Chen, W. et al. (Eds.) (2017). Proceedings of the 25th International Conference on Computers in Education.
New Zealand: Asia-Pacific Society for Computers in Education

Effects of Prior Knowledge of High Achievers on Use of e-Book Highlights and Annotations

Misato OI*, Fumiya OKUBO, Yuta TANIGUCHI, Masanori YAMADA, Shin'ichi KONOMI
Faculty of arts and science, Kyushu University, Japan
*oimisato@gmail.com

Abstract: To identify "good performance," this study analyzed the highlighting and annotating action logs of undergraduates during their e-book usage. To reveal "good performance," the study focused on the learning behavior of high achieving students. Few highlights and annotations were observed for both rich knowledge and poor knowledge high achievers. Moreover, in the spontaneous usage of e-books outside the classroom, high and poor knowledge students did not display differences in highlights and annotations.

Keywords: e-book, preview, review, annotation, highlight

1. Introduction

In the past decade, the popularity of big data and analytics has increased (Picicciano, 2012), including in education. Many researchers have highlighted the potential of big data in education (e.g., Daniel, 2015; Long & Siemens, 2011; Picciano, 2012). For example, "big data can provide institutions of higher education the predictive tools they need to improve learning outcomes for individual students as well ways ensuring academic programmes are of high-quality standards" (Daniel, 2015).

To improve learning and teaching, 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). The e-book system enables students to browse e-book materials before/during/after lectures, anywhere and anytime, using their personal computer (PC) or smartphone. All user actions performed on this e-book system, such as page flips and opening a material, are recorded as learning logs and automatically sent to the University's database when a network connection is available. The e-book system provides the additional functions of bookmarking, highlighting, annotating, and searching. By the end of April 2017, approximately 45,000,000 logs (Moodle: 28,000,000; Mahara: 1,000,000; BookLooper: 16,000,000) were collected from approximately 20,000 students. Analysis of educational big data from M2B provided insights into the activities of students, such as browsing patterns with respect to quiz scores (e.g., Shimada, Okubo, & Ogata, 2016), effective learning behavior (e.g., Oi, Okubo, Shimada, Yin, & Ogata, 2015a; Oi, Yin, Okubo, Shimada, Kojima, Yamada, & Ogata, 2015b; Oi, Yamada, Okubo, Shimada, & Ogata, 2017a; Oi, Yamada, Okubo, Shimada, & Ogata, 2017b, Yamada, Shimada, Okubo, Oi, Kojima, & Ogata, in press), and predictive modeling (e.g., Okubo, Shimada, Yin, & Ogata, 2015).

To develop an effective feedback and/or intervention system, the concept of "good performance" needs to be clarified to reveal to students their goal (Sadler, 1989). The aim of the present study is to identify "good performance" by analyzing e-book logs of M2B. In previous studies (Oi et al., 2015a, b; Oi et al., 2017a, b), we focused on learning behavior, namely, covering the same content before (preview) and after (review) its learning in a class session. Undergraduates' performance of such preview and review was analyzed based on e-book logs 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 are as follows: (1) preview is more deeply related with academic achievements than review (Oi et al., 2015a; 2017b), (2) relatively low achievers attempted to perform previews, but they tended to give up easily on the endeavor (Oi et al., 2017b).

As a first step to identifying "good performance," the present study analyzed the highlighting and annotating action logs of undergraduates during e-book usage. To understand new information, annotating and highlighting are useful techniques, and these are assumed as a valuable part of the process of learning (Glover, Xu, & Hardker, 2007). A study reported that university students commonly used annotation to identify key parts of the document (Ovsiannikov, Arbib, & McNeill, 1999). As pointed out by a classical study (Ausubel, 1960), if a student has prior knowledge of the contents of a course, it may help the student's learning by acting as an *advance organizer*. In other words, students' prior knowledge of contents of a course could affect their learning behavior, and that "good performance" may differ according to the amount of knowledge processed. For example, a student who has less knowledge has to verify technical words and the relationships between fundamental concepts, particularly while performing a preview. However, a student who has considerable knowledge does not need to verify such words and relationships. Based on this hypothesis, we examined differences between rich vs. poor knowledge students and their preview vs. review performances. To reveal "good performance," we mainly focused on the learning behavior of high achieving students.

