Analysis of Game Development Activity using Team-Based Learning

Akiko Teranishi*, Minoru Nakayama†, Theodor Wyeld‡ and Eid A. Mohamad*
*Electrical and Computer Engineering
New York University Abu Dhabi, PO BOX 129188, Abu Dhabi, United Arab Emirates
Email: at132@nyu.edu
†Information and Communications Engineering
Tokyo Institute of Technology, Meguro, Tokyo 152-8552, Japan
Email: nakayama@ict.e.titech.ac.jp
‡Department of Screen&Media
Flinders University, GPO Box 2100, Adelaide 5001, South Australia, Australia

***Abstract*—Game development activity using Team-based Learning (TBL) was investigated in order to identify factors contributing to the usability of the product. In this study, three teams from two different countries are compared. As the related factors, the followings were examined to analyze the relationships with usability scores: (1) learning reflection, (2) social media communications within teams, and (3) participants' characteristics and information literacy. Usability scores were conveyed using a System Usability Scale (SUS) by evaluations from the other teams. The participants' characteristics and information literacy were measured before starting the project as a pre-test. The discussions and communications via social media of each group were categorized as: Proposal, Permission, Encouragement, and Acknowledgment, using protocol analysis to examine their contributions towards the usability scores. After completing the study project, a learning reflection questionnaire was completed by all participants to evaluate efficacy, satisfaction and achievement of learning, and difficulties.**

1. Introduction and Related Work

Human Computer Interaction (HCI) environments in educational contexts often engage team-based learning (TBL). A TBL environment positively influences students by making them interact with each other and share their knowledge and skills. It is a well-known strategy for enhancing the quality of student learning [1]. Gomez et al [2] introduced computer-supported team-based learning (CS-TBL) to extend learning in the hybrid classroom where students meet both face-to-face and online. It has been shown to influence students' motivation, enjoyment, and team contribution to learning. HCI learning environments are more interactive using such CS-TBL strategies when compared to conventional course studies when learning about usability concepts [3].

Cross-cultural TBL was used to teach HCI using a 3-D Collaborative Virtual Environment [4]. Multicultural teamwork was also examined between Australians, Norwegians, and Taiwanese members from previous work [5]. Cross-cultural learning environments were shown to enhance students' general knowledge and enrich their learning experiences.

Participants' reflections, satisfaction and achievement in learning are considered impactful factors for educational effectiveness in both distance learning and web-based learning environments [6] [7] [8]. Online communications [9] [10] and personalities [11] [12] [13] have been analyzed as psychologically influential factors in collaborative learning. The instruction and evaluation methods for the TBL in the current study has been captured by visualizing and modeling these measuring factors.

In the current study, the following factors are investigated to clarify the instruction strategy and support of students' learning of HCI in an environment using cross-cultural TBL:

1) The relationships between the SUS scores and the participants' reflection were measured to analyze summative evaluations.
2) The social media communications among the teams were analyzed to investigate the relationships between the SUS scores and the categorized communications as formative evaluations.
3) The participants' attributes, such as characteristics and information literacy, were evaluated in order to better understand usability as diagnostic evaluation.

2. Study procedure

2.1. Implementation cycles

The study was conducted using an online 2D mobile phone game development task, where participants were asked to develop a game as a collaborative exercise with their teammates. Each team started with the same level of a basic html and javascript browser game, provided by the experimenters. The game was developed using an online

2375-0138/17 $31.00 © 2017 IEEE
DOI 10.1109/iV.2017.73

Figure 1. Basic html and javascript browser games

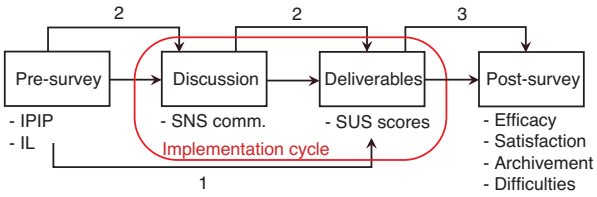

Figure 2. Conceptual diagram

game editor provided by the experimenters and/or by changing the code directly. Teams were required to implement the game with usability testing between versions. Code changes and updates were based on usability testing by the other teams. Participants included 12 students from United Arab Emirates (UAE) and Japan. Participants were divided into 3 teams depending on their nationalities: 1. Arabic team (consisted of 3 Arab students); 2. Japanese team (consisted of 4 Japanese students); and, 3. International team (consisted of 3 Japanese students and 2 Arab students). In this study, we refer to the Arabic team as "domestic team 1" and the Japanese team as "domestic team 2".

