Finding Traces of High and Low Achievers by Analyzing Undergraduates' E-book Logs

Misato Oi[1], Masanori Yamada[1], Fumiya Okubo[1], Atsushi Shimada[1], Hiroaki Ogata[2]

[1]Kyushu University, Fukuoka, Japan
[2]Kyoto University, Kyoto, Japan

oimisato@gmail.com
mark@mark-lab.net
fokubo@artsci.kyushu-u.ac.jp
atsushi@limu.ait.kyushu-u.ac.jp
hiroaki.ogata@gmail.com

Abstract: We investigated the learning behavior of undergraduates with e-book logs. E-book logs from 99 undergraduates taking an information science course were collected. First, we analyzed differences between nine high-achieving students and three low-achieving students. A log recorded before a class session in which the same e-book was used as a textbook was considered a preview log, and one recorded after a class session was considered a review log. The analysis of preview frequency indicates that the low achievers did not perform the previews, but many high achievers frequently did. The review frequency demonstrates that regardless of high and low achievements, students performed reviews. We added the logs of six relatively low achievers and analyzed more details of the preview logs of high and low achievers. The number of page flips and durations of preview logs revealed that relatively low achievers tried to perform previews, but they gave the endeavor up easily.

Keywords: E-book logs, Preview, Review.

1 Introduction

According to Daniel and Willingham (2012), "The race to replace traditional textbooks with electronic versions is on" (p. 1570). As of 2010, Amazon.com has sold more e-books than print books (Bounie, Eang, Sirbu, & Waelbroeck, 2013), and e-book sales continue to show strong and steady growth (Reynolds, 2011). In recent years, many countries (e.g., Japan, South Korea, and Singapore) have implemented and begun assessment of information and communication technology (ICT)-based education and learning materials in schools—and especially of electronic textbooks, or e-(text) books (Nakajima, Shinohara, & Tamura, 2013). The present study focuses on one aspect of e-book use in an educational environment, that is, the digital footprints of students, which can reveal the details of the learning behaviors of students.

To improve teaching and learning, Kyushu University introduced a single platform learning system (Mitsuba, or M2B) that was based on a common learning management system (Moodle), an e-portfolio system (Mahara), and an e-book system (BookLooper). BookLooper allows students to browse e-book materials not only in classrooms but also across time and locations. By the end of 2015, approximately 5,320,000 logs were collected from approximately 20,000 students from various academic courses (e.g., Information Science, Earth and Planetary Science, and History) with the cooperation of approximately 10,000 teachers and other staff members of the university. We utilize this educational big data in our research including analysis of browsing patterns against quiz scores (e.g., Shimada, Okubo, & Ogata, 2016), investigation of effective learning behavior (e.g., Oi, Okubo, Shimada, Yin, & Ogata, 2015; Oi, Yamada, Okubo, Shimada, & Ogata, 2017; Yamada, Yin, Shimada, Kojima, Okubo, & Ogata, 2015), and predictive modeling (e.g., Okubo, Shimada, Yin, & Ogata, 2015).

To ensure effective learning, it is important to cover the same content before and after learning it in a class session (see the review in Shinogaya, 2012). Hereafter, we refer to learning before a class session as a "preview" and learning after the session as a "review." In order to investigate learning behaviors and achievements of students, most previous studies used subjective measures such as questionnaire responses (e.g., Ihmeideh, 2014; Shinogaya, 2014; Woody, Daniel, & Baker, 2010). However, from these questionnaires, it is difficult to learn the details of students' learning behavior. To address this issue, we analyzed e-book logs, which reveals students' behavior in and out classrooms objectively (Oi et al., 2015; 2017). The logs of e-books were obtained from undergraduate students who attended "Information Science" course. E-book logs were categorized as follows: if a log was recorded before a class session in which the same e-book was used as a textbook, it was a preview log, and if after, a review log. The main findings can be summarized as follows: (1) students who obtained consistently good achievement more frequently switched between different e-books and different pages within e-books than low achievers, but (2) this difference was found only for preview logs, not review logs. These results suggested the general tendencies of high and low achievers; however, the details of their behavior are still unclear.

The present study tries to shed light on this problem. We selected high- and low-achiever students who attended Information Science course and analyzed details of their e-book logs. Furthermore, we examined whether students' fundamental knowledge of contents of the course affected their learning behavior. Before beginning the course, if a student already has fundamental knowledge of the contents of the course, it may help his/her learning by acting like an advance organizer (Ausubel, 1960). To examine this issue, we introduced a basement test that assessed the fundamental knowledge of students before the course starts. note that the first paragraph of a section or subsection is not indented. The first paragraphs that follows a table, figure, equation etc. does not have an indent, either.

2 Methods

We analyzed logs from Information Science course (from 2016.04.12 to 2016.07.26). The objective of the course was to understand the fundamentals of ICT. One hundred and ten students participated in this course. The number of sessions in the course was 14. For assessment of students' fundamental knowledge of ICT, before beginning the first lecture, the students took a basement test that consisted of some questions from the Information Technology Engineers Examination1[1]. Figure 1A shows the distribution of the scores and its quartile of the basement test. Students also took a midterm and end-term examination during the 8th and 14th sessions, respectively.

After all of the sessions in the course, students were given their final score, which was converted into a grade (i.e., A: 90–100, B: 80–89, C: 70–79, D: 60–69, and F: less than 60). The final scores were calculated for each student from his/her mid-term examination score (30%), end-term examination score (30%), short report (10%), and attendance (20%). Figure 1B shows the distribution of the final score and the grade.

For analyses, we excluded logs from students who did not take the basement test (n = 4), the mid-term examination (n = 4), or the end-term examination (n = 2), and who did not submit any short reports and took grade "F" (n = 1). We considered the score of the basement test to represent the amount of students' fundamental knowledge of ICT (i.e., the contents of the course). To categorize students who had much knowledge or less knowledge, the students were divided into four groups according to the quartile of the scores of the basement test.

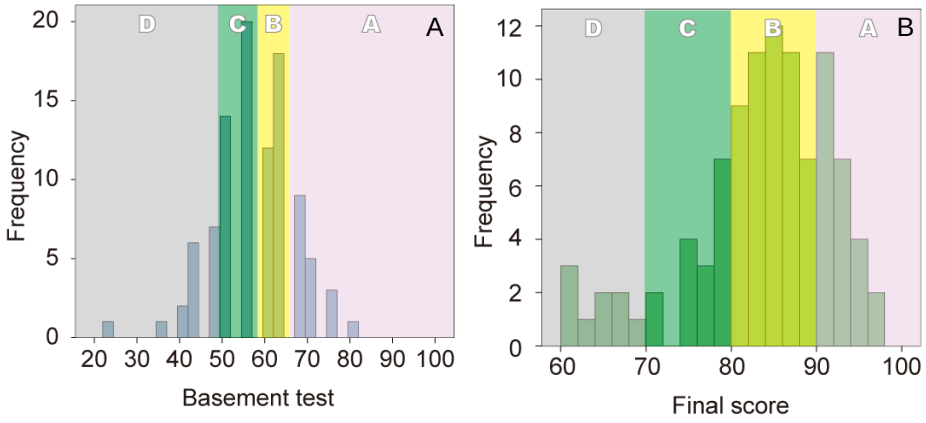

Fig. 1. Distribution of scores and the quartiles of (A) the basement test and (B) the final score

3 Results and Discussion

Table 1 breaks down the number of the students according to a combination of the quartile of the basement test and the grade. We focused on four groups: two high-

[1] See website for details of the examination. https://www.jitec.ipa.go.jp/index-e.html

achieving groups and two low-achieving groups. The first group consisted of students who maintained high scores (A and A). They had fundamental knowledge before the beginning of the course and took the first grade when the course ended. The second group consisted of students who did not have knowledge at the beginning but took the first grade when the course ended (D and A). The third group consisted of students who took the worst grade despite having fundamental knowledge (A and D). The fourth group of students who did not have fundamental knowledge took the worst grade (D and D).

Table 1. Proportion of the quartile of the basement test and the grade.

Basement exam	Final grade			
	A	B	C	D
A	5	10	1	2
B	8	15	5	2
C	7	15	8	4
D	4	10	2	1

Based on our previous studies (Oi et al., 2015; 2017), we categorized the previews and reviews in e-book logs. First, we summarized whether students performed a preview and/or review for each session. Regardless of the number of preview or review for a session, we simply counted the frequencies of preview and review as performed or not for each session.

Figure 2 shows the frequencies of the (A) previews and (B) reviews of the four groups. One remarkable feature of the preview patterns is that none of the students in the low-achieving (AD and DD) groups performed previews. In contrast, three of the five students in the AA group performed previews for more than half (i.e., six) of the sessions, but one student did not perform previews. This student probably had enough knowledge and could understand the sessions without preview. The difference between AA and DA groups is not very salient, but only the students in AA group performed previews for more than half of the sessions. This may reflect the other aspect of the fundamental knowledge of the AA group that is it might help them to understand the contents of the e-book before they were taught the contents in the class session, so they more frequently and easily performed previews than the students in DA group did. For reviews, all of the students in the high-achieving (AA and DA) groups performed them for all of the sessions. Unlike in the case of the previews, the students in the low-achieving groups performed reviews.

None of the low-achieving students performed previews. However, the number of students was only three. To further investigate the learning behavior of the low achievers, we added DC students and CD students to our analyses. Figure 3 shows the frequencies of the (A) previews and (B) reviews of CD and DC groups. Three of six students performed previews only once or not at all. However, the remaining three performed previews for half or more of the sessions. These results indicate that these three students paid attention in the sessions, but their previews did not work well. For reviews, all of the students in CD and DC groups performed them for more than half of the sessions.