2. Methods

2.1 Participants

E-book logs were collected from 110 undergraduates enrolled in an information science course (from April 4 to July 26, 2016, 14 sessions); these logs were also analyzed in Oi et al. (2017a, b). The objective of the course was for students to understand the fundamentals of information and communication technology (ICT). For the assessment of students' prior knowledge of ICT, the students took a placement test before beginning the first session; the test comprised questions from the Information Technology Engineers Examination (https://www.jitec.ipa.go.jp/index-e.html). Students also took a midterm and an end-term examination during the 8th and 14th sessions, respectively. The teacher did not provide clear instructions to the students regarding the use of highlights or annotations.

After completing 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 reports (10%), and attendance (20%).

For analyses, we excluded logs from students who did not take the placement test (n = 4), the mid-term examination (n = 4), or the end-term examination (n = 2); those who did not submit any short report; and those who received a grade "F" (n = 1). We considered the scores of the placement test to represent the level of students' prior knowledge of ICT (i.e., the contents of the course). In categorizing students with rich or poor knowledge, they were divided into four groups according to the quartile of the scores of the basement test.

Table 1 summarizes the number of students according to a combination of the quartile of the basement test and the grade.

Table 1: Proportion of the quartile of the placement test and the grade.

Placement test	Final grade				
	A	B	C	D	Sum
A	5	10	1	2	18
B	8	15	5	2	30
C	7	15	8	4	34
D	4	10	2	1	17
Sum	24	50	16	9	99

2.2 E-book Logs

The total number of e-book logs was 447,650. Table 2 presents a sample of the e-book logs.

Table 2: Sample of e-book logs.

logid	userid	operationname	operationdate	contentsid	deviceid	memo	page_no
UTG7z	4RJqBr	OPEN	2015/3/31 14:57	012ABC	76UjvV		0
GbycT	4RJqBr	PREV	2015/3/31 14:57	012ABC	76UjvV		0
My0bl	4RJqBr	PREV	2015/4/2 10:21	012ABC	UFQq7C		1
qUQxf	4RJqBr	NEXT	2015/4/2 10:21	012ABC	UFQq7C		1
1JCv7	4RJqBr	ZOOM	2015/4/2 10:21	012ABC	UFQq7C		1
zN3Gl	4RJqBr	ZOOM	2015/4/2 10:21	012ABC	UFQq7C		1
GLJPt	4RJqBr	ZOOM	2015/4/2 10:21	012ABC	UFQq7C		1
AHLfX	4RJqBr	ADD MARKER	2015/4/2 10:21	012ABC	UFQq7C		1
nMKVx	4RJqBr	ADD MEMO	2015/4/2 10:21	012ABC	UFQq7C	1 ギガ	1
94xjjM	4RJqBr	PORTRAIT	2015/4/2 10:21	012ABC	UFQq7C		1

3. Results and Discussion

3.1 Number of e-Book Logs

First, we confirmed whether the number of e-book logs varied vis-à-vis students' final grade and rank in the placement test. Figure 1 presents the average number of e-book logs for each group. To account for groups with a small *Ns* (e.g., $n = 1$ for A-C [Placement-Final grade] and D-D), *SDs* are not shown. To examine the differences, we performed one-way ANOVAs on the number of e-book logs with groups of placement test and final grade, respectively, as between-subject factors. As the *Ns* of some groups were considerably small, we did not employ a two-way ANOVA. The ANOVAs revealed a significant difference among the final grade, $F(3, 95) = 11.42$, $p < .0001$, $\eta^2 = 0.27$. However, the difference among the groups of the placement test was not significant, $F < 1$. Figure 2 presents the average number of e-book logs for each final grade group. Multiple comparisons on difference among the final grades with Bonferroni correction (corrected $p < .05$) revealed that the number of e-book logs of group A was significantly higher than that of the other three groups. No other significant differences were observed between the groups. These results indicate that (1) high achievers use e-books more frequently than middle and low achievers, and (2) prior knowledge of the course did not significantly affect the number of e-book logs.