The duration of the project was 3 weeks with 3 development cycles. Each cycle lasted 7 days: 6 days for development and 1 day for usability testing. In the first cycle, all teams started with the basic game and developed it further based on the first usability testing result, which was evaluated on the first day. At the end of each cycle, all participants evaluated another team's game. The evaluation results were collected and sent to each team for the next cycle. The experimenter explained the overall project at the beginning of the study in both UAE and Japan. The html and javascript browser games consisted of three basic and conceptually similar games (see Figure 1):

- Globe Game (Domestic Team 1): The main character (green circle) scores points by catching its friends (yellow circles) and at the same time tries to avoid the enemies (orange squares) which randomly attack. Once the main character is hit by the enemies, the game is over.
- Maze Game (Domestic Team 2): The main character (green square) scores points by catching its friends (blue squares) and at the same time tries to avoid the enemies (orange squares) which randomly attack. Once the main character is hit by the enemies, the game is over. The movement is restricted by the maze shape.
- Catcher Game (International Team): The main character (green square) scores points by catching its friends (orange squares) which randomly fall down from the top of the screen by enemies (orange squares). Once the main character fails three times (lives shown as red squares on the top right) to catch the friends, the game is over.

There were 12 participants in this study. They all signed the Informed Consent Form, authorizing their social media communication for data analysis. Participants who completed the team-based work received a compensation of $100. Participants were between 18 and 24 years old.

2.2. Methodology

The study flow is illustrated in Figure 2. In this study, formative evaluations using team communications and summative evaluations using learning reflections were measured to analyse the interrelationship between factors. The participants' attributes, such as characteristics and information literacy, were evaluated as diagnostic evaluation.

The participants were instructed to complete the International Personality Item Pool (IPIP) test to measure their personalities [14] and an Informational Literacy (IL) test to measure their skills and knowledge [15]. During the game implementation cycles, each participant interacted with his/her teammates to improve the usability of the games. In this study, the SNS communications between teammates was provided for analysis. Before starting the project, and after each development cycle, the participants individually evaluated the usability of the other teams' games. There were 3 development cycles in total. Four SUS scores were provided in total. At the end of the study, each participant was asked to answer a questionnaire about their learning experiences; we call this their learning reflection. Evaluation and impact of each factor will be discussed.

2.2.1. Measuring the participations' attribute. Measuring participants' attributes, such as personality, were examined its influence on performance based on previous work [11] [12] [13]. In this study, the participants' characteristics and information literacy were measured to analyze the relationships between usability scores and these attributes. Before starting the study, the participants were asked to complete 2 tests: the IPIP and IL. These tests were conducted to examine any differences between teams regarding characteristics and information literacy. IPIP inventory is used to evaluate the participants' characteristics in terms of "Extraversion", "Agreeableness", "Conscientiousness", "Emotional stability", and "Intellect" [14]. The IL is performed to assess a participants' ability for computer skills and general knowledge of information technology [15]. The IL test consists of 32 questions. Eight defining factors can be elicited from the IL test: "interest and motivation," "fundamental operation ability," "information collecting ability," "mathematical thinking ability," "information control ability," "applied operation ability," "attitude,"

and “knowledge and understanding.” In addition, based on the questions in the IL test, “operational confidence and knowledge understanding”, as IL-SF1 and, “attitude issues”, as IL-SF2 were extracted as secondary factors, as shown in a previous study [13].

2.2.2. SNS communications. The discussions and communications by the teams in this TBL exercise provided information about the processes and strategies of activities among teams [9] [10]. In this study, communications via social networking services (SNS) between teammates were provided to the experimenters. Based on these communications, the relationships between the usability scores and communications were analyzed to understand the effectiveness of each of the factors.

All groups decided to use social media applications to interact with teammates, whereas the experimenter did not specifically define the types of communication they should use. Most of the social media applications had similar functions and features. All 3 teams decided to use different social media applications. They used Skype messenger, Facebook messenger, and Line, which is another SNS smartphone application. Both domestic teams (Japanese and Arabic) were able to communicate without social media, but they decided to use it as one of their communication tools anyway. All teams used a text messaging system, which included the ability to attach files simultaneously, such as screenshots and word/text documents. The Arabic team communicated via Facebook messