

Measures to Promote Practice of Quiz and Evaluation Thereof in Blended Learning

<http://dx.doi.org/10.3991/ijet.v9i5.3854>

H. Yamamoto, M. Nakayama and Y. Shimizu
Tokyo Institute of Technology, Tokyo, Japan

Abstract—To improve learning outcomes from e-Learning, the authors created an environment where quizzes could be practiced outside class hours in Blended-Learning. As a result of instructions for the active usage of the quiz with practice promoting measures in which results of the quizzes practiced in the term were added to the record for accreditation, there were observed increase in the percentage of participation in the quizzes and substantial increase in the percentage of paper submission. Besides, the results of the written term examination were improved, not only in results concerning knowledge and understanding in threefold choice and fill-in-the-blanks questions, but also in evaluation of “Self-assertiveness,” “Multiplicity,” and “Pictorial explanation” in short essays, in particular. In the comparison according to questions of the term examination between the higher-ranking and lower-ranking groups in scores in the quizzes, the higher-ranking group achieved the better result in each of the questions. There was a positive correlation between the total number of quiz practices and the score therein. It was also suggested that the quizzes were independently practiced.

Index Terms—Blended-learning, Learning performance comparison, Quiz, Higher-education, e-Learning

I. INTRODUCTION

Blended-Learning, which is a learning system having a combination of classroom education and e-Learning, has resulted in enhancement in learning effect, e.g., with development in understanding by utilization of contents in addition to conventional face-to-face class [1][2][3]. Blended-Learning, having variations, has been recently attracting the attention in fields including higher education of commuting system [4]. Nevertheless, Blended-Learning often requires scrupulous guidance to students, follow-up to learning, etc [1][5].

A report from the University Council of Japan chiefly pointed out the insufficiency in learning time off campus, saying, “The purport of the credit system, ‘One credit shall normally be organized to contain 45-hour learning inside and outside a classroom,’ has not been thorough enough, and the ingenuity in instruction for ensuring the learning of students inside and outside classrooms has been insufficient. [6]” In 2008, the Central Council for Education submitted a report titled “Towards the enhancement of undergraduate education” stating that it was a task to promote independent learning by saying, “It is a critical problem how to stimulate students poor in willingness to learn and sense of purpose in order to make them have an orientation and an attitude toward independent learning. [7]”

After that, the universities engaging in education on the front lines have reported the situation in Japan where the

independent learning is extremely scarce, etc., based on the comparison of learning time except class hours between Japan and overseas countries. In addition, they have pointed out the necessity to enhance the motivation by formation of willingness to learn and have proposed that such classes require specific tools, practical methods and the like [8].

Generally, records for accreditation in university classes are often evaluated on basis of only results of term examinations. Results of questionnaires to students, however, have demonstrated that many students think, “The records should not be determined on basis of a single examination or a single assignment. [9]”

In view of such background, the purposes of the research were set as follows:

- a. To build a Blended-Learning system and to implement it in the university education;
- b. To measure concrete effects of instructing active use of the system as the practice promoting measures for enhancing the attitudes of students to learning; and
- c. To clarify the numerical effects thereof on basis of the percentages of participation in the quizzes, the percentages of paper submission, the results of the term examination, the comprehensive evaluation in the quizzes and the term examination, etc.

As a result, significant increases were observed in the percentages of participation in the quizzes, the percentages of paper submission, and the results of the term examination. In the relation between the quizzes and the term examination, the higher-ranking group in the quizzes gained significantly higher scores in all the questions of the term examination. It was also suggested that the instructions for the active usage resulted in continued and independent practice of the quizzes. Further, it was found that attempts aiming at scores not less than 80 points in the quizzes were carried out even with increase in the number of the practices.

The characteristics of tests in e-Learning are as follows:

1. Easiness of practice even at home (particularly for review)
2. Availability of real-time automatic marking (Feedback-motivation)
3. Unlimitedness in the number of practices (Self-motivation)
4. The final score would be retained as a record (Successful-feeling-motivation)
5. Variability in questions at each time (prevention of cheating)

PAPER
MEASURES TO PROMOTE PRACTICE OF QUIZ AND EVALUATION THEREOF IN BLENDED-LEARNING

TABLE I.
DIFFERENCES IN PRACTICE BETWEEN USAGE INSTRUCTED G AND NON-INSTRUCTED G

	Item	Usage instructed G	Non-instructed G
Quizzes	1) Taken into consideration in the accreditation	Underwent description on the intention to do so	No particular notice
	2) Possible number of the practices for each session	Try any number of times until you are satisfied	
	3) Record of results	The final score would be retained as a record	
	4) Number of presented questions	100 scores in total for 10 quizzes, and the total point was 1000	
	5) Method of presentation	10 questions were presented at random from among 30 questions	
6)	Contents of class and number of sessions	"Information society and occupation", 15 in total	
7)	Contents for preparation and review	Could be learned at any time and even outside the campus	
8)	Written term examination	Questions were completely identical, 90 minutes, 100 perfect scores	

II. IMPLEMENTED BLENDED-LEARNING SYSTEM

A. Overview of system

In this research, the authors initially arranged an environment in which simple quizzes could be practiced with use of ICT and built up a system by which the results could immediately be fed back to the students. The arrangement was made such that the students could take the quizzes any number of times and such that the results of the practices were recorded finally. The required time for each quiz was on the order of ten minutes. The environment was built in which images, sounds and pictures of the classes were converted into contents and stored in a server and in which learning could be carried out at any hour as required according to the results of the quizzes. The system with use of ICT was the Learning management System developed by Blackboard Inc.

B. "Usage instructed G" and "Non-instructed G" (G is abbreviation of Group)

As written in the report from the Central Council for Education described in the Introduction, the stimulation of students poor in willingness to learn and sense of purpose is an important factor. In the research, therefore, "Usage instructed G" and "Non-instructed G" were set and the effects of the instructions to the students were measured by comparison between both. Table 1. shows major differences between them.

In Table I.,

1) As to whether the results of the quizzes would be "taken into consideration in the accreditation," (Reward-motivation) the "Usage instructed G" underwent description on the intention to do so and the repeated instructions for the practice of the quiz until the fourth lesson of the class (Emphasis of motivation). Thereafter, however, the instructions for the practice were not actively given in particular. The "Non-instructed G" received no particular notice about the records for the accreditation but were told that they had better practice the quizzes for better understanding of the contents of the class.

2) Regarding "Possible number of the practices for each session," (Self-motivation) the "Usage instructed G" were explained that

the quizzes "can be practiced any number of times until you are satisfied." The "Non-instructed G" received no particular notice about that.

3) Regarding "Record of results" of the quizzes, the "Usage instructed G" were explained that the final score in the practices would be retained as a record (Successful-feeling-motivation) and that the students could refer to their scores any time (Feedback-motivation). The "Non-instructed G" received no particular notice about that.

4) The "Number of presented questions" for each group was one hundred in total for ten quizzes, and the maximum number of total points was 1000.

5) In the "Method of presentation," ten questions were presented at random from among 30 questions in each session of the class. When two students simultaneously practice a quiz, accordingly, the same questions are not presented to them. When one student opens the second quiz, questions different from those of the first are presented.

6) The "Contents of class and number of sessions" were "Information society and occupation" and fifteen in total, including the term examination, respectively, for both "Usage instructed G" and "Non-instructed G."

7) The "Contents for preparation and review" could be learned at any time and even outside the campus for both groups.

8) For both groups, the "Written term examination" had the duration of 90 minutes and consisted of four questions in general. The maximum number of total points was 100 with each question having 25 points. The questions for both groups were completely identical.

"Usage instructed G" is the group where, up until the fourth lesson, students were repeatedly instructed to perform the quiz in order to effectively motivate them and encourage follow-through. "Non-instructed G", on the other hand, were provided with an e-Learning motivation feature, but for that group motivation and follow-through were not emphasized.

Both "Usage instructed G" and "Non-instructed G" were originally supposed to have face-to-face classes, and the system that had already been used in classes using only e-Learning was diverted to the Blended-Learning.

The relevant class was "Information society and occupation" with two credits, intended for first-year undergraduates. The class was an elective course that could voluntarily be taken by the students in both groups. The students were divided into the two groups without special filtering, conditioning, etc. The students were not informed of the discrimination between "Usage instructed G" and "Non-instructed G."

III. RESULTS OF EFFECTS OF INSTRUCTIONS

Evaluation was made regarding the percentages of participation in the quizzes, the percentages of paper submission, etc. that related to general learning activities, the results of the written term examination, the relation between the quizzes and the results of the term examination, and the situation of independent learning for the quizzes. This section will describe the methods of the evaluation and the results thereof.

A. Total evaluation of learning activities

Herein, the total evaluation was made concerning the percentages of participation in the quizzes, the percentages of paper submission, and the results of the written term examination. Table 2. shows the results thereof.

As shown in Table II., the numbers of the registration were 40 in "Usage instructed G" and 37 in "Non-instructed G."

A pretest was administered using simple threefold choice questions. The results were 80.6 points in "Usage instructed G" and 81.4 points in "Non-instructed G." As t-value was 0.18, no significant difference was found between both groups. This confirmed that there was no difference in the results of the pretest between "Usage instructed G" and "Non-instructed G."

The ratio of the number of times of quiz participation refers to the ratio of the total number of times of the actual practice by the students (numerator) to the total number of the quizzes given ten times in all the ten lectures (denominator). The ratio of 369/400 (92.3%) in "Usage instructed G" was higher than that of 300/370 (81.1%) in "Non-instructed G." The result of chi-square test was $p < 0.01$.

The ratio of the number of submitted papers refers to the percentage of submission calculated on the condition that two papers each on an A4 sheet were assigned to submit in the term. The subjects of the papers and the method of the practice were common to both groups. The "Usage instructed G" with the ratio of 80/80 (100%) attained the perfect result in which all of them submitted two times, in contrast to the "Non-instructed G" with the ratio of 65/74 (87.8%). The result of chi-square test was $p < 0.01$.

The average score of the written term examination refers to the average score per one student. The added points based on the results of the quizzes and "taken into consideration in the accreditation" were not included in the results of the term examination. Used as the records of the written term examination were the results of marking by experts other than the teachers in charge. Its details will be described in the next section "3.2 Contents prepared for written term examination and method of marking" and later.

B. Contents prepared for written term examination and method of marking

The written examination with the duration of 90 minutes was given to both groups at the end of the term. In the examination, notebooks, PCs, etc. were not allowed and the students were made to write the papers by hand using pencils. The examination consisted of four questions including two threefold choice and fill-in-the-blanks questions of so-called short-answer style for objective evaluation concerning knowledge and understanding and two

TABLE II.
COMPARISON OF TOTAL EVALUATION REGARDING LEARNING ACTIVITIES

	Usage instructed G	Non-instructed G	Difference t Test Judgment
Numbers of the registration	40	37	+3
Pretest(points)	80.6	81.4	-0.8 t=0.18 n.s.
The ratio of the number of times of quiz	369/400 (92.3%)	300/370 (81.1%)	$\chi^2=21.0$ $p < 0.01$
The ratio of the number of submitted papers	80/80 (100%)	65/74 (87.8%)	$\chi^2=10.3$ $p < 0.01$
The average score of the written term examination	81.0	68.0	+13.0 t=3.95 $p < 0.01$

Note: The denominators of ratios of number of times of quiz participation and of ratios of number of submitted papers are total numbers.

short essays of so-called verbal problem style for evaluation as to whether description was given based on multifaceted consideration, self-assertiveness, etc.

For the marking of the short essays, two experts were requested to independently make the evaluation from five viewpoints in order to ensure objectivity.

1) Prepared contents and aims thereof

Important contents prepared for the written term examination and the aims thereof were as follows [10]:

Question 1 was for asking about basic knowledge on the learned contents. One correct answer was to be chosen from three alternatives.

Question 2 consisted of 15 small questions asking about technical knowledge on information technology. In each small question, three blanks were to be filled in.

Question 3 was for asking how the answerers would like to utilize information technology on the assumption that they got into the real world and whether they could describe their own thoughts clearly, specifically and multifacetedly with use of pictures or drawings.

Question 4 was for estimating insight and inventiveness for perceiving trends of development of information technology.

The questions were aimed at such that most students could finish answering with seriousness within 90 minutes. In order that the leakage of the questions to the outside might be prevented, the question sheets and the answer sheets were integrated and the sheets for all the examinees were collected after the examination was over. Consideration with the increase in the number of questions was given such that bias with respect to the overall scope of the class might be prevented and such that the students had difficulty in keeping the contents of the questions until the next year.

2) Method of marking and allocation of marks

The marking was carried out in objective and unbiased manners as follows:

1) For the threefold choice questions of Question 1 and the fill-in-the-blanks questions of Question 2, correct/incorrect answers could undoubtedly be determined and the marking was therefore done according to the list of errata.

2) For the short essays of Questions 3 and 4, the two experts were requested to independently carry out the marking.