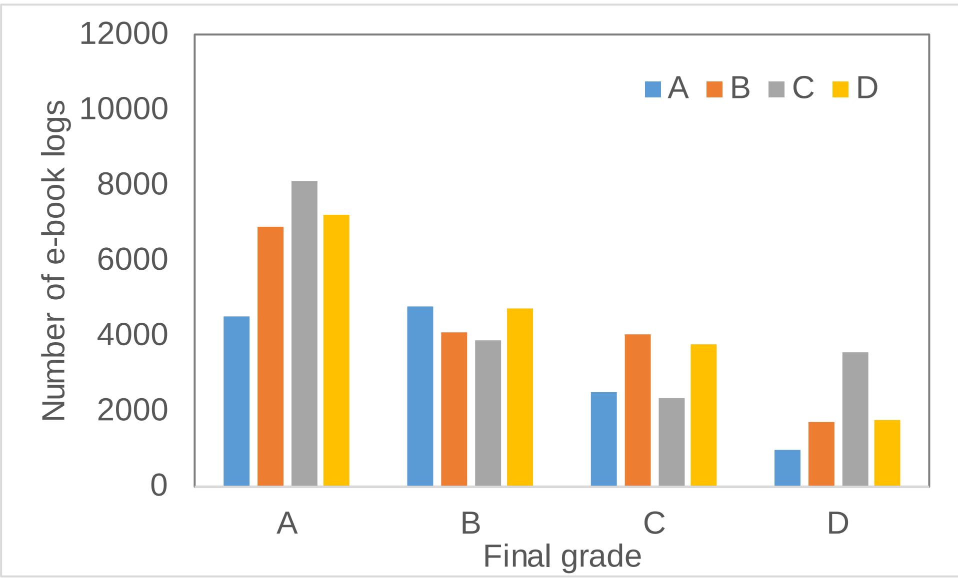

Figure 1. Average number of e-book logs for each group.

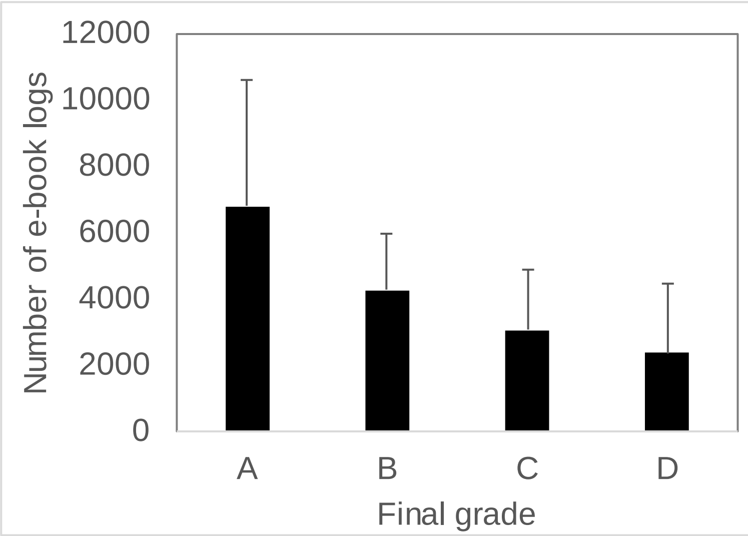

Figure 2. Average number of e-book logs for each final grade.

3.2 Highlights and Annotations

We focused on high achievers with a final grade of A. To examine whether the level of prior knowledge affects high achievers' annotations and highlights, we analyzed the frequency of highlights and annotations of the rich knowledge group (i.e., A-A) and poor knowledge group (i.e., D-A) (see Table 1) during preview, class session, and review. Table 3 presents the number of highlights and annotations for each student. Only two students from the rich knowledge (A-A) group and one from the poor knowledge (D-A) group used highlights during the spontaneous e-book usage outside the classroom (i.e., both preview and review). In other words, the use of highlights by both rich and poor knowledge high achievers was relatively minimal. Although DA01 added 42 highlights, we could not determine whether this log indicated a characteristic of poor knowledge high achievers or simply that of the individual student.

Both groups showed few annotations throughout their preview and review. Only AA03 showed annotation logs during preview; however, the text fields of the annotations were blank. In other words, AA03 clicked the annotation button but did not write anything. Moreover, during the review and class session, six students clicked on the annotation button but no text was written. We checked the annotation command logs of all students: 209 of the 588 logs contained text. For example, an annotation text of one student is "Merge sort is three times faster." (This annotation text was translated from Japanese). The remaining 379 logs did not have text.

