An Analysis of Students' Opinions on Using a Real-time Survey Tool

Yoko Okita

Juntendo University

yohkita.hawaii@gmail.com

Abstract

This study examined the usefulness of a real-time polling tool, Respon (2021), quantitatively and qualitatively, in a large hybrid Japanese composition class in Tokyo in 2021. The participants were 164 first-year university students. The result of the evaluation scores showed that the use of Respon was received well by the students. Text mining of the students' written comments revealed invisible relationships among the students' perceptions of using Respon and three academic grade levels: low, middle, and high. The high- and middle-grade students wrote words related to understanding the content more frequently than expected compared with the low-grade students, who appeared not to utilize real-time polling to understand the content of the class in the same manner as the high- and middle-grade students. The high-grade students seemed more interested in knowing others' opinions than the middle- and low-grade students. Words related to feeling of participation were written by 40% to 46% of the students across the academic grade levels. Overall, use of real-time polling in a hybrid class created a sense of participation among the students. These results are useful for understanding the learning processes of students and for planning future instruction.

Keywords

real-time polling, text mining, hybrid class

1 Introduction

This study sought to evaluate the usefulness of a real-time polling tool in a large hybrid university class based on the evaluation scores and text mining of the students' written comments. Real-time polling was employed to make a large hybrid class interactive and to enable teachers to obtain the students' written comments digitally. This was an excellent opportunity to evaluate the usefulness of real-time polling in a hybrid context using text mining technologies. Furthermore, examining the relationship between the students' written comments and their academic grades provided useful information on which aspects of real-time polling were appreciated by which kinds of students.

Managing large classes online is extremely challenging for both teachers and students who have little experience of online teaching and learning (Buckley et al., 2021). According to Menon et al. (2004), the benefits of using a personal digital polling tool in a class include promoting interactive learning, keeping the students' attention focused, and increasing their satisfaction. Real-time polling allows teachers to understand the levels of the students' prior knowledge and their understanding of the content during the class, enabling teachers to tailor their teaching to the students (Stover et al., 2014). Stanley (2013) reported that using clickers improved the students' achievements in classes over the course of a semester. Real-time polling improved the students' understanding of the content of the course, and increased their levels of participation (Stover et al., 2014). The use of real-time polling creates interactive and learner-centered classrooms.

Student feedback is commonly used to evaluate the quality of teaching and to improve future teaching practice, as well as the learning outcomes at universities (Gronberg et al., 2021; Wu & Dawson, 2022). Student feedback provides valuable information about the students' perspectives of learning experiences (Santhanam et al., 2022); student feedback surveys usually consist of quantitative Likert-type questions and qualitative open-ended questions. Quantitative numerical outcomes are widely used because they can be calculated quickly, analyzed statistically, and used as indicators of teaching quality. However, qualitative written comments tend not to be used as often because of the difficulty of analyzing unstructured text data and extracting useful educational information (Santhanam et al., 2022; Stupans et al., 2016). However, qualitative student feedback can provide insights into students' learning experiences in a specific context and can be used to improve teaching and learning outcomes (Steyn et al., 2019). Student feedback also reveals unexpected aspects that quantitative questions cannot identify (Steyn et al., 2019; Wright, 2022).

Due to advances in computer technologies that can be incorporated into natural language processing, students' written feedback can be collected digitally and analyzed using text mining

techniques. Many different text mining techniques have been developed and reported in studies of the text mining of students' feedback (Gronberg et al., 2021). Text mining extracts important information from unstructured texts (Ferreira‐Mello et al., 2019), and provides not only word frequencies but also word clustering and visualized associations of words (Zaitseva et al., 2022). The text mining of students' written comments provides useful information that can improve course designs and learning outcomes.

Little research has examined student feedback by employing both quantitative and qualitative analyses. Quantitative numerical results provide holistic views of students' opinions easily, whereas qualitative results provide detailed information about the students' perceptions and experiences of using real-time polling in a hybrid class. The present study examined student feedback regarding the usefulness of real-time polling during lectures both quantitatively and qualitatively. In this study, the students' feedback questions consisted of a quantitative question and a qualitative question. The quantitative question measured their satisfaction levels using a 5-point Likert scale, and the qualitative question asked the students to write free comments. The students' quantitative and qualitative responses were collected digitally using Respon (2021).