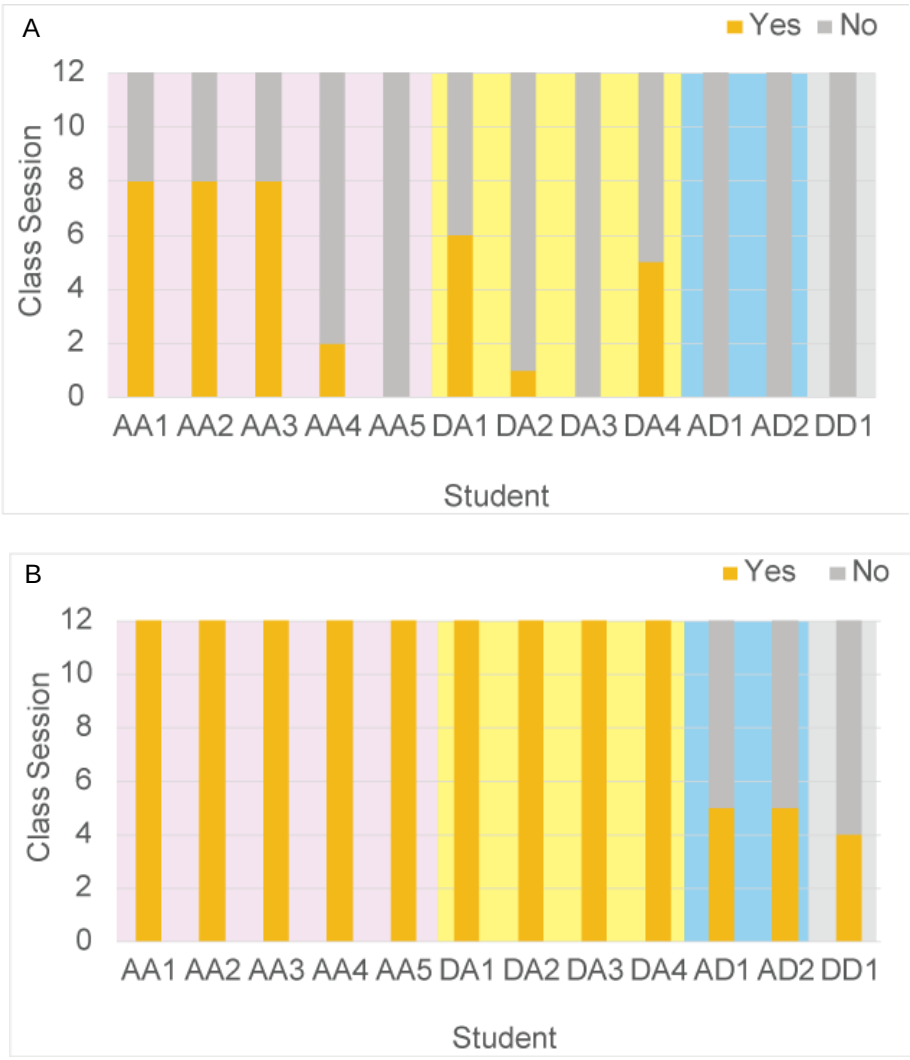

Fig. 2. Frequencies of (A) preview and (B) review of four groups.

To investigate more details of the learning behavior of the high and low achievers, we further analyzed preview logs, because the differences between the high and low achievers were more prominent in the previews than in the reviews, as previous studies have indicated (Oi et al., 2015; 2017). We selected logs of three students in the AA group who performed previews for more than half of the sessions as high achievers, and three students in CD and DD groups who performed previews relatively frequently as low achievers. According to our previous studies, "one" preview was defined as follows. When students opened an e-book, a preview started, and when the

student changed to another e-book, or when an interval between two logs passed for more than one hour, a preview ended. Then, we calculated the duration (s) and number of page flips for each preview. Figure 4 shows the number of page flips and duration of each preview.

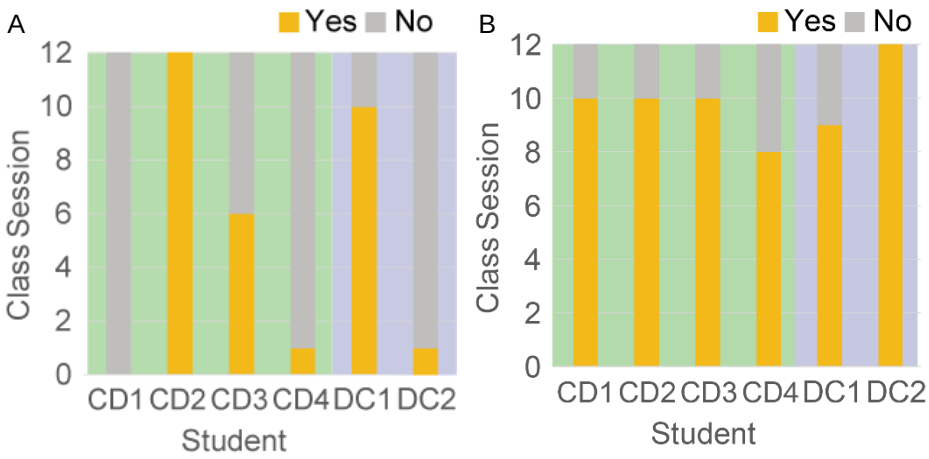

Fig. 3. Frequencies of (A) preview and (B) review of CD and DC groups.

