Lexical Analysis of Student's Learning Activities during the Giving of Instructions for Note-Taking in a Blended Learning Environment

Minoru Nakayama, Kouichi Mutsuura, and Hiroh Yamamoto

***Abstract*—Note-taking activity affects student's learning performances in blended learning courses, which consist of face-to-face sessions and online learning materials. To promote the effectiveness of note-taking, a lecturer gave students instructions during the course. According to the results of a lexical analysis of the contents of notes taken by students, the lecturer's instructions had a significant effect on some of the indices of features of notes taken. This effectiveness can be observed when the lecturer gives instructions as opposed to not giving them. Also, the relationships between student's characteristics and indices of features of content of notes taken were analyzed.**

***Index Terms*—Note taking, blended learning, lexical analysis, learning performance, student's characteristics.**

I. Introduction

The modern learning environment using information communication technologies (ICT) offers various types of learning experiences for university level education and for informal methods of learning [1]. Those environments are easy for students to learn, their learning processes are not simple.

To maintain learning effectiveness in the online learning environment, the analysis of student's actual learning activities is absolutely necessary. For monitoring the learning situation, student's access logs of online materials are often analyzed, these approaches are sometimes effective [2] and sometimes not [3]. This shows that an effective extraction of data features is required.

Though the online learning environment is currently a requirement for many universities world wide, the student behavior through the environment needs to be evaluated, since they imply to present their learning performance. These analyses are currently being conducted in the MOOC (Massive Open Online Course) environment [4], [5]. Regarding this analytical trend, note-taking activity has always been analyzed, as even in the online learning environment note-taking is a key learning activity [6]–[9], since the note-taking activity stimulates constructive learning [10]. In addition, the relationship between some factors of note-taking and learning performance at universities world-wide has been identified in previous studies [8], [9], [11]. Again, note-taking is a time-honoured and commonly used and skill [12].

The results of several surveys which the authors have conducted presented note-taking behaviour contribution to learning performance. In addition to those results, the effectiveness of instructions about note-taking skills given was measured [13]. According to the results, both learning performance and the contents of notes may be improved when students recognize good note-taking skills, as explained by a competent instructor. To confirm this hypothesis, two experimental surveys were conducted.

The following topics are addressed in this paper:

- How does instruction regarding note-taking improve the contents of notes?
- How do students summarize the contents of their notes are according to these instructions?
- How do student's individual characteristics contribute to their note-taking activities?

To respond to these questions, a lexical analysis of the contents of notes taken by students was conducted.

II. Method

A. Courses

The surveys were conducted over two years during Blended Learning courses at a Japanese university. The subject was Information Network and the courses were Bachelor level credit courses. The course consisted of weekly face-to-face sessions for 15 weeks [14]. Participants were also encouraged to take an on-line test (OT) for each session outside of the lecture room, as a function of the learning management system (LMS). Also, a final exam (FE) was given at the end of the course.

B. Note-Taking Instructions

All participants were required to present their notebooks in order to track the progress of their learning. In most sessions, the lecturer reviewed and assessed the contents of notes for 14 weeks. The contents were evaluated using a scale, and the sum of the scores is defined as the individual note-taking assessment score (NT-A). To determine the possibility of improving note-taking activities by having the lecturer give instructions, two survey conditions were developed: with and without instructions. The first year course was conducted without any instructions being given or feedback about notes taken, and this condition is defined as"without instruction". The second year course was conducted twice, with instructions concerning note-taking techniques and the

Manuscript received June 19, 2014; revised September 10, 2014.

M. Nakayama is with the Department of Human System Science and Cradle, Tokyo Institute of Technology, Tokyo, 152-8552 Japan (e-mail: nakayama@cradle.titech.ac.jp).

K. Mutsuura is with the Faculty of Economics, Shinshu University, Matusumoto, Japan (e-mail: mutsuur@shinshu-u.ac.jp).

H. Yamamoto is with Shinshu University, Matstumoto, Japan (e-mail: yama77722jp@ybb.ne.jp).

display of examples of good notes shown at the beginning and mid-point of the course. This condition is defined as "with instruction".