In the Japanese context, the text mining software KHCoder (2022) has been widely used in academia because it is free and has many useful analytical functions, ranging from word counting to the visualized mapping of word networks. Little research has explored the relationship between students' written feedback and their academic outcomes. KHCoder includes a clustering analysis that groups related words in unstructured texts. In the present study, the author attempted to identify the relationships between word clusters and academic grades using KHCoder. Identifying the types of students who appreciated various aspects of real-time polling in a large hybrid class reveals the underlying nature of the students' learning processes in this specific context. The results of this study were expected to provide useful information for implementing the effective use of real-time polling in classrooms.

2 Methods

2.1 The Participants and the Class

Two-hundred-forty-two first-year university students enrolled in a mandatory Japanese composition class in the first half of the 2021 school year in Tokyo. The class met for 90 minutes once a week for 13 weeks. The students were divided into two groups based on their student ID numbers. One group participated in a face-to-face class in the classroom, and the other participated in an online class via Zoom (2021). The students took turns attending the face-to-face class or the online class. The students were required to read the course material prior to each class. During the class, the teacher explained the background of the material or the technical terminology involved based on the results of the real-time polling questions. Five to ten polling questions were asked in 10 of the 13 classes. Summarizing the reading material was assigned as homework in each class.

The students were classified according to five groups based on their average academic scores at the end of the 2021 school year, namely Group A (100-90 points), Group B (89-80 points), Group C (79-70 points), Group D (69-60 points), and Group F (59 points or less). The students' average academic grade scores at the end of the 2021 school year were used as the indicator of their academic outcomes because these scores would represent the students' overall academic outcomes more accurately than would their grade scores for a single class (Japanese composition).

2.2 Real-time Polling Tool

Respon (2021) is a smartphone-based, real-time polling system. The students installed the Respon application on their smartphones. After opening the application, the students entered their student ID numbers to access the application, following which a keypad appeared on the screen. Each question had a nine-digit ID number. After entering the ID number for a question, the students could access the answer page for the question. The question types were multiple choice or free writing. The results of the students' responses appeared immediately on the teacher's Respon web site. The teacher could choose whether a web page with the students' responses would be projected onto the screens in the classroom or not. A web page that presented the answer to a multiple-choice question consisted of a graph and a table, as shown in Figure 1. The students' written comments appeared on a web page in the order of the earliest first. An example of a web page showing the students' free writing of comments is presented in Figure 2. The teacher could choose whether the students' comments were shown anonymously or not. In this study, the students' written comments were shown anonymously in all the lectures. The teacher could download the students' responses to the multiple-choice questions and the written comments linked to their ID numbers after the class.

Answering questions was counted as attendance points regardless of whether an answer was correct or incorrect.

Figure 1

An Example of a Web Page Showing the Results for a Multiple-choice Question

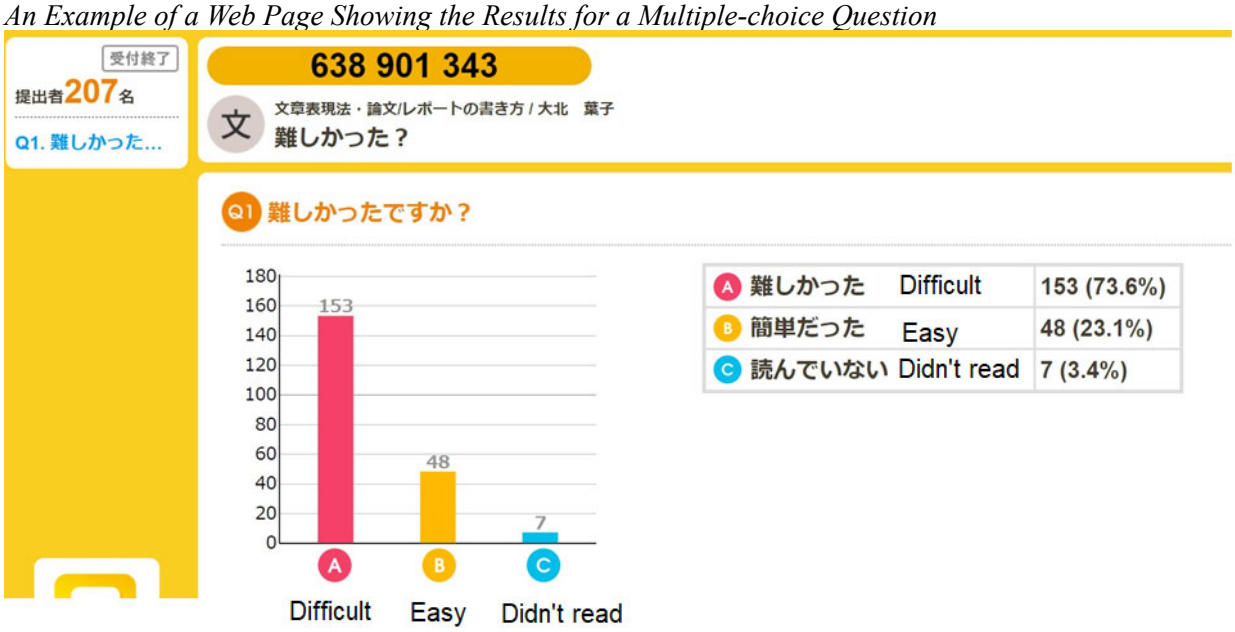
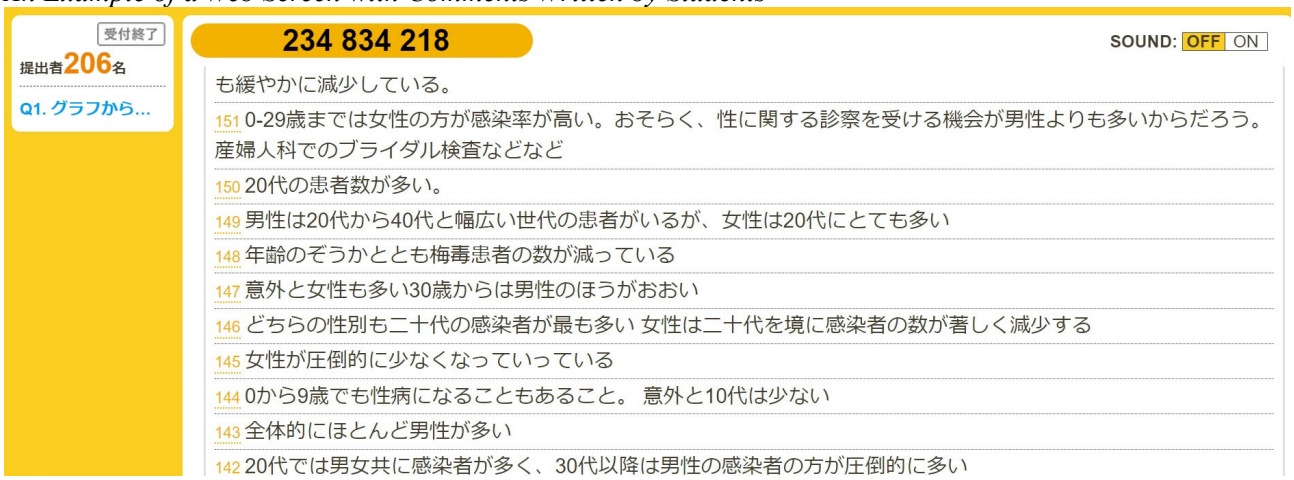


Figure 2

An Example of a Web Screen with Comments Written by Students



2.3 Data Collection Procedures

In the last class of the 13 weeks of classes, the students were asked to answer one quantitative and one qualitative question to evaluate the usefulness of Respon (2021). The quantitative question asked the students about their perceptions of the usefulness of Respon using a 5-point Likert scale, as follows: 1 = Not useful at all, 2 = Not useful, 3 = Undecided, 4 = Useful, and 5 = Very useful. The qualitative question asked the students to write free comments about the usefulness of Respon in the classes. The students' responses to the quantitative and qualitative questions were not shown to the students in the class, and their responses did not affect their grades for the class. The number of students who answered or did not answer the qualitative and quantitative questions in the five academic grade groups are presented in Table 1. The data from 164 students who provided written comments were analyzed for text mining. The average age of the 164 students was 18.42 years (SD = 0.793). As very few low-grade Grade D (n = 4) and Grade F (n = 3) students answered the questions, the students in Groups C, D, and F were combined as one group of low-grade students in the qualitative and quantitative analyses. Three of the 164 students did not answer the quantitative questions, leaving 161 students who did answer the quantitative question. The data from 161 students were analyzed for the quantitative analysis. Thirty-seven students answered the qualitative questions but did not answer the qualitative question. Forty-one students did not answer either the qualitative or the quantitative questions.