3) The short essays were evaluated from the five viewpoints for each question. The criteria for marking the short essays were as follows:

- a. Compliance with preparation: The contents of the short essays are in compliance with the intention of the preparation?
- b. Learned contents: The learned contents are included therein?
- c. Self-assertiveness: Opinions of the learner were asserted therein?
- d. Multiplicity: The contents have been described from multifaceted viewpoints?
- e. Pictorial explanation: Pictorial expressions and/or the like are included therein?

4) As for the allocation of the marks, 25 points were allocated to each of Questions 1 through 4, so that the maximum number of total points was 100.

3) *Correlation regarding viewpoints for evaluation between two experts*

For the short essays, the two experts were requested to independently carry out the marking in order to ensure objectivity. A correlation coefficient was calculated in order to determine whether there was no great divergence between their marking results.

Table III. shows the consequence. In the research, the evaluation was made in terms of the items that exhibited $r=0.4$, i.e., presumably moderate correlation coefficient, or greater values.

C. *Comparison of question-by-question scores of written term examination*

Table IV shows the results of the comparison of the question-by-question scores of the written term examination.

Compared in Table IV were the question-by-question breakdown scores of “term examination” in the Comparison of total evaluation regarding learning activities of Table II.

In the threefold choice questions of Question 1, “Usage instructed G” scored 22.8 points that was +2.8 relative to 20.0 scored by “Non-instructed G.” The level of significance was 1%. For the questions of Question 1 asking about the basic knowledge, both groups exhibited high percentages of correct answers not lower than 80%.

In the fill-in-the-blanks questions of Question 2, “Usage instructed G” scored 20.3 points with the difference of +3.0 from 17.3 scored by “Non-instructed G.” The level of significance was 1% as in Question 1.

In the short essay of Question 3, “Usage instructed G” scored 20.6 points with the difference of +4.6 from 16.0 scored by “Non-instructed G.” The level of significance was 1% as in Questions 1 and 2.

In the short essay of Question 4, “Usage instructed G” scored 17.3 points with the difference of +2.6 from 14.7 scored by “Non-instructed G.” The level of significance was at 1%.

In total, “Usage instructed G” scored 81.0 points with the difference of +13.0 from 68.0 scored by “Non-instruc-

TABLE III.
CORRELATION IN EVALUATION SCORES BETWEEN TWO EXPERTS

	correlation coefficient	significance level
a) Compliance with preparation	0.55	p<0.001
b) Learned contents	0.22	P<0.05
c) Self-assertiveness	0.47	p<0.001
d) Multiplicity	0.41	p<0.001
e) Pictorial explanation	0.89	p<0.001
sum total	0.85	p<0.001

TABLE IV.
COMPARISON OF RESULTS OF TERM EXAMINATION

	Usage instructed G n=40	Non-instructed G n=37	Difference t Test Judgment
Question 1 Threefold choice	22.8 (1.61)	20.0 (4.63)	+2.8 t=3.48 p<0.01
Question 2 Fill-in-the-blanks	20.3 (3.73)	17.3 (4.69)	+3.0 t=3.04 p<0.01
Question 3 Short essay 1	20.6 (4.26)	16.0 (4.96)	+4.6 t=4.29 p<0.01
Question 4 Short essay 2	17.3 (4.48)	14.7 (4.25)	+2.6 t=2.58 p<0.01
Sum total	81.0 (7.80)	68.0 (15.4)	+13.0 t=4.57 p<0.01

Numerical values designate average scores (standard deviation)

ted G.” A significant difference was observed at the level of significance of 1%.

It was found from the above that “Usage instructed G” yielded significantly higher results than “Non-instructed G” did, in all of the Questions 1 through 4.

D. *Comparison of short essays in written term examination according to viewpoints*

For the comparison of the short essays according to the evaluation viewpoints, breakdowns of the marking of Questions 3 and 4 were shown in Table V. “Comparison of short essays according to marking viewpoints”.

For “Compliance with preparation,” “Usage instructed G” scored 10.7 points with the difference of +0.3 from 10.4 scored by “Non-instructed G,” whereas there was observed no significant difference. Both groups answered in line with the purport of the preparation and made the high scores.

For each of “Self-assertiveness,” “Multiplicity,” and “Pictorial explanation,” the score made by “Usage instructed G” was higher than that made by “Non-instructed G.” By t-test, differences were observed in “Self-assertiveness” and “Multiplicity” at the significance level of 5% and in “Pictorial explanation” and “Total” at the significance level of 1%.