In Figure 4, each bar and each dot indicates each preview. For example, student CD3 in the low-achieving group performed previews six times (Figure 4B). The results show that even though the low achievers performed previews, both their duration and page flips were almost 0 for approximately half of their previews. The high achievers showed such a pattern for a few cases. These low durations and page flips of the low achievers suggest that the students tried to perform previews but gave up for almost half of them. If we prepare more suitable learning materials for the students (e.g., summary of the textbook with annotations), they will probably be able to accomplish their previews. Further analyses of the details of high achievers' usage of their e-books according to their logs, especially in the case of the DA group, may help in the making of such learning materials.

4 Conclusions

We investigated the learning behavior of undergraduates with e-book logs. The results can be summarized as follows. The very low achievers did not perform previews. However, many high achievers performed previews frequently. Relatively low achievers tried to perform previews, but they gave the endeavor up easily. Regardless of high and low achievements, students performed reviews. These results imply that e-book logs can reveal the details of the learning behavior of students.

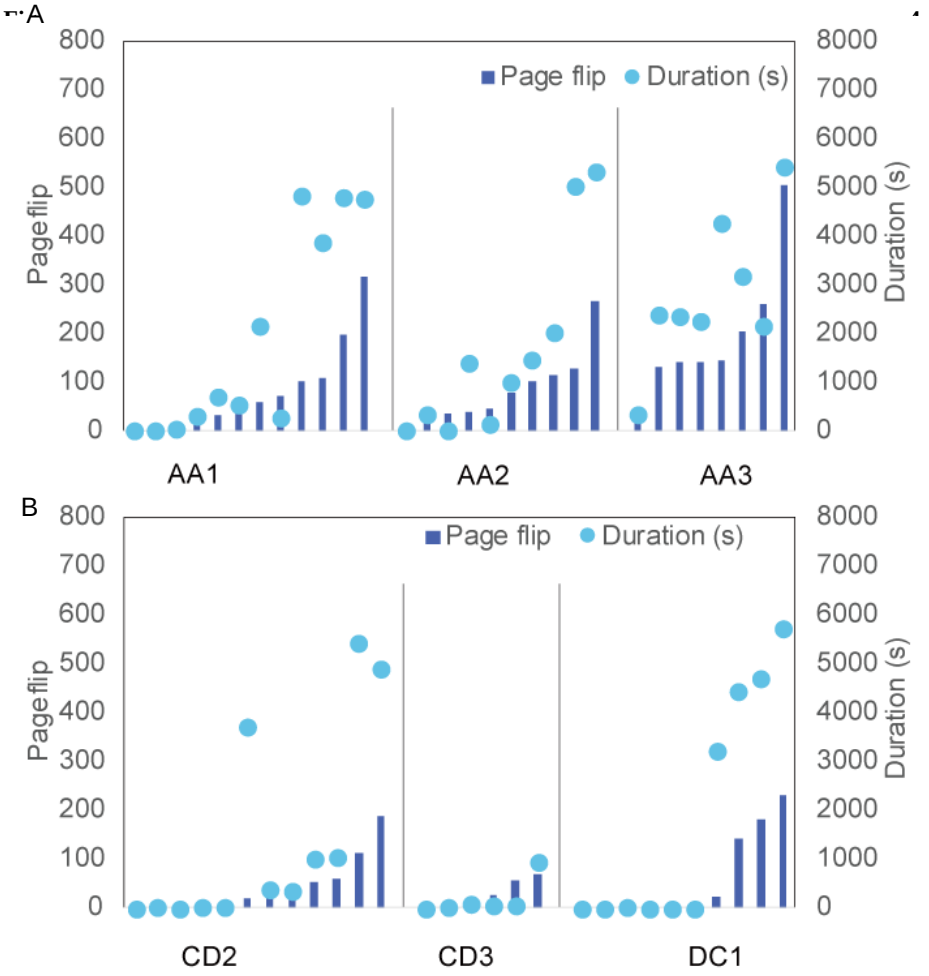

Fig. 5. Number of page flips and duration for each preview of the (A) high-achieving and (B) low-achieving groups. Each bar and each dot indicates each preview.

