An Application of DBR to the Development of a Framework for English Language Education

Sei Sumi and Thomas Schalow
University of Marketing and Distribution Sciences
Gakuen Nishimachi 3-1, Nishi-ku, Kobe-city, Hyogo 651-2188 Japan
seisumi@lang-tech.net   ijinkan@mac.com

ABSTRACT

This paper explains how to develop an integrated framework for English language education based on the four phases of design-based research (DBR). In the first phase of the DBR we conducted a qualitative investigation to identify relevant problems at the Japanese university where this research was undertaken. To ameliorate these problems in the second phase, we (a) developed an original placement test, (b) created a classification technique by employing a Bayesian algorithm, and (c) established an e-learning system and comprehensive educational policy. In the third phase the developed placement test was tested, and all the test items were analyzed using a Rash model. In the last phase our framework for English language education was put into practice. Our findings confirmed that the new framework not only resulted in a substantial reduction in budgetary and manpower costs, but also successfully contributed to the improvement of the students' learning outcomes and integration of the entire process of English language education.

KEYWORDS

English language education, design-based research, test development, e-learning, education in Japanese

1 INTRODUCTION

English language education in Japanese universities has been changing rapidly in recent years as a result of a number of government initiatives and directives. The Ministry of Education, Culture, Sports, Science and Technology recently released a report on the globalization of Japanese universities.[1] In the report reforms in English language education were particularly emphasized. These reforms were in line with Japanese government plans to employ standardized tests such as TOEFL to set minimum English requirements for new candidates for positions as government bureaucrats.[2] These changes will have a tremendous impact on the shape of English language education at Japanese universities.

In addition to changes resulting from these government initiatives, change is being driven by the sharp decline in the student population. Japanese universities have started lowering their standards of admission and the level of exams to enroll as many students as possible. There are even many students who never take any kind of test to get into universities, if they possess a talent for sports or music. Consequently, it has become difficult to require students to have a basic command of English. Some students have not attained even a high school level of English ability. Many universities are addressing this challenge by offering small-group lessons and remedial education classes based on students' English ability. A key common concern for these universities is how to raise fundamental English ability in areas such as vocabulary and grammar.

To respond to these changes in the direction of English language education our university, the University of Marketing and Distribution Sciences, has initiated several projects.[3] In 2013 we developed and deployed a novel framework for English language education based on a comprehensive assessment of English language education, from the placement test for freshmen to the final achievement test for sophomores. This paper reports on our project.

2 DESIGN-BASED RESEARCH

The framework for our English language program was developed using design-based research (DBR). The origin of DBR is most closely associated with Brown (1992) and Collins (1992).[4] [5] In the early stages of development DBR was labeled a design experiment (DE) in consideration of the research focus, practice, and underlying epistemology.[6] DE focuses on the understanding of a "learning ecology", which consists of complex elements of an interactive system in teaching and learning, with different types and levels.[7] In contrast to other research methodologies that reduce complex teaching and learning practices to a simple cause-and-effect models with a list of separating factors, DE aims to offer a holistic explanation for design-based practice, and has greatly influenced educational studies.

DE has grown in diversity since it first appeared in the field of educational studies. Consequently, it has become difficult to find theoretical or methodological coherence among efforts purporting to use DE. In recognition of this diverse nature of DE, the Design-Based Research Collective (2003) was created to provide new theoretical and methodological advances.[8] Their work clearly defined the character of DBR, and showed the direction for ensuing studies based on DBR.

DBR, with its flexible and practical characteristics (Cobb, Confrey, diSessa, Lehrer, & Schauble, 2003), has proven especially useful in educational and other settings that face new challenges brought on by technological innovation.[9] DBR is therefore now practically applied to teaching practices, and though the number of such teaching practices is quite limited (Hung, 2011), their variety is growing.[10] For example, in the field of mathematics education Gravermeijek and Cobb (2006) applied a DBR framework to teaching practice and contributed to identifying a local instruction theory.[11] Yutdhana (2005, p. 170) introduced DBR as an effective alternative to Computer-assisted Language Learning research and described its two key features as follows[12];

1. Designing learning environments and
2. Developing theories of learning.

According to Amiel and Reeves (2008), DBR comprises four phases of study (Figure 1).[13] Following these four phases, we conducted a qualitative research study, designed a solution, tested it, documented the results, and provided our assessment of their theoretical and practical implications.

2.1 Phase 1: Analysis of practical problems

In Phase 1 we conducted interviews to identify

Design-based research

| Analysis of practical problems by researchers and practitioners in collaboration | Development of solutions informed by existing design principles and technological innovations | Iterative cycles of testing and refinement of solutions in practice | Reflection to produce "design principles" and enhance solution implementation |

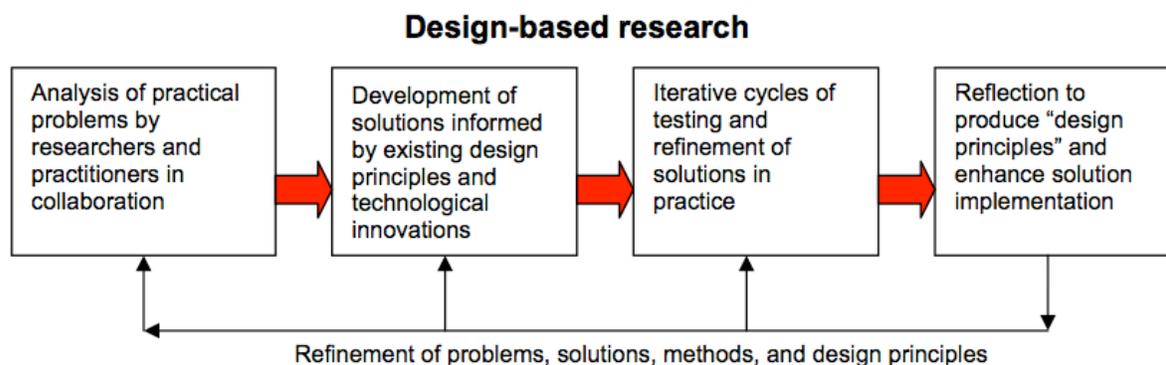Refinement of problems, solutions, methods, and design principles

Figure 1. The four phases of DBR (adopted from Amiel & Reeves, 2008, p. 34)

problems in the placement test the university had used during the 2012 academic year. The university had used a private, standardized test every year since 2008 to divide freshmen into three levels based on their scores; advanced classes (20%), intermediate classes (60%), and basic classes (20%). After the placement test students take compulsory English classes at this level for two years.

Many Japanese universities presently use private, standardized tests such as TOEIC or TOEFL as the basis for their placement tests.[14] These tests are well known among Japanese students and private companies in Japan, and have many advantages. Perhaps the main advantage for universities is that they offer a standardized, if not always accurate, means to assess the language abilities of students. Students were thought to benefit by being placed in language classes they seem to be best suited for. For these reasons, we used a standardized test for the placement test, as well as for the achievement test. The achievement test is conducted two years after the placement test, after students finish their compulsory English lessons.

However, student interviews revealed a number of problems with the placement test. Since the English language abilities of students vary widely, many students felt burdened with the length of the exam, the unfamiliar testing style, and the high-level test items. The university also felt encumbered by the need to use a standardized placement test because the month during which the placement test is conducted coincides with the busiest time for enrolling new students. The university expends considerable manpower and a substantial amount of financial resource to conduct the placement test. The placement examination results are returned a couple of days after the test is conducted, and manpower resource must once again be expended in checking results and dividing all the students into the three levels. Moreover, in addition to the fact the placement test imposed a burden on the students and administrative staff, all test items remain undisclosed to the public, preventing the university from identifying students' weaknesses and using the data for educational purposes after the exam.

The original purpose of the placement test was simply to divide students into three levels based on their scores. To achieve this purpose the students and administrative staff have thus far been willing to accept the burden of the examination. Through the interview process, however, we found that students are questioning the validity of using a standardized test for the placement examination, while the university is questioning the benefits of dividing the students into skill levels that may not reflect working knowledge of the language. Taking all these problems into consideration, we decided to develop an original placement test and devise a framework for English language education that is compatible with the diverse English language abilities of students, thereby contributing to an improvement in the students' learning outcomes.

2.2 Phase 2: Development of solutions

In Phase 2, in an effort to ameliorate the problems identified in Phase 1, we (a) developed an original placement test, (b) devised a classification technique employing a Bayesian algorithm, and (c) established an e-learning system and language education policy to take student needs and abilities into consideration.

a. Development of the placement test

In order to develop an original placement test, we decided to put the entrance examinations used by our university from 2004 to 2012 into digital format, creating a databank of approximately 1,000 test items. All test items were arranged according to the correct answer rate, and 10% of the total items were randomly selected. We called this test the β test.

We recognized numerous advantages of recycling past entrance examination test items for the placement test. All the test items for the entrance examination represent the university's minimum standard of English. The vocabulary and grammar employed in each test item also adhere to a

minimum standard. By using similar entrance examinations every year, we are able to measure students' English ability over time. Moreover, because all entrance examinations are published and available to the public in the following year, we are able to use the results to better understand the students' weaknesses at both the individual and class levels.

b. Development of the classification technique

Several methods exist for classifying students into different groups based on test scores. The simplest approach is to use the raw score. However, using this method it is impossible to discern a meaningful difference between identical scores. If, for example, students A and B each completed 50 of 100 questions on a test, but student A chose to answer more difficult test items, while student B was only able to answer much easier test items, they might both achieve similar scores when the examinations were graded. It would probably be wrong to conclude, as raw test scores would, that the English ability of student A and B is the same.

To avoid this problem, we applied a naive Bayes classifier to the categorization of the students' test scores. A naive Bayes classifier allows users to classify data based on a probability model that has been widely used in solving problems such as the classification of spam email. With this algorithm we can classify students based on their test answer patterns.

A classification technique that enables us to categorize students into three levels can be achieved by a developing Bayes' theorem (1) for an advanced class (2), an intermediate class (3), and a basic class (4).

$$P(H_i|D_j) = P(D_j|H_i)P(H_i) / P(D_j) \tag{1}$$

$$P(D_j|H_1)P(H_1) > P(D_j|H_2)P(H_2) > P(D_j|H_3)P(H_3) \tag{2}$$

$$P(D_j|H_1)P(H_1) < P(D_j|H_2)P(H_2) > P(D_j|H_3)P(H_3) \tag{3}$$

$$P(D_j|H_1)P(H_1) < P(D_j|H_2)P(H_2) < P(D_j|H_3)P(H_3) \tag{4}$$

In the equations, $P(H_1)$, $P(H_2)$, and $P(H_3)$ are prior probabilities. $P(H_1)$ is for advanced classes, $P(H_2)$ is for intermediate classes, and $P(H_3)$ is for basic classes. A prior probability indicates a correct answer rate at each level on the placement test. A prior probability can be subjectively set after consideration of factors such as the students' scores at each level, instructors' expectations, and test difficulty. An advantage of Bayes' theorem is that it enables users to employ subjective factors in order to build a probability model. The $P(H_i|D_j)$ section in the equations indicates the likelihood of each test item, which informs us of level i, indicating the student's correct answer rate on a test item j.

A probability model comprising prior probability and likelihood can be trained through practical application using Bayesian updating. The trained probability model is applicable to a number of fields, such as the development of a curriculum and materials. The theoretical background of Bayes' theorem, allowing us to employ users' expectations and subjectivity for updating the model, has many similarities with the DBR framework, which integrates research and practice to draw a better solution by following four phases.

c. Developing an e-learning system and instructional design

The e-learning system we developed also used the databank of all past entrance examination test items, producing enough lessons for two years of compulsory study. To achieve this, we rearranged all test items by the correct answer rate, divided them by 20 test items, and successively created 48 sets of test items. In addition, we divided the 48 sets of test items into four categories, and named each category Basic, Standard, Intermediate, and Advanced. The following list indicates the four categories and sets of test items included. The percentage in the brackets shows the range of the correct answer rate.

Basic: SET001 to SET012 (99.1%–66.4%)
Standard: SET013 to SET024 (66.3%–52.7%)
Intermediate: SET025 to SET036 (52.6%–39.5%)
Advanced: SET037 to SET048 (39.4%–3.2%)

During their two years of compulsory English lessons students study English I in the first semester of the first year, English II in the second semester of the first year, English III in the first semester of the second year, and English IV in the second semester of the second year. We allocated the test items of Basic to English I, Standard to English II, Intermediate to English III, and Advanced to English IV. Students spent approximately 20 to 30 minutes studying each set of test items every week. The test items studied in the prior week were tested the following week. The instruction design was designed based on the "Cyclic Model of Learning."[15]

The cyclic model of learning (CML, Takeuchi, 2007)[16] was designed with a special focus on the local context, namely, Japanese EFL classrooms. The most distinctive feature of the CML is that it integrates in-class teaching practices with students' out-of-class self-learning, with the aid of technology. In addition, the CML attempts to elicit student participation in lessons and to promote self-learning outside the classroom, both of which are considered to be indispensable to successful EFL learning (Takeuchi, 2002).[17]

All test items were uploaded onto a Moodle web system.[18] Figure 2 shows an example of the Basic test items. In addition to the test items, videos of study-abroad programs, and learning materials for basic grammar, vocabulary and extensive reading were uploaded onto the Moodle website. Any computer or smart phone with online capability was able to access the study and test contents.