Table 1

The Number of Students Who Answered or Did Not Answer the Qualitative and Quantitative Questions in the Five Academic Grade Groups

Qualitative question	*Quantitative question*	*A 100-90*	*B 89-80*	*C 79-70*	*D 69-60*	*F 59 or lower*	*Total*
Answered	Answered	20	102	32	4	3	161
Answered	Did not answer	0	1	2	0	0	3
Did not answer	Answered	5	18	10	2	2	37
Did not answer	Did not answer	0	8	10	9	14	41

2.4 Data Analyses

A one-way ANOVA test was conducted to identify the differences in the students' evaluations of the usefulness of Respon (2021) according to their academic grades. KHCoder (2022) was used to analyze the students' written comments. First, hierarchical clustering was conducted to obtain word clusters that consisted of related words in the free writing texts. Second, a cross-tabulation analysis of the clusters and the students' year-end academic grades was conducted to examine the relationships among word clusters and academic grades. Third, chi-square tests were performed to determine whether the proportion of students who provided written comments and that of those who did not were equal across the three academic grades in each cluster.

3 Results

3.1 Quantitative Analysis

The average score for the evaluation of the use of Respon (2021) was 3.97 (SD = .941, N = 161). It appeared that the use of Respon in the classes was received well by the students.

A one-way ANOVA revealed no significant difference in the evaluation scores for the use of Respon across the three academic grades, $F(2, 158) = 1.967$, $p = .143$. The students' perceptions of the usefulness of Respon were not affected by their academic grades. Table 2 presents the mean scores, the standard deviations, and the medians of the students' perceptions regarding the usefulness of Respon for each of the academic grade groups.

Table 2

Mean, Standard Deviations and Medians for the Evaluation Scores Regarding the Use of Respon According to Academic Grade

Academic Grade	*n*	*Mean score*	*SD*	*Mdn*
A	20	4.10	.852	4
B	102	4.04	.916	4
C	39	3.71	1.025	4

3.2 Text Mining Analyses

One-hundred-sixty-four students wrote 237 sentences. In total, 4,153 words were extracted from 237 sentences. Table 3 shows the list of the 10 most frequently written words. The two most frequently used words were not related to the evaluation of the use of Respon (2021); however, the third to the tenth most frequently used words reflected positive evaluations.

Table 3

The 10 Most Frequently Used Words

Rank	*Words (Japanese)*	*Words (English)*	*Frequency*	*Rank*	*Words (Japanese)*	*Words (English)*	*Frequency*
1	授業	class	69	6	良い	good	22
2	レスポン	Respon (2021)	46	7	工夫	derive	20
3	参加	participation	26	8	使う	use	19
4	意見	opinion	23	9	回答	answer	18
5	内容	content	23	10	集中	concentrate	18

Figure 3 shows the cluster dendrogram of the words. Seven clusters were obtained. The words included

in Cluster 1 were participation, class, concentration, lecture, student, and motivation; Cluster 1 was labeled "Feeling of participation". The words included in Cluster 2 were feeling, a little bit more, number, and long; Cluster 2 was labeled "Numbers". The words included in Cluster 3 were listening, teacher, myself, thinking, time, understand, devise, especially, answer, using, and answering; Cluster 3 was labeled "Thinking". The words included in Cluster 4 were a little, not, good, confirmation, can do, every time, and numerical characters; Cluster 4 was labeled "Content confirmation". The words included in Cluster 5 were many, sentences, feel, Respon, content, and understanding; Cluster 5 was labeled "Content understanding." The words included in Cluster 6 were readings, questions, Chinese characters, easy, and reduce; Cluster 6 was labeled "Chinese characters". The words included in Cluster 7 were others, people, thoughts, knowing, opinions, see, answers, mind, and being able to see; Cluster 7 was labeled "Knowing others' opinions". Clusters 2 and 6 reflected negative opinions. Cluster 2 showed that the students complained about the long nine-digit ID numbers for the questions in Respon, while Cluster 6 showed that the students complained about the implementation of frequent Chinese character tests.

Figure 3

Cluster Dendrogram of Words

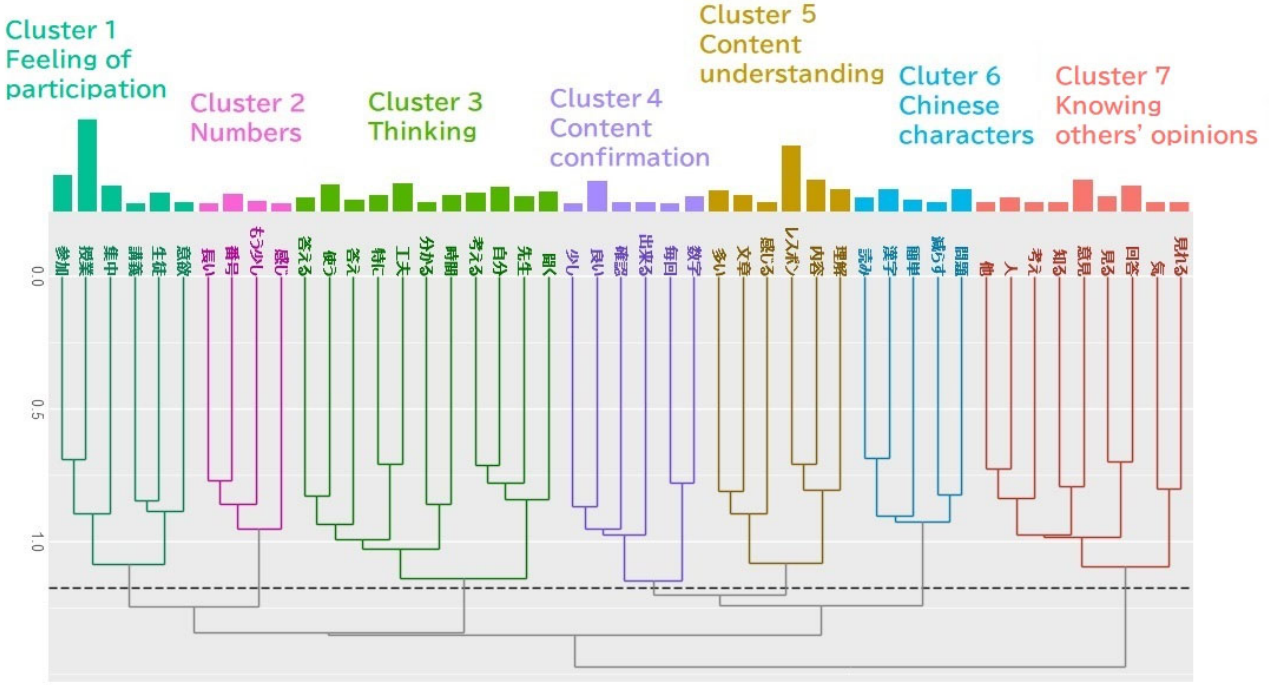


Figure 4 shows the babble chart for cross-tabulation between the word clusters and the academic records. Figure 5 shows the heat chart for cross-tabulation between the word clusters and the academic grades. The numbers in the bars in Figure 5 show the percentages of the students who wrote words that were included in each cluster within each grade category.

Table 4 shows the number of students who wrote words that were included in each cluster. The chi-square tests compared the proportion of students who wrote such words and the proportion of those who did not across the three academic grades within each cluster. A significant difference was found in Cluster 5, $X^2(2, N = 161) = 10.27$, $p = .006$. A post hoc residual analysis showed that the number of the words written by the middle-grade students was greater than expected, whereas the number of words written by the low-grade students was less than expected. A weak significant difference was found in Cluster 7, $X^2(2, N = 161) = 5.40$, $p = .066$. A post hoc residual analysis at an alpha level of .10 showed that the number of the words written by the high-grade students was greater than expected, whereas the number of words written by the low-grade students was less than expected. The words in Clusters 5 and 7 tended to be written more frequently as the students' academic grades increased. No significant difference was found between the proportions of written words and academic grades in Cluster 1, $X^2(2, N = 161) = .328$, $p = .848$, Cluster 2, $X^2(2, N = 161) = .588$, $p = .745$, Cluster 3, $X^2(2, N = 161) = .204$, $p = .360$, Cluster 4, $X^2(2, N = 161) = .547$, $p = .760$, and Cluster 6, $X^2(2, N = 161) = 2.73$, $p = .254$.

Figure 4

Babble Chart of Cross-tabulation between Word Clusters and Academic Grades

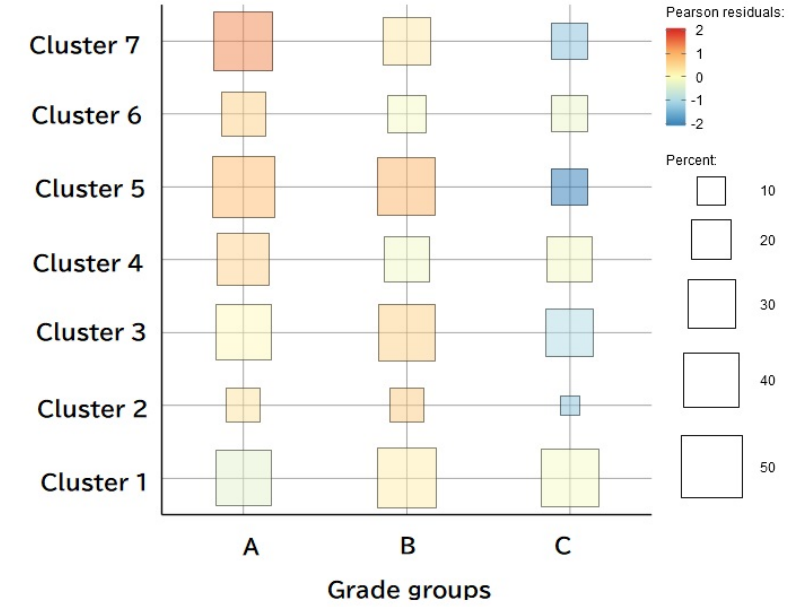
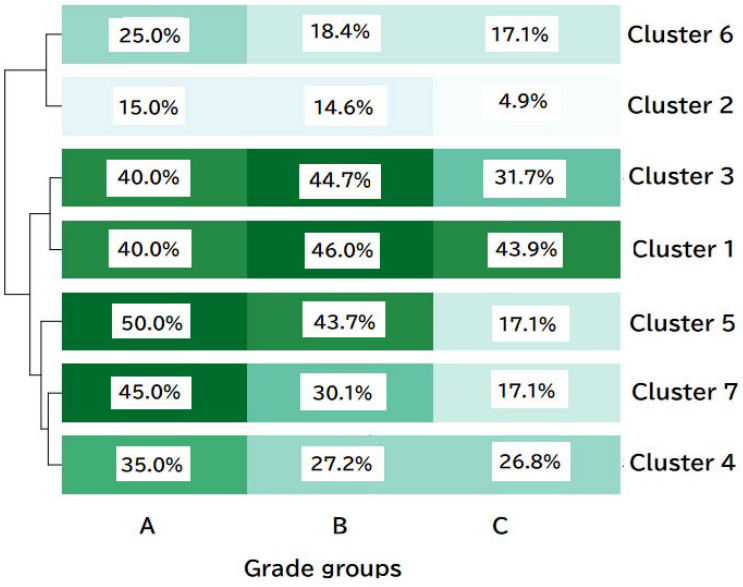


Figure 5

Heat Chart of Cross-tabulation between Word Clusters and Academic Grades



Note. The numbers in the bars show the percentages of students who wrote words that were included in each word cluster within each grade category.

Table 4

The Number of the Students Who Wrote Words that were included in Each Cluster and the Results of the Chi-square Tests among the Three Academic Grades within Each Cluster (N = 164)

Grade	*n*	*Cluster 1*	*Cluster 2*	*Cluster 3*	*Cluster 4*	*Cluster 5*	*Cluster 6*	*Cluster 7*
A	20	8	5	8	7	10	3	9
B	103	48	19	46	28	45	15	31
C	41	18	7	13	11	7	2	7

4 Discussion

This study examined the students' evaluation scores and written comments regarding the usefulness of real-time polling in a large hybrid class. The average score for the evaluation of the use of Respon (2021)

showed that the students well apricated the use of Respon in the class. The result of a one-way ANOVA test showed that academic grade levels did not affect the students' evaluations of the usefulness of Respon. The results of the qualitative question showed that using a real-time polling tool in a hybrid class was generally received well by the students.

The first and second of the ten most frequently written words were not related to the evaluation of real-time polling. As the frequently written words did not provide useful information about the students' feedback, further detailed analyses were necessary to identify hidden information in the students' written comments.

