriahi, N., Moskal, A. C. M., Schulte, J., Siemens, G., & Vigentini, L. (2018). OnTask: Delivering data-informed, personalized learning support actions. *Journal of Learning Analytics*, *5*(3), 235–249. https://doi.org/10.18608/jla.2018.53.15

Pardo, A., Jovanovic, J., Dawson, S., Gašević, D., & Mirriahi, N. (2019). Using learning analytics to scale the provision of personalised feedback. *British Journal of Educational Technology*, *50*(1), 128–138. https://doi.org/10.1111/bjet.12592

Rienties, B., Cross, S., & Zdrahal, Z. (2017). Implementing a learning analytics intervention and evaluation framework: What works? In B. K. Daniel (Ed.), *Big data and learning analytics in higher education* (pp. 147–166). Springer. https://doi.org/10.1007/978-3-319-06520-5_10

Sancho, M.-R., Camba Batate, A., & Sabate, F. (2015). Contextualizing learning analytics for secondary schools at micro level. In A. Pester, V. Alexandrov, J. J. Noguez & L. J. Neri (Eds.), *Proceedings of 2015 International Conference on Interactive Collaborative and Blended Learning* (pp. 70–75). IEEE. https://doi.org/10.1109/ICBL.2015.7387638

Schnitzler, K., Holzberger, D., & Seidel, T. (2020). All better than being disengaged: Student engagement patterns and their relations to academic self-concept and achievement. *European Journal of Psychology of Education*, *36*(3), 627–652. https://doi.org/10.1007/s10212-020-00500-6

Staker, H. (2011). *The rise of K-12 blended learning*. Innosight Institute.

Yang, Y., Li, H., Majumdar, R., & Ogata, H. (2023). GOAL system for online self-direction practice: Exploring students' behavioral patterns and the impact on academic achievement in the high school EFL context. *Journal of Computers in Education*. https://doi.org/10.1007/s40692-023-00272-0

Yang, Y., Majumdar, R., Li, H., Flanagan, B., & Ogata, H. (2022). Design of a learning dashboard to enhance reading outcomes and self-directed learning behaviors in out-of-class extensive reading. *Interactive Learning Environments*, 1–18. https://doi.org/10.1080/10494820.2022.2101126

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Publisher's Note