A final examination was conducted at the end of each semester. Students thus take achievement tests five times, including the placement test, during their two years of compulsory English lessons. All test scores are analyzed according to levels, classes, faculties, and tutors every time the final examination is conducted. The test results are then applied to the improvement of test items, materials, and lessons.
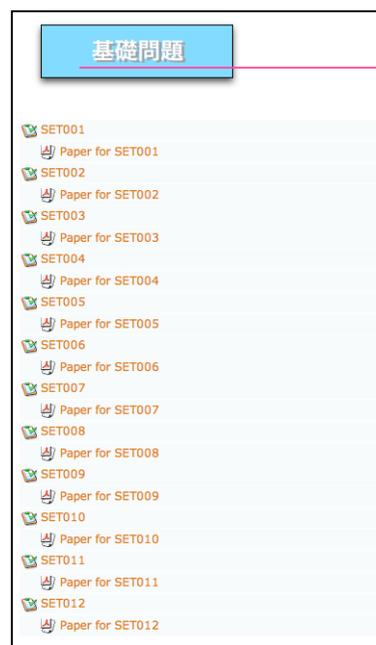

Figure 2. Example of Basic test items

2.3 Phase 3: Testing and refinement of solutions

In Phase 3, the β test was tested, and each test item was analyzed (N = 945, M = 41.19, SD = 9.31). Figure 3 shows the overall picture of the result. All test items were analyzed using a Rash model. The findings showed the average of the examinees' ability was -0.47, the standard deviation was 0.54, and the reliability of the examinees' ability was .78. The average item difficulty level was 0, the standard deviation was 1.32, and the reliability of the item difficulty level was .97.

Figure 4 displays a persons and item map, which allows us to understand the result of the test on a single scale. As shown in the figure, misfit items can be identified both in the area of the high- and low-difficulty levels. While misfit items found at the low-difficulty level had to be refined, misfit items found at the high-difficulty level were items that all students need to study throughout the two-year compulsory English lessons. Therefore, we retained those items without modification. No misfit item (*infit MSNQ* ±0.75) was found. Based on the results, we decided to employ 100 modified test items, and developed the placement test (ver. 1.0).

Figure 3. Result of the β test.

3 PRACTICE

In Phase 4, the placement test (ver. 1.0; $N = 653$, $M = 37.31$, $SD = 11.45$) was conducted on freshmen entering in the academic year that began in 2013 (March, 2013). According to an analysis of a Rash model, the average of the examinees' ability was -0.62, the standard deviation was 0.60, and the reliability of the examinees' ability was .84. In addition, the average item difficulty level was 0, the standard deviation was 0.90, and the reliability of the item difficulty level was .99. Figure 5 shows that the misfit items in the area of the low-difficulty level had improved. The overall result shows that there was no misfit item that was typically over or under (*infit MSNQ* ±0.75). We therefore divided students into three class levels (i.e., Advanced, Intermediate, and Basic) by employing a probability model developed using a naive Bayes classifier.
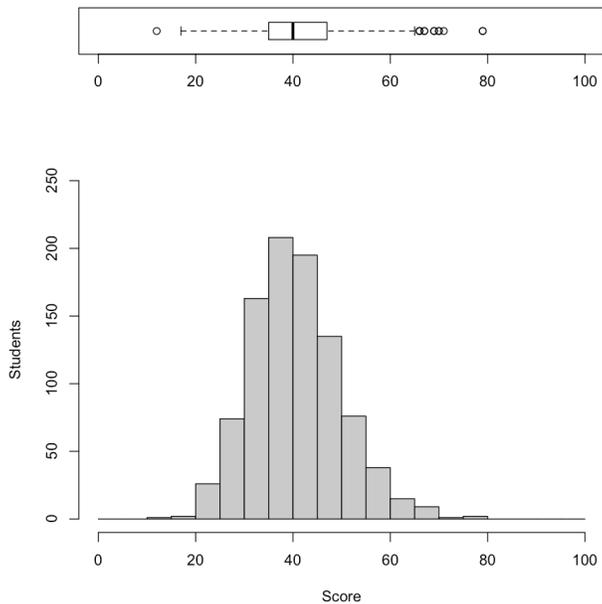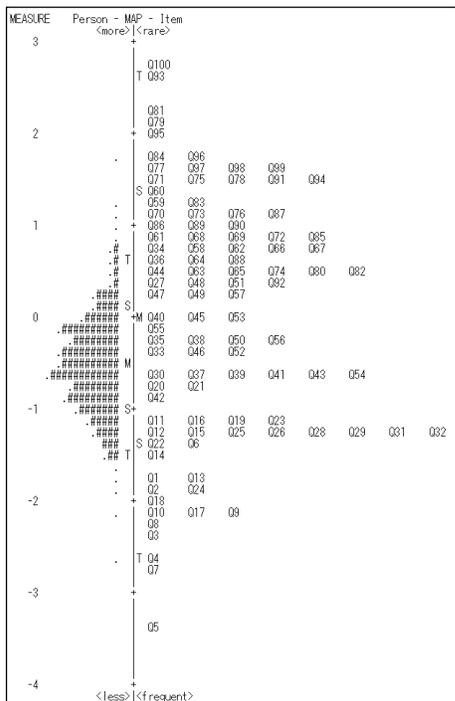
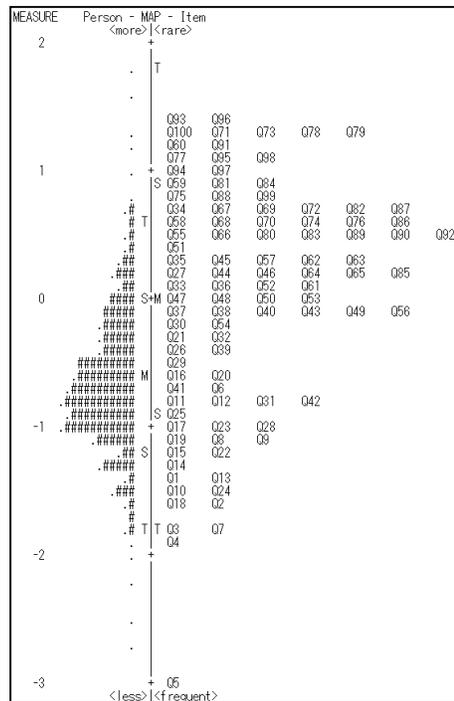

Figure 4. Person and item map
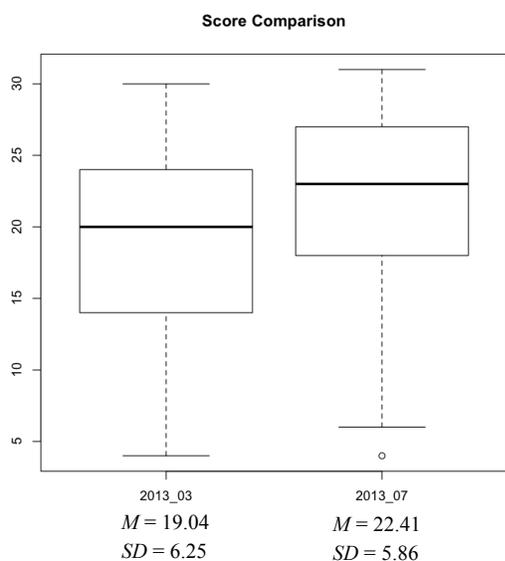

Figure 5. Result of the placement test ver. 1.0

Figure 6. Result of the placement test and the achievement test

Figure 6 shows the result of 31 shared test items in the placement test in March, 2013 and the achievement test in July, 2013. 451 freshmen took both tests, and we compared the scores of the shared items. As Figure 6 shows, a tendency toward improvement of the students' vocabulary and grammar results can be recognized.

4 SUMMARY

As a result of our efforts, we were able to achieve three project outcomes:

1. Development of an original placement examination (ver. 1.0)
2. Development of a probability model based on a naive Bayes classifier
3. Development of an integrated framework for two-year compulsory English lessons, with applications to actual lessons

In regard to the first point, it is worth noting we managed to drastically reduce budgetary and manpower costs. Beginning in the 2014 academic year we will conduct a computer-based placement test using Moodle software, in place of the previous private, standardized test. Since we have developed a new Moodle module for a naive Bayes classifier, scores of any test conducted on Moodle can be classified automatically. Further

budgetary and manpower savings will thus be possible.

Regarding the second point, we classified students into appropriate class levels. The probability model proposed by our project can be trained and refined through successive application. With a more sophisticated probability model, we are currently working on the development of a naive Bayes-based computer-adaptive teaching system. More integrated and sophisticated teaching support will be offered once the new system becomes available.