The valid number of participants is 32 for without instruction and 24 for with instruction.

C. Characteristics of Students

Student's characteristics have an impact on their learning activities. Some indices were surveyed previously, during our prior studies. The constructs are: Personality [15], [16], Information Literacy [17] and Learning Experience [18]. Additionally, an inventory of note-taking skills was surveyed to extract three factor scores.

Personality: The personalities of students were measured using a public domain item pool, the International Personality Item Pool (IPIP) inventory [16]. This questionnaire can produce a five factor personality model [15], and measure scores of the five components: "Extroversion" (IPIP-1), "Agreeableness" (IPIP-2), "Conscientiousness" (IPIP-3), "Neuroticism" (IPIP-4) and "Openness to Experience" (IPIP-5).

Information Literacy: Information literacy inventories were defined and developed by Fujii [17]. The survey consists of 32 question items, and 8 factors are extracted. These 8 factors can be summarized as two secondary factors: Operational Skills (IL-1), and Attitudes toward Information Literacy (IL-2) [19].

Learning experience: Students' online learning experiences were measured using a set of questions and three factors identified. Three factors are Factor 1 (LE-F1): Overall Evaluation of the e-learning experience, Factor 2 (LE-F2): Learning Habits, and Factor 3 (LE-F3): Learning Strategies [18].

Note-taking skills: A set of constructs for measuring the note-taking skills of students was surveyed. These note-taking skills may consist of note-taking abilities, attitudes and techniques, andoriginal inventories of these have been developed by the authors [14], [20]. This construct consists of the following three factors: F1: Recognizing note taking functions, F2: Methodology of utilizing notes, and F3: Presentation of notes.

D. Contents of Notes Taken

Student's notes were scanned and stored as image files, and the contents were read and recorded manually as electronic text files. Figures and tables were extracted. The lecturer's hand-written notes to be presented to students during face-to-face sessions were also transformed into electronic text files.

The texts of the student's and the lecturer's notes were lexically analyzed using a Japanese morphological term analysis tool [21]. In this paper, the frequency of nouns each session is calculated. A comparison of the frequencies in students' notes and the lecturer's presentation produces two indices, defined as follows [14], [22]:

- Word ratio: the ratio between the number of terms written and the number of terms given (the number of terms students recorded vs. the number of terms the lecturer presented)
- Coverage: the coverage ratio was calculated as a percentage of the number of terms recorded by students.

As an additional lexical analysis, co-occurring noun terms were surveyed using the following procedure: For example, noun strings A-B and B-C were extracted from the string A-B-C. A pair of nouns appearing consecutively in a sentence is defined as a consequential noun, or a 2-gram.

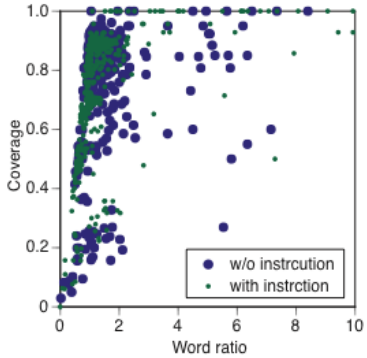

Fig. 1. Relationship between word ratio and coverage of student's notes.

III. Results

A. Relationship between Word Ratio and Coverage

Both word ratios and coverages were calculated for every session in which students took notes. They correlate with each other, as high coverage requires a higher word ratio in comparison with the previous survey [14], [22]. The relationships between the two experimental conditions, with and without note-taking instructions being given, are summarized in Fig. 1. The horizontal axis represents the word ratio, and the horizontal axis represents the coverage. Though there are some deviations in the scatter grams, an overall correlation relationship can be observed. The correlation coefficients are $r = 0.38$ for with instruction and $r = 0.18$ for without instruction. Note-taking instructions that can promote better note-taking behaviour have already been reported in previous surveys by the authors.