Text mining of the students' written comments in accordance with their academic grades provided useful insights into the relationships among the students' perceptions of using real-time polling and their academic grades. The middle-grade students wrote words that were included in Cluster 5 significantly more frequently than expected, whereas the low-grade students wrote words that were included in Cluster 5 significantly less frequently than expected. This result suggests that the middle-grade students utilized the questions provided by Respon during classes to understand the content of the class, and that the low-grade students did not use the questions in the same way that the middle-grade students did. Fifty percent of the high-grade students also wrote words that were included in Cluster 5. Therefore, it could be assumed that the high-grade students also used Respon to understand the content of the class in the same way that the middle-grade students did. It appears that the low-grade students did not make use of Respon as a tool for understanding the content of the class.

The high-grade students wrote words that were included in Cluster 7 more frequently than expected, whereas the low-grade students wrote words that were included in Cluster 7 less frequently than expected, with an alpha level of .1. Although no statistical significance was obtained, 45.0% of the high-grade students wrote words that were included in Cluster 7, whereas 30.1% of the middle-grade students wrote words that were included in this cluster. This result suggests that only the high-grade students were interested in others and were willing to know others' opinions. Learning from outside information is an important learning strategy (Ichikawa, 2017: Shafto et al., 2012), and it is widely accepted that social interaction is essential in language learning (Rupley & Russell, 1979). No collaborative work was given to the students during the course. However, as the high-grade students appeared to be seeking opportunities to improve their knowledge via interactions with their peers, willingness to learn from others might be an important characteristic of high-grade students.

More than 40% of the students across all three grade levels wrote words that were included in Cluster 1. This result suggests that using real-time polling created a sense of participation in the class for all the students regardless of their academic grades, which is in line with the results of previous research (Menon et al., 2004).

The words that the low-grade students wrote second most frequently were words that were included in Cluster 3, even though the percentage of the students who wrote (31.7%) these words was lower than were the percentages of the high- and middle-grade students who did so, at 40.0% and 44.7%, respectively. It is assumed that real-time polling caused the low-grade students to think to some extent.

In summary, the high- and middle-grade students used real-time polling to improve their learning to a greater extent than did the low-grade students. Having a sense of belonging to the class was appreciated by all the students.

Clusters 2 and 6 included negative feedback about Respon. Cluster 2 reflected complaints about the frequent Chinese character tests in the class. Twenty-five percent of the high-grade students complained, and word "easy" was included in the words in Cluster 2. It is likely that the Chinese character tests were too easy for the high-grade students. This result is in line with Price's (2022) suggestion that teachers should avoid implementing the same activity frequently.

The results of the text mining showed aspects of the students' perceptions of real-time polling that one quantitative question would not have been able to reveal. The findings derived from the text mining in this study will be used as questionnaire items in future research. A limitation of this study was the failure to obtain written comments from 78 students; specifically, 47 of these 78 students were low-grade students. Feedback from low-grade students is important in order to investigate effective educational measures to support unsuccessful students; thus, it is necessary to find ways to obtain feedback from low-grade students.

5 Conclusions

This study evaluated the usefulness of real-time polling in a large hybrid class by using quantitative evaluation scores and text mining of the students' written comments. The result of the quantitative test showed that real-time polling was received well by the students regardless of their academic grades, while

the text mining revealed the relationships underlying the students' comments and their academic grades. The low-grade students wrote words that were related to understanding the content significantly less often than the middle-grade students. The low-grade students did not appear to use real-time polling to improve their learning to the same extent as did the middle-grade students. The high-grade students appeared to be more willing to learn from others than were the middle- and low-grade students. Therefore, willingness to learn from others might be a characteristic of high-grade students. Between 40% and 46% of the students wrote comments including the wording "feeling of participation." Real-time polling appeared to promote student participation in the large hybrid class. The low-grade students appreciated the use of real-time polling to obtain a sense of participation; however, they did not appreciate it as a learning tool. The quantitative questions provided an overall evaluation, while the text mining revealed the students' different perceptions of real-time polling according to their academic grades and provided useful information for developing instruction and for creating questionnaire items to evaluate instruction.