In regard to the third point, we managed to develop an integrated framework for English language lessons. Using the framework we are able to identify students' English language abilities and deficits regularly and successively over time. Data taken from the test are useful not only for students, but also for tutors and the university in attempts to improve lessons and the syllabus. Also, the framework underpins the minimum standard of students' English ability both in vocabulary and grammar, and can contribute to the improvement of their learning outcomes.

Certain limitations to our efforts did, however, become clear as we continued our research. The framework created through the project mainly focuses on developing and monitoring students' basic English language abilities in areas such as vocabulary and grammar. Additional efforts are required to add communicative or practical support and materials to the framework.

Although limitations were found, we managed to develop an integrated framework for English language lessons. The framework endorses the two-year compulsory English course of study, from the placement test prior to entrance to the achievement test in the second year. The system framework, which enables us to regularly monitor students' English language abilities, provides us with important data to direct our English education in an effective manner. With the globalization of education and the need to address the decline in students' basic English abilities, we

are confident the integrated framework we developed through the project will aid Japanese universities with similar budgetary, manpower, and curriculum concerns.

5 REFERENCES

1. Japanese Ministry of Education, Culture, Sports, Science and Technology.: Reformation of Educational Plan for Development of Global Human Resources (2013). http://www.kantei.go.jp/jp/singi/keizaisaisei/skkkaigi/dai7/siryou07.pdf

2. Task Force for Developing Global Human Resources.: Strategy for Developing Global Human Resources (2012). http://www.kantei.go.jp/jp/singi/global/suisin/dai2/siryou01_1.pdf

3. Yamamoto, K., Azuma, J., Sumi, S,.: Development and Application of a Blended Learning System for English Learners. Journal of the University of Marketing and Distribution Sciences (Humanities, Arts & Sciences) 24 (2), 33–37 (2012).

4. Brown, A. L.: Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. The Journal of the Learning Sciences 2 (2), 141–178 (1992).

5. Collins, A.: Toward a Design Science of Education. In: Scanlon, E., O'Shea, T. (eds.) New Directions in Educational Technology, pp. 15-22. Springer, Berlin (1992).

6. Bell, P.: On the Theoretical Breadth of Design-based Research in Education. Educational Psychologist 39 (4), 243–253 (2004).

7. Cobb, P., Confrey, J., diSessa, A., Lehrer, R., Schauble, L.: Design Experiments in Educational Research. Educational Researcher 32 (1), 9–13 (2003).

8. The Design-Based Research Collective.: Design-based Research: an Emerging Paradigm for Educational Inquiry. Educational Researcher 32 (1), 5–8 (2003).

9. Cobb, P., et. al.. op. cit. (2003)

10. Hung, H. T.: Design-based Research: Designing a Multimedia Environment to Support Language Learning. Innovations in Education and Teaching International 48 (2), 159‑169 (2011).

11. Gravemeijer, K., Cobb, P.: Design Research from the Learning Design Perspective. In: van den Akker, J., Gravemeijer, K., McKenney, S., Nieveen, N. (eds.) Educational Design Research: The Design, Development and Evaluation of Programs, Processes and Products, pp. 17-51. Routledge, London (2006).

12. Yutdhana, S.: Design-based Research in CALL. In: Egbert, J.L., Petrie, G.M. (eds.) CALL Research Perspectives, pp. 169–178. Lawrence Erlbaum Associates, Mahwah, NJ (2005).

13. Amiel, T., Reeves, T.C.: Design-based Research and Educational Technology: Rethinking Technology and the Research Agenda. Educational Technology & Society 11 (4), 29–40 (2008).

14. Regional Science and Center for Higher Education.: (2013) Application of TOEFL® Test for Development of Global Human Resources. 3-01–3-09.

15. Sumi, S., Takeuchi, O.: Using an LMS for foreign teaching/learning: An attempt based on the Cyclic Model of Learning. The Journal of Information and System in Education 7, 59–66 (2008).

16. Takeuchi, O.: Mizukara manabu shisei wo minitukerutameniha. [Autonomy and self-learning: Necessities and methods] Teaching English Now 8, 2–5 (2007).

17. Takeuchi, O.: Gaikokugo shutoku no seihi: Kyoshitsugai no doryoku ga kagi. [Success or failure in foreign language learning: Some keys for success] Mainichi Shinbun (2002).

18. http://moodle.umds.ac.jp

ACKNOWLEDGMENT