To determine the effectiveness of giving instructions, the ratios between the first and the second halves of the sessions were compared since the lecturer advised students at beginning and mid-point of the courses. The change in ratios is summarized in Fig. 2. The horizontal axis represents two groups of sessions in sequence, the horizontal axis represents mean ratios. The coverage remains at the same levels across the two session conditions regardless of whether or not instructions are given. The word ratio increases during the courses, with the increasing rates for with instructions given higher than the ones for without instructions given. If the differences were influenced by the experimental conditions, the instructions affected the word ratios in student's notes.

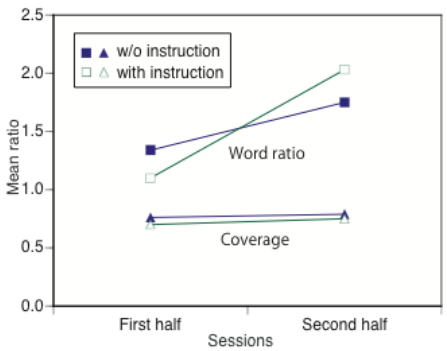

Fig. 2. Comparison of mean ratios of word ratio and coverage between the first and second halves of the course.

$$
\begin{array}{c}
 \\ T1 \\ T2 \\ T3 \\ T4 \\ T5 \\ T6 \\ T7
\end{array}
\begin{array}{c}
\begin{array}{ccccccc} T1 & T2 & T3 & T4 & T5 & T6 & T7 \end{array} \\
\begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}
$$

Fig. 3. An example of an adjacency matrix.

B. Comparison of Adjacency Matrices

Both student's notes and the lecturer's presentation were converted to an adjacency matrix, which indicates the connections between nouns. Fig. 3 contains an example of a matrix of a lecturer's presentation (Session 13).

Generally, students do not record all terms the lecturer presents, though they do record some related, original terms. Therefore, the relationship between lecturers and students in two adjacency matrices can be illustrated in Fig. 4. When students made notes using terms which were not mentioned by the lecturer, the number of terms was larger than the number of terms in the lecturer's notes. The difference between the two matrices represents the behaviour of the note-takers.

The differences can be mathematically measured as edit distances, otherwise known as Levenshtein distances. The distances for sub-matrices L and S were evaluated separately [23]. As a result, the two indices are defined as follows:

- Additional distance means the sum of the number of additional nodes or edges in a matrix.
- Insufficient distance means the sum of the number of reduced nodes or edges in a student's matrix in comparison with the lecturer's matrix.

Both distances are influenced by the total number of terms in the lecturer's presentation, so that the relative distances are calculated using the number of terms the lecture presented in each session.

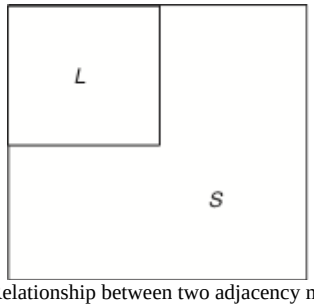

Fig. 4. Relationship between two adjacency matrices.

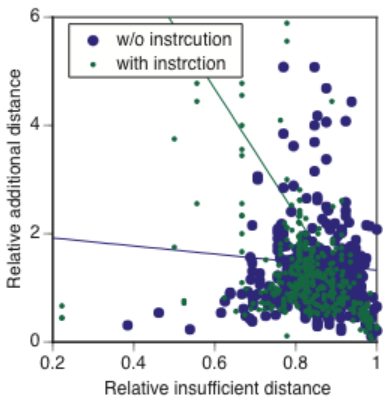

Fig. 5. Relationship between insufficient and additional distances.

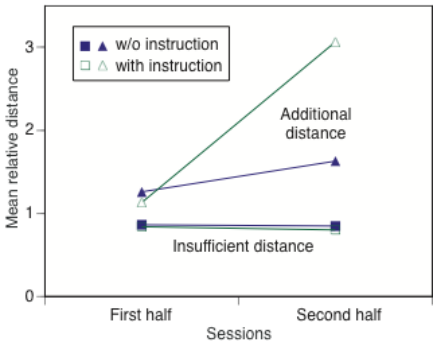

Fig. 6. Comparison of mean distances across sessions.