References

Allais, S. (2014). A critical perspective on large class teaching: The political economy of massification and the sociology of knowledge. *Higher Education*, *67*(6), 721-734. https://doi.org/10.1007/s10734-013-9672-2

Buckley, K., Stone, S., Farrell, A. M., Glynn, M., Lowney, R., & Smyth, S. (2021). Learning from student experience: Large, higher education classes transitioning online. *Irish Educational Studies*, *40*(2), 399-40. https://doi.org/10.1080/03323315.2021.1916566

Ferreira-Mello, R., André, M., Pinheiro, A., Costa, E., & Romero, C. (2019). Text mining in education. *Wiley Interdisciplinary Reviews*, *9*(6). https://doi.org/https://doi.org/10.1002/widm.1332

Gronberg, N., Knutas, A., Hynninen, T., & Hujala, M. (2021). Palaute: An online text mining tool for analyzing written student course feedback. *IEEE Access*, *9*, 134518-134529. https://doi.org/10.1109/access.2021.3116425

Ichikawa, S. (2017). Ninchi shinrigaku wa kyoiku ni doo kookensuruka [How cognitive psychology has contributed to education]. *Japanese Psychological Review, 60*(4), 367-378.

KHCoder (2022). *KHCoder* (Version 3 Beta 04f). Computer Software. https://khcoder.net/

Menon, A. S., Moffett, S., Enriquez, M., Martinez, M. M., Dev, P., & Grappone, T. (2004). Audience response made easy: Using personal digital assistants as a classroom polling tool. *Journal of the American Medical Informatics Association: JAMIA, 11*(3), 217-220.

Price, T. J. (2022). Real-time polling to help corral university learners' wandering minds. *Journal of Innovating Teaching & Learning*, *15*(1), 98-109. https://doi.org/10.1080/JRIT-03-2020-0017

Respon (2021). *Respon.* Mobile App Software. https://respon.jp/

Rupley, W. H. & Russell, M. T. (1979). The interaction of language, cognitive, and social development. *Language Arts*, *56*(6), 697-99.

Santhanam, E., Lynch, B. & Jones, J. (2018). Making sense of student feedback using text analysis: Adapting and expanding a common lexicon. *Quality Assurance in Education: An International Perspective*, *26*(1), 60-69. https://doi.org/10.1108/QAE-11-2016-0062

Santhanam, E., Lynch, B., Jones, J. & Davis, J. (2022). From anonymous student feedback to impactful strategies for institutional direction. In E. Zaitseva, B. Tucker, & E. Santhanam (Eds.), *Analyzing student feedback in higher education: Using text-mining to interpret the student voice* (pp. 167-179). Routledge.

Shafto, P., Goodman, N. D., & Frank, M. C. (2012). Learning from others: The consequences of psychological reasoning for human learning. *Perspectives on Psychological Science*, *7*(4), 341-351. https://doi.org/10.1177/1745691612448481

Stanley, D. (2013). Can technology improve large class learning? The case of an upper-division business core class. *Journal of Education for Business*, *88*(5), 265-270. https://doi.org/10.1080/08832323.2012.692735

Steyn, C., Davies, C. & Sambo, D. (2019). Eliciting student feedback for course development: The application of a qualitative course tool evaluation tool among business research students. *Assessment & Evaluation in Higher Education, 26*(1), 11-24 http://dx.doi.org/10.1080/02602938.2018.1466266

Stover, S., Noel, D., McNutt, M., & Heilmann, S. G. (2015). Revisiting Use of real-time polling for learning transfer. *Journal of Teaching and Learning with Technology*, *4*(1), 40-60. https://doi.org/https://doi.org/10.14434/jotlt.v4n1.13002

Stupans, I., McGuren, T., & Babey, A. M. (2016). Student evaluation of teaching: A study exploring student rating instrument free-form text comments. *Innovative Higher Education*, *41*(1), 33-42.

https://doi.org/https://doi.org/10.1007/s10755-015-9328-5
Wright, S. (2022). Automating insights: Analyzing the National Student Survey data using NVivo. In E. Zaitseva, B. Tucker, & E. Santhanam (Eds.), *Analyzing student feedback in higher education: Using text-mining to interpret the student voice* (pp. 37-50). Routledge.
Wu, Y. & Dawson, G. (2022). You articulate, we implement: Adding constructive feedback coaching and automated text analysis in the course evaluation loop. In E. Zaitseva, B. Tucker, & E. Santhanam (Eds.), *Analyzing student feedback in higher education: Using text-mining to interpret the student voice* (pp. 19-36). Routledge.
Zaitseva, E., Santhanam, E., & Tucker, B. (2022). Discovering student experience: beyond numbers through words. In E. Zaitseva, B. Tucker, & E. Santhanam (Eds.), *Analyzing student feedback in higher education: Using text-mining to interpret the student voice* (pp.1-15). Routledge.
Zoom (2022). *Zoom*. Cloud App. https://zoom.us/jp-jp/meetings.html