Table 3: Frequency of highlights and annotations during preview and review.

	AA01	AA02	AA03	AA04	AA05	DA01	DA02	DA03	DA04
Highlight									
Preview	15	0	3	0	0	5	0	0	0
Class	3	0	6	1	0	6	4	0	5
Review	0	0	6	0	0	42	0	0	0
Annotation									
Preview	0	0	3	0	0	0	0	0	0
Class	0	0	0	0	2	0	0	0	1
Review	1	0	0	1	0	0	1	1	0

High achievers did not write annotation text, and a third of them used the highlight function in preview. To confirm whether the use of few annotations and highlights is a characteristic of high achievers, we summarized the number of students who used these functions for each grade throughout all of the logs. Table 4 summarizes the number of students who used highlights or annotations. Almost half of the students did not use these functions, and thus, we could not directly compare the number of annotations and highlights with respect to the final grades. We performed Chi square tests on the ratio of the students who used the functions and those who did not. The Chi square tests revealed the absence of significant differences among the grade groups for both the highlights, $\chi^2(3) = 6.463$, $p = .091$, and the annotations, $\chi^2(3) = 6.151$, $p = .104$. Thus, there was no significant difference in the grades of the ratio of students who used the functions of highlighting and annotating.

An implication of the above results is that without explicit instructions to use highlights and annotations, students may not actively use these functions of the e-book system. However, these analyses were performed based on the e-book logs; records of usage of other electronic devices and paper notes were not included. More enriched functions (e.g., free sketching and automatic summarization of annotations and highlights) could perhaps encourage students to use annotation and highlights and thus enable deeper analyses of learning behavior.

Table 4: Number of students who used highlights and annotations.

	A	B	C	D
Highlight				
Yes	16	32	6	3
No	8	18	10	6
Annotation				
Yes	16	27	5	3
No	8	23	11	6

4. Conclusion

As a first step to identify "good performance," the present study analyzed the highlighting and annotating action logs of undergraduates during e-book usage. To determine "good performance," we mainly focused on the learning behavior of high achieving students. Few highlights and annotations were observed for both rich knowledge and poor knowledge high achievers. In the spontaneous usage of e-books outside the classroom, high and poor knowledge students did not exhibit differences in their use of highlights and annotations.

Acknowledgements

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.

References

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

Daniel, B. (2015). Big data and analytics in higher education: Opportunities and challenges. *British Journal of Educational Technology*, *46*(5), 904–920. doi: 10.1111/bjet.12230

Glover, I., Xu, Z. J., & Hardaker, G. (2007). Online annotation: Research and practices. *Computers & Education*, *49*(4), 1308–1320.

Long, P., & Siemens, G. (20). Penetrating the fog: Analytics in learning and education. *EDUCAUSE Review*, *46*, 30-32, 34, 36, 38, 40.

Oi, M., Okubo, F., Shimada, A., Yin, C., & Ogata, H. (2015a). 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 3–December 4, 2015).*

Oi, M., Yin, C., Okubo, F., Shimada, A., Kojima, K., Yamada, M., & Ogata, H. (2015b). Analysis of links among e-books in undergraduates' e-Book logs. *Workshop of the 23rd International Conference on Computers in Education*, *(Hangzhou, China, November 30–December 4, 2015).*
Oi, M., Yamada, M., Okubo, F., Shimada, A., & Ogata, H. (2017a). 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)*.
Oi, M., Yamada, M., Okubo, F., Shimada, A., & Ogata, H. (2017b) Finding traces of high and low achievers by analyzing undergraduates' e-book logs. *Cross-LAK2017*, *(Vancouver, Canada, March 12, 2017)*.
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.
Ovsiannikov, I. A., Arbib, M. A., & McNeill, T. H. (1999). Annotation technology. *International Journal of Human-Computer Studies*, *50*(4), 329–362.
Picciano, A. G. (2012). The evolution of big data and learning analytics in American higher education. *Journal of Asynchronous Learning Networks*, *16*(3), 9–20.
Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*, 119–144.
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.
Yamada, M., Shimada, A., Okubo, F., Oi, M., Kojima, K., & Ogata, H. (in press). Learning analytics of the relationships among self-regulated learning, learning behaviors, and learning performance. *Research and Practice in Technology Enhanced Learning.*