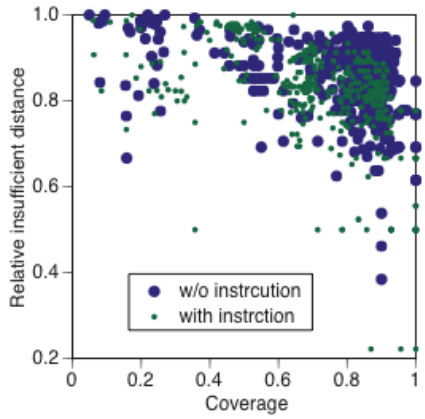

Fig. 7. Relationship between coverage and insufficient distance.

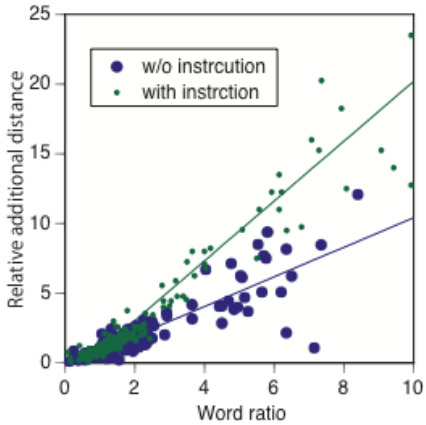

Fig. 8. Relationship between word rate and additional distance.

TABLE I: CORRELATION COEFFICIENTS BETWEEN STUDENT'S CHARACTERISTICS AND INDICES OF TEXT FEATURES OF NOTES TAKEN

	Word rate	Coverage	Insufficient D	Additional D
IPIP-3	-	-	0.43	-
	-	-	-	-
IL-f	-	-	-	-
	-	0.53	-.53	-
IL-s	-	-	-	-
	0.43	-	-	0.41
NT-F1	-	-	0.35	-
	-	-	-	-
NT-F2	-	-	-	-
	0.57	-	-	0.56
NT-F3	0.35	-	-	-
	-	-	-	-
LE-F1	-	-	-	-
	-	-	-	-
LE-F2	-	-.41	0.44	-
	-	-	-	-
LE-F3	-	-	-	-
	-	0.52	-	-
NT-A	0.76	-	0.48	0.79
	0.70	0.36	-	0.56

Upper column line: without instruction
Lower column line: with instruction

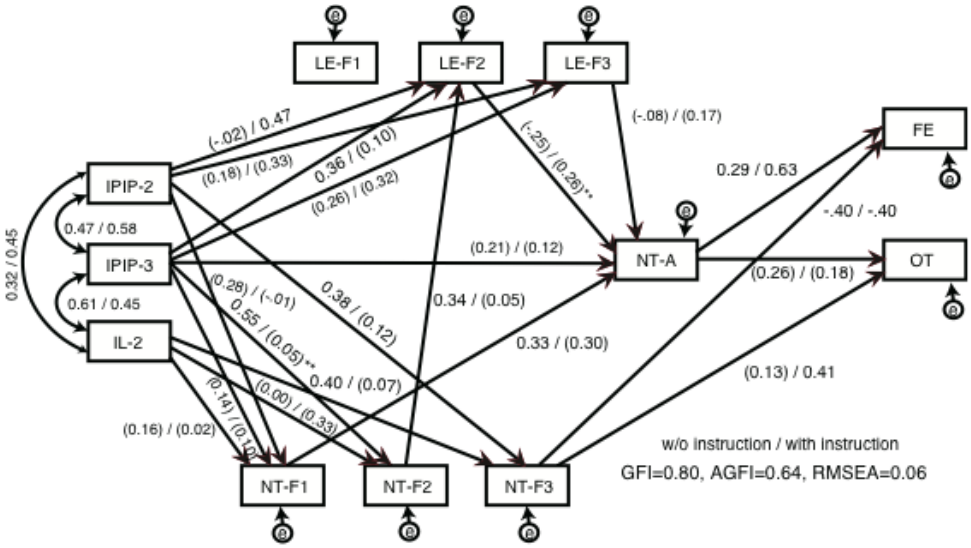

Fig. 9. Causal relationships between student's characteristics and learning performance.