E. Relation between quizzes and learning results

The scores made by “Usage instructed G” in the written term examination were significantly higher than those made by “Non-instructed G.” It was investigated whether factors therein related to the quizzes.

Inspected therein were the question-by-question scores in the written term examination that had been achieved by the higher-ranking group and the lower-ranking group in the scores of the quizzes. Table VI shows the results thereof.

In the threefold choice questions of Question 1, as seen from Table VI, the higher-ranking group made +2.7 points relative to the lower-ranking group, and the level of significance was 1%. In the fill-in-the-blanks questions of Question 2, the higher group made +2.0 points relative to the lower group, and the level of significance was 5%.

In the short essay of Question 3, the higher group made 19.6 points, which was +2.5 relative to 17.1 of the lower group, and the level of significance was 5%. In the short essay of Question 4, the higher group made 17.0 points, which was +2.0 relative to 15.0 of the lower group, and the level of significance was 5%. In total, the higher group in the quizzes made 79.2 points, which was +9.2 relative to 70.0 of the lower group, and the level of significance was 1%. It was found that the higher-ranking group in the quiz scores achieved the higher scores in the written term examination and that there was a relation therebetween. That is, it was observed that the higher the scores of the quizzes the higher the question-by-question scores in the written term examination.

F. Relation between number of quiz practices and average score

In the quiz system, a learner may practice the quiz any number of times until being satisfied. Some students stop the practice after the first time in each session of the class by being satisfied with their achievement of 80 or 90 points, whereas some other students practice the quiz a number of times until achieving 80 or 100 points without satisfaction with their scores of 70 points or lower. Only the “Usage instructed G” was divided into higher-ranking group and lower-ranking group according to the average values of the number of quiz practices, and the average scores of both groups were investigated. Table VII. shows the results thereof.

As seen from Table VII., the average number of quiz practices of 20 students in the higher-ranking group was 7.43, and that of 20 students in the lower-ranking group was 2.30. Thus there was the difference as great as 5.13. The average scores in the quizzes were 90.2 points in the higher group and 78.6 in the lower group, and there was the difference of +11.6. In both the number and the score, the differences were significant at the significance level of 1%.

Even in only “Usage instructed G,” it was understood that the score in the quizzes of the higher group with the larger number of quiz practices was higher than that of the lower group with the smaller number. Fig. 1. shows the correlation between the total number of quiz practices and the score.

The correlation coefficient in Fig. 1. was $r=0.51$, $p<0.01$, showing that there was a positive correlation

TABLE V.
COMPARISON OF SHORT ESSAYS ACCORDING TO MARKING VIEWPOINTS

	Usage instructed G n=40	Non-instructed G n=37	Difference t Test Judgment
Compliance with preparation	10.7(1.90)	10.4(2.44)	+0.3, t=0.46, n.s.
Self-assertiveness	9.8(2.23)	8.7(2.43)	+1.1, t=2.00, p<0.05
Multiplicity	9.4(2.25)	8.0(3.02)	+1.4, t=2.28, p<0.05
Pictorial explanation	8.0(3.07)	3.6(3.70)	+4.4, t=5.68, p<0.01
sum total (full marks 50)	37.9(7.19)	30.7(8.48)	+7.2, t=3.95, p<0.01

Numerical values designate average scores (standard deviation)

TABLE VI.
COMPARISON AS TO TERM EXAMINATION BETWEEN HIGHER- AND LOWER-RANKING GROUPS IN QUIZ SCORE

	Scores of quiz		Difference t Test Judgment
	higher-ranking groups n=39	lower-ranking groups n=38	
Question 1 Threefold choice	22.8 (1.54)	20.1 (4.70)	+2.7 t=3.41 p<0.01
Question 2 Fill-in-the-blanks	19.8 (3.83)	17.8 (4.88)	+2.0 t=1.85 p<0.05
Question 3 Short essay 1	19.6 (4.67)	17.1 (5.81)	+2.5 t=2.18 p<0.05
Question 4 Short essay 2	17.0 (3.78)	15.0 (5.05)	+2.0 t=1.90 p<0.05
Sum total	79.2 (9.18)	70.0 (15.9)	+9.2 t=3.03 p<0.01

Numerical values designate average scores (standard deviation)

TABLE VII.
COMPARISON OF NUMBER OF PRACTICES AND SCORES BETWEEN HIGHER- AND LOWER-RANKING GROUPS AS TO NUMBER OF QUIZ PRACTICES [SUBJECT: USAGE INSTRUCTED G]

	higher-ranking groups n=20	lower-ranking groups n=20	Difference t Test Judgment
Average practices	7.43 (4.83)	2.30 (1.23)	+5.13 t=4.60 p<0.01
Average scores	90.2 (7.76)	78.6 (11.9)	+11.6 t=3.67 p<0.01

Numerical values designate number of practices and average scores (standard deviation)

between the number of quiz practices and the score. As seen from the figure, there was a case that was solely isolated with the total number of quiz practices greater than 250. The correlation coefficient calculated with the case excluded as an outlier was $r=0.62$, $p<0.01$. It can be said that the efforts of the students in the practice of the quizzes were rewarded with the results.

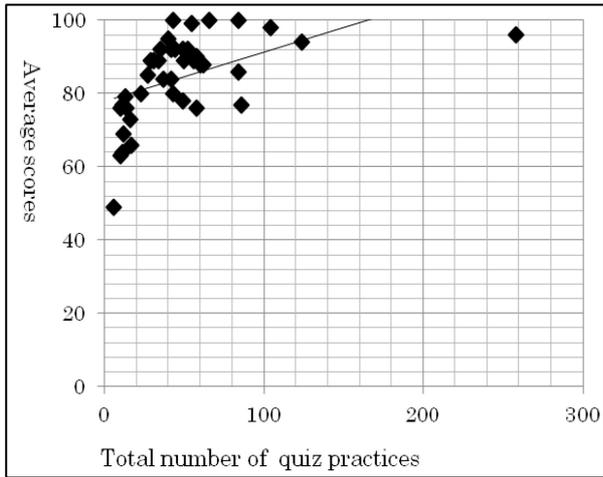


Figure 1. Correlation between total number of quiz practices and scores

G. Situation about the number of quiz practices

With relation to the number of quiz practices, the scores made by “Usage instructed G” and “Non-instructed G” in the former five sessions and the latter five sessions were investigated. Table VIII. shows the results thereof.

As seen from Table VIII., the “Usage instructed G” made the average score of 81.2 points in the former five sessions and the average of 87.6 in the latter five, exhibiting +6.4 increase. By contrast, the “Non-instructed G” made the average score of 77.6 points in the former five and the average of 69.3 in the latter five, exhibiting 8.3 decrease. That is, the trends of the average scores of “Usage instructed G” and “Non-instructed G” from the former five to the latter five were found to be inverse. The “Non-instructed G” made the better score in the former but underwent the decrease in the score in the latter. Inversely, the “Usage instructed G” made the better score in the latter five than in the former five. This suggests that the instructed group continued independent learning after the repeated instructions for the practice of the quizzes during the former four sessions were ceased.

Subsequently, the change in the number of quiz practices in each class session was investigated for only the “Usage instructed G.” Table IX shows the results.

As seen from Table IX, the average number of practices in the former five sessions was 4.35, while that in the latter five was 5.38. The number was larger in the latter half without the instructions for the usage, and it is thought this was because the students independently continued learning.

At an interview with a student in “Usage instructed G,” the student said, “I gave myself over to the practices until achieving the score of 100 points.”

The results in Tables 8. and 9. suggest that “Usage instructed G” came to learn continuously and independently so as to keep trying until achieving satisfactory scores while they practiced the quizzes.

Fig. 2. shows the relation between the score in the first practice and the number of the practices for the quiz given in each class session.

There were many students who stopped the practice after achieving high scores in the first practice, whereas there were many students who practiced the quiz until they were satisfied.

TABLE VIII. COMPARISON OF AVERAGE SCORES IN QUIZZES IN FORMER AND LATTER FIVE SESSIONS

	Former 5 session	Latter 5 session	Difference t Test Judgment
Usage instructed G n=40	81.2 (29.6)	87.6 (24.9)	+6.4 t=2.22 p<0.05
Non-instructed G n=37	77.6 (31.4)	69.3 (42.7)	-8.3 t=-2.15 p<0.05

Numerical values designate average scores (standard deviation)

TABLE IX. COMPARISON OF QUIZ PRACTICE BY USAGE INSTRUCTED G IN FORMER AND LATTER HALVES [SUBJECT: USAGE INSTRUCTED G]

Class session	first	second	3rd	fourth	fifth	average
Number of practices	1.75	7.58	3.53	3.63	5.25	4.35 (6.97)
Class session	sixth	seventh	eighth	ninth	tenth	average
Number of practices	3.68	4.80	7.75	4.08	6.60	5.38 (7.30)
					difference	1.03 t=1.45 p<0.10

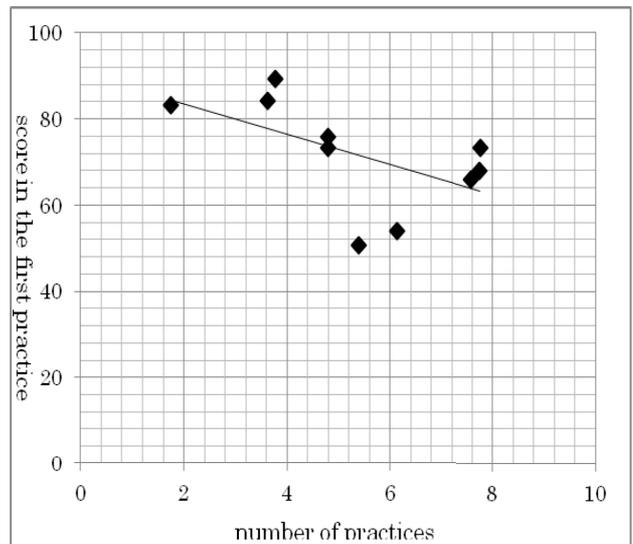


Figure 2. Relation between score in first quiz practice and number of practices in each class session

The number of the practices is the average value in each session that covers the number for the students satisfied with only the first practice and the number of the practices continued until the achievement of satisfaction. In total, the average score was 71.8 points and the standard deviation was 12.6. As seen from Fig. 2., the number of quiz practices tended to be small when a high score was made in the first practice and tended to be large when a low score was made in the first practice.

This indicates that the quiz having difficult questions led to low scores in the first practice and thus resulted in the increase in the number of quiz practices.

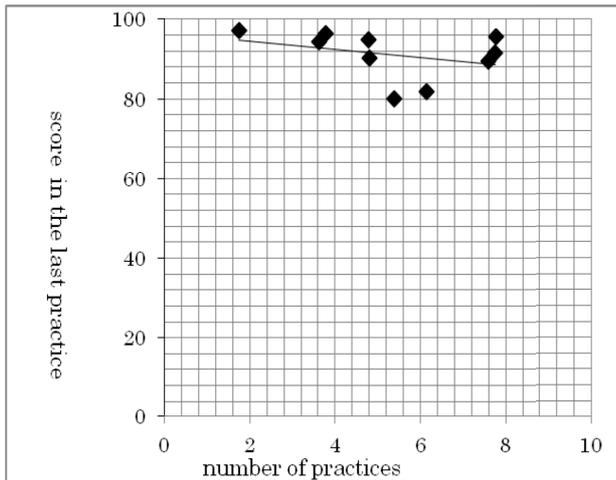


Figure 3. Magnetization as a function of applied field. Note how the caption is centered in the column

Fig. 3. shows the relation between the score in the last quiz practice and the number of practices in each class session.

As seen from Fig. 3., the scores in the last practice were not lower than 80 points in general irrespective of the numbers of practices. It is conceived that the students repeatedly practiced the quiz until they achieved 80 points or higher scores, regardless of the difficulty of the questions. A large number of students worked hard until achieving 100 points, and it is conceived that the above-mentioned student said, "I gave myself over to the practices," based on true feelings. This is backed up with the fact that the average number of times of achieving the perfect score of 100 points was as high as 5.5 per person. In total, the average score was 91.2 points and the standard deviation was 6.03.

The average number of practices by all the students in each class session was 4.86, and the highest number was 50. In an anonymous questionnaire issued to the students after the completion of all the sessions, a question therein reading "What degree has time you get for preparation and review changed in comparison with that for course including only 'face-to-face' class?" caused 61% of the respondents to answer "Greatly increased" or "Slightly increased." Thirty-nine percent answered, "Not changed," and answers "Slightly decreased" and "Greatly decreased" comprised 0% [10].

IV. CONCLUSION

Regarding Blended-Learning, the authors made the division into the usage instructed group and the non-instructed group and the comparison of learning results etc. between the two groups. Consequently, the increases in the percentage of participation in the quizzes, the percentage of paper submission, etc. were greater in the usage instructed group than in the non-instructed group. In the comparison of the results of the term examination, the former group exhibited the increase not only in the understanding and knowledge in the threefold choice, the fill-in-the-blanks questions, etc. but also in the scores concerning the multifaceted statement with self-assertiveness and the clear pictorial explanation in the short essays.

The comparison of the learning results according to the questions between the lower-ranking group and the high-

er-ranking group in the quizzes made clear that the higher group had the significantly higher scores in each question than the lower group did. That is, it was observed that the higher group in the quizzes achieved the higher scores in each question of the written term examination.

The comparison of the scores between the higher- and the lower-ranking groups in the number of quiz practices in the usage instructed group revealed that the higher group in the number achieved higher points in the scores than the lower group did. Concerning the practice of the quizzes, it was suggested that the repeated instructions for the usage in the former half brought about the independent practice.

The scores in the first practice of the quizzes were found to vary considerably according to the difficulty of the questions. The scores in the final practice revealed that most of the students made the practices as many times as they like with the aim of achieving scores not lower than 80 points.

In the Blended-Learning of this type, it is difficult to improve learning results only by creating an environment allowing practice of quizzes with provision of contents for preparation and review. It is thought to be important to take measures to encourage students to continue learning, e.g., to motivate them by the addition of results of quizzes in middle of the term to the record for accreditation and to repeat instructions for the practice at first.

One of our objectives in the future is required to determine whether or not students developed independent learning

REFERENCES

- [1] I. MIYAJI, H. YAO, and K. YOSHIDA, "The Practice and Learning Effects of Education by Blending of Lecture and e-Learning," *Transactions of Japanese Society for Information and Systems in Education*, 22(4): pp. 254-263, 2005
- [2] T. SAITO, and S. KIM, "A Meta-Analysis On e-Learning Effectiveness in Higher Education," *Japan Journal of Educational Technology*, 32(4): pp. 339-450, 2009
- [3] K. ADACHI, "Analysis of the Classification of the Learners' Activities in Blended Learning," *Japan Journal of Educational Technology*, 31(1): pp. 29-40, 2007
- [4] I. MIYAJI, "Editing and Writing," *Toward Blended Learning from e-Learning*. Kyoritsu Shuppan Co., Ltd. (JAPAN), 2009
- [5] T. KATASE, H. YAMAMOTO, and K. MUTSUURA, "The Motivation Effects of the Blended Learning Environments in Promoting Self-Regulated Learning," *Japan Journal of Educational Technology*, 34(Supple): pp. 17-20, 2010
- [6] UNIVERSITY COUNCIL REPORT, "About a University Image of the 21st Century and a Future Reform Policy," (Reference day: July 17, 2012), 1998 http://www.mext.go.jp/b_menu/shingi/old_chukyo/old_daigaku_index/toushin/1315917.htm
- [7] REPORT of CENTRAL COUNCIL for EDUCATION, "For Construction of the First Degree Education," (Reference day: July 17, 2012), 2008 http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/1217067.htm
- [8] M. KANEKO, "The University which Lets You Learn It," *IDE Modern Higher Education*, No.515, pp. 4-11, 2009
- [9] J. NISHIGAKI, "Difference of the Recognition about the Scholastic Evaluation of the Teacher and the Student," *Journal of Educational Research Shinshu University*, 10: pp. 13-23, 2004
- [10] H. YAMAMOTO, M. NAKAYAMA, and Y. SHIMIZU, "Association Between Quiz and Term-end Examination with the Learning Environment by Using ICT," *Research Report of JSET Conferences (JSET 10-4)* pp.1-8, 2010
- [11] H. YAMAMOTO, "Educational Promotion Policy of ICT Utilization at Whole Shinshu University," *Research Report of National Institute of Multimedia Education (39)*: pp38-53, 2008

PAPER
MEASURES TO PROMOTE PRACTICE OF QUIZ AND EVALUATION THEREOF IN BLENDED-LEARNING

AUTHORS

H. Yamamoto is with Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Tokyo, 152-8552 Japan (yama77722jp@ybb.ne.jp). Correspondent author.

M. Nakayama, is with Graduate School of Decision Science and Technology, Tokyo Institute of Technology, Tokyo, 152-8552 Japan

Y. Shimizu is with Auditor, Tokyo Institute of Technology, Tokyo, 152-8550 Japan

This work was supported by JSPS KAKENHI Grant-in-Aid for Challenging Exploratory Research Number 23650529 Submitted, 24 January 2014. Published as resubmitted by the authors 14 September 2014.