References

1. Ausubel, D. P. (1960). The use of advance organizers in the learning and retention of meaningful verbal material. *Journal of Educational Psychology*, *51*(5), 267–272. doi:10.1037/h0046669
2. Bounie, D., Eang, B., Sirbu, M., & Waelbroeck, P. (2013). Superstars and outsiders in online markets: An empirical analysis of electronic books. *Electronic Commerce Research and Applications*, *12*, 52–59. doi:10.1016/j.elerap.2012.11.004
3. Daniel, D. B., & Willingham, D. T. (2012). Electronic textbooks: Why the rush? *Science*, *335*(6076), 1569–1571. doi:10.1126/science.335.6076.1569

4. Ihmeideh, F. M. (2014). The effect of electronic books on enhancing emergent literacy skills of pre-school children. *Computers & Education*, *79*, 40–48. doi: 10.1016/j.compedu.2014.07.008
5. Nakajima, T., Shinohara, S., & Tamura, Y. (2013). Typical functions of e-Textbook, implementation, and compatibility verification with use of ePub3 materials. *Procedia Computer Science*, *22*, 1344–1353. doi:10.1016/j.procs.2013.09.223
6. Oi, M., Okubo, F., Shimada, A., Yin, C., & Ogata, H. (2015). Analysis of preview and review patterns in undergraduates' e-book logs. *Proceedings of the 23rd International Conference on Computers in Education (ICCE) (Hangzhou, China, November 30 – December 4, 2015)*, 166–171.
7. Oi, M., Yamada, M., Okubo, F., Shimada, A., & Ogata, H. (2017). Reproducibility of findings from educational big data: A preliminary study. *Proceedings of the 7th International Learning Analytics and Knowledge (LAK) (Vancouver, Canada, March 13 – 17, 2017)*.
8. Okubo, F., Shimada, A., & Yin, C. (2015). Visualization and prediction of learning activities by using discrete graphs. *Proceedings of the 23rd International Conference on Computers in Education (ICCE) (Hangzhou, China, November 30 – December 4, 2015)*, 739–744.
9. Reynolds, R. (2011). Trends influencing the growth of digital textbooks in US higher education. Publishing Research Quarterly, *27*(2), 178–187. doi:10.1007/s12109-011-9216-5
10. Shimada, A., Okubo, F., & Ogata, H. (2016). Browsing-pattern mining from e-book logs with non-negative matrix factorization. *Proceedings of the 9th Educational Data Mining (EDM) (Raleigh, NC, USA, June 29 – July 02, 2016)*, 636–637.
11. Shinogaya, K. (2012). Learning strategies: A review from the perspective of the relation between learning phases. *Japanese Journal of Educational Psychology*, *60*, 92–105.
12. Shinogaya, K. (2014). Students' strategies in preparation and lectures: Direct and moderating effects of teachers' teaching strategies. *Japanese Journal of Educational Psychology*, *62*, 197–208.
13. Yamada, M., Yin, C., Shimada, A., Kojima, K., Okubo, F., & Ogata, H. (2015). Preliminary research on self-regulated learning and learning logs in a ubiquitous learning environment, *Proceedings of the 15th IEEE International Conference on Advanced Learning Technologies* (*ICALT2015) (Hualien, Taiwan, July 6 – July 9 2015)*, 93–95. doi:10.1109/ICALT.2015.74
14. Woody, W. D., Daniel, D. B., & Baker, C. A. (2010). E-books or textbooks: Students prefer textbooks. *Computers & Education*, *55*(3), 945-948. doi:10.1016/j.compedu.2010.04.005

Acknowledgments. These research results have been achieved under the theme of "Research and Development on Fundamental and Utilization Technologies for Social Big Data" (178A03), as Commissioned Research for the National Institute of Information and Communications Technology (NICT), Japan.