The relationship between the two distances is summarized in Fig. 5. The horizontal axis represents relative insufficient distance and the vertical axis represents relative additional distance. The two lines illustrate the regression relationship between the two distances. Regarding these calculations, the correlation coefficients are r = -0.45 for with instruction and r = 0.04 for without instruction. The correlation can illustrate new terms written as replacements as additional distances instead of omitting these terms as being insufficient distances when the concurrent appearance of terms is considered. Both word rate and coverage show the number of terms the lecturer presented during the session which students took. The distance metric suggests that connections between terms create concepts, and term replacement activities may reflect better understanding of the contents of the information provided by the lecturer.

To determine the effectiveness of instructions given, mean distances across the sessions were compared. The results are summarized in Fig. 6, in the same format as in Fig. 2. When instructions were given to students, mean additional distances increased significantly in the second halves of the sessions. Again, the additional distance shows additional recording of their own new terms, when the connection with the terms the lecturer presented is considered.

C. Relationship between Term Ratio and Distances of Adjacency Matrices

Both word ratio and coverage are easily calculated using term frequency, without considering the co-appearance of other terms. To confirm the degree of occurrence, the relationships between them were analyzed. First, the relationship between coverage and insufficient distance was evaluated. When students correctly recorded terms in their own notes or recorded term networks, the insufficient distances are shorter, though the coverage increases. But both coverage and insufficient distance increase when students chose to replace the presented terms with others. These relationships are illustrated in Fig. 7. The horizontal axis represents coverage and the vertical axis represents insufficient distance. The plots illustrate the total number of negative correlation relationships.The correlation coefficients are calculated as r = -0.51 for with instruction and r = -0.33 for without instruction. The absolute value for with instruction is larger than the one for without instruction. Instructions being given can make participants pay greater attention to recording term network terms.

Second, the relationship between word rate and additional distance is evaluated. The word ratio reflects student's writing activities. The additional distances also increase when these terms are linked with other terms. The relationship is illustrated in Fig. 8. The horizontal axis represents word rate and the vertical axis represents additional distance. As the regression lines in the figure illustrate, they are linearly correlated, and the correlation coefficients are r = 0.97 for with instruction and r = 0.90 for without instruction. Regarding the slopes of the regression lines, recording new words contributes to increases in additional distances when instructions have been given.

These results show that note-taking instructions for students can encourage them to create term networks.

D. Causal Relationship between Learning Performance and Student's Characteristics

The above analyses confirm that giving note-taking instructions improves the effectiveness of student's note-taking activities. Previous studies suggest that a participant's characteristics may also affect their performance. To examine the effectiveness of the lecturer's instructions and the participant's characteristics, a causal analysis using a structural equation modeling technique was introduced. There are many indices of student's characteristics, and some factors were extracted step by step, using correlation analysis. All parameters were estimated using structural equation modelling (SEM) software (AMOS).

As a result, a causal path can be created, as in Figure 9. The paths consist of some factors of personality, information literacy, note-taking skills, learning experience and learning performance, including note-taking scores. Regarding the indices of fitness of the model (GFI: Goodness of Fitting index) (GFI=0.80, AGFI=0.64, RMSEA=0.08), a possible relationship is exhibited. The arrows indicate paths between variables, and path coefficients with and without note-taking instructions being given are indicated. Non significant coefficients are indicated using parentheses. By comparing the coefficients between with and without instruction,

significant differences between the two paths are revealed: IPIP-3(Conscientiousness) to NT-F2(Methodology of utilizing notes), and LE-F2(Learning habits) to NT-A (Note-taking assessment). Though the coefficients are significant for the condition without instruction, some coefficients changed to not significant for the condition with instruction. The instructions given about note-taking techniques may affect note-taking activity to a greater extent than the student own characteristics. This result is evidence of improvement in student's note-taking activities due to the lecturer's instructions. There are no significant changes in the coefficients for the final exams (FE), though there are some significant changes in coefficients in relation to Note-taking assessment (NT-A). How to give more effective instructions and suggestions related to this will be subjects of our further study.

To determine the relationship between student's characteristics and note-taking activity, correlation coefficients between these were calculated. The significant coefficients are summarized in Table I. As mentioned in the above section, there are significant correlations between coverage and insufficient distance, and between word ratio and additional distance. Therefore, significant coefficients regarding those relationships appear simultaneously when instructions are provided. The results of correlation analysis suggest that student's characteristics affect the indices of text features of notes taken. Regarding this result, a possible causal relationship should be identified, and effective instructions for better note-taking should be developed. These will be subjects of our further study.

IV. Conclusion

To improve student's learning performance in a blended learning course, the effectiveness of the lecturer's note-taking instructions was evaluated using lexical analysis of the contents of student's notes.

Regarding the results of our analysis, the following points have been extracted.

1) The results of lexical analysis of student's notes suggest that word ratio, which means the ratio of the number of terms students recorded against the number of terms the lecturer presented, increased while the additional distances, which consider the co-appearance of terms, also increased in the with instruction condition.
2) The lecturer's instructions had significant correlation relationships with the indices of note-taking features instead of when no instructions were given.
3) The causal paths between student's characteristics and learning performance were analyzed using the SEM technique. The lecturer's instructions affected the causal relationship between student's characteristics and note-taking. Also, student's characteristics correlated with indices of the content of features of notes taken.

More effective instructions and additional suggestions about improving learning activities will be subjects of our further study.

Those relationships should be confirmed in other types of online learning environments, such as fully online course and courses using social media. Also, the development of improving instruction using these results will be a subject of our further study.

Acknowledgment

This research was partially supported by the Japan Society for the Promotion of Science (JSPS), Grant-in-Aid for Scientific Research (B-26282046: 2014-2016).

References

[1] M. Nakayama and R. Santiago, "Two categories of e-learning in Japan," *Educational Technology Research and Development*, vol. 52, no. 3, pp. 100–111, 2004.

[2] M. Ueno, "Online outlier detection for e-learning time data," *IEICE Transaction*, vol. J90-D, no. 1, pp. 40–51, 2007.

[3] M. Nakayama, H. Kanazawa, and H. Yamamoto, "Detecting incomplete learners in a blended learning environment among Japanese university students," *International Journal of Emerging Technology in Learning*, vol. 4, no. 1, pp. 47–51, 2009.

[4] D. T. Seaton, S. Nesterko, T. Mullaney, J. Reich, and A. Ho, "Characterizing video use in the catalogue of MITx MOOCs," *E-Learning Papers*, no. 37, pp. 33–41, 2014.

[5] D. T. Seaton, Y. Bergner, I. Chuang, P. Mitros, and D. E. Pritchard, "Who does what in a massive open online course?" *Communication of the ACM*, vol. 57, no. 4, pp. 58–65, 2014.

[6] K. A. Kiewra, "Students' note-taking behaviors and the efficacy ofproviding the instructor's notes for review," *Contemporary Educational Psychology*, vol. 10, pp. 378–386, 1985.

[7] K. A. Kiewra, "A review of note-taking: The encoding-storage paradigm and beyond," *Educational Psychology Review*, vol. 1, no. 2, pp. 147–172, 1989.

[8] K. A. Kiewra, S. L. Benton, S.-I. Kim, N. Risch, and M. Christensen, "Effects of note-taking format and study technique on recall and relational performance," *Contemporary Educational Psychology*, vol. 20, pp. 172–187, 1995.

[9] K. Kobayashi, "What limits the encoding effect of note-taking? Ametaanalytic examination," *Contemporary Educational Psychology*, vol. 30, pp. 242–262, 2005.

[10] P. Tynajä, "Towards expert knowledge? A comparison between a constructivist and a traditional learning environment in the university," *International Journal of Educational Research*, vol. 31, pp. 357–442, 1999.

[11] P. A. Nye, T. J. Crooks, M. Powley, and G. Tripp, "Student notetaking related to university examination performance," *Higher Education*, vol. 13, pp. 85–97, 1984.

[12] P. Weener, "Note taking and student verbalization as instrumental learning activities," *Instructional Science*, vol. 3, pp. 51–74, 1974.

[13] M. Nakayama, K. Mutsuura, and H. Yamamoto, "Effectiveness of instructional suggestions for note-taking skills in a blended learning environment," in *Proc. 12th European Conference of ELearning*, Nice, France, 2013, pp. 333–339.

[14] M. Nakayama, K. Mutsuura, and H. Yamamoto, "Evaluation of student's notes in a blended learning course," *International Journal of New Computer Architectures and their Applications*, vol. 1, no. 4, pp. 1080–1089, 2011.

[15] L. Goldberg, "A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models," *Personality Psychology in Europe*, vol. 7, pp. 7–28, 1999.

[16] IPIP. (2004) A scientific collaboratory for the development of advanced measures of personality traits and other individual differences. *International Personality Item Pool*. [Online]. Available: http://ipip.ori.org

[17] Y. Fujii, "Development of a scale to evaluate the information literacy level of young people – Comparison of junior high school students in Japan and Northern Europe," *Japan Journal of Educational Technology*, vol. 30, no. 4, pp. 387–395, 2007.

[18] M. Nakayama, H. Yamamoto, and R. Santiago, "The impact of learner characteristics on learning performance in hybrid courses among Japanese students," *The Electronic Journal of e-Learning*, vol. 5, no. 3, pp. 195–206, 2007.

[19] M. Nakayama, H. Yamamoto, and R. Santiago, "Impact of information literacy and learner characteristics on learning behavior of Japanese students in online courses," *International Journal of Case Method Research & Application*, vol. XX, no. 4, pp. 403–415, 2008.

[20] M. Nakayama, K. Mutsuura, and H. Yamamoto, "Causal analysis of student's characteristics of note-taking activities and learning performance during a fully online course," in *Proc. 2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communication*, Liverpool, UK, 2012, pp. 1924–1929.

[21] MeCab: Yet another part-of-speech and morphological analyzer. [Online]. Available: http://mecab.sourceforge.net
[22] M. Nakayama, K. Mutsuura, and H. Yamamoto, "Visualization analysis of student's notes taken in a fully online learning environment," in *Proc. 16th International Conference of Information Visualisation*, Montperier, France, 2012, pp. 434–439.
[23] M. Nakayama, K. Mutsuura, and H. Yamamoto, "A note taking evaluation index using term networks in a blended learning environment," in *Proc. 2014 Eighth International Conference on Complex, Intelligent and Software Intensive System*, Birmingham, UK, 2014, pp. 486–490.

Minoru Nakayama is a professor at Human System Science and Cradle, Tokyo Institute of Technology, Japan. He graduated from Tokyo Gakugei University in 1983 and he completed his M.E. program in 1985, and received his PhD of Eng. degree from Tokyo Institute of Technology in 1989. His research concerns educational technology.

Kouichi Mutsuura was born in Tokushima, Japan in 1950. He received the B.E. and M.E. degrees in communication engineering from the University of Electro-communications in 1974 and 1976, respectively, and Ph.D. from Osaka University in 1992. From 1976 to 1992, he has been with Takuma National College of Technology. Since 1992 he has been with Shinshu University. He is currently a professor in the Faculty of Economics at Shinshu University. He is engaged in research on random signal processing and packet radio networks. Dr. Mutsuura is a member of the Institute of Electronics Information and Communication Engineers, and Japanese Society for Information and Systems in Education.

Hiroh Yamamoto received his B.E. degree in communication engineering from Shinshu University, Nagano, Japan in 1966. He received his PhD of engineering degree from Tokyo Institute of Technology in 1999. From 1966 to 2002, he has been in Hitachi Electronics Services Co., Ltd. He experienced a director of the company. Since 2002 he has been with Shinshu University, where he was a professor of Research Center for Education Programs. He is currently professor emeritus of the university. His research interest is focused on learning assessment. He is a member of JSET (Japan Society for Educational Technology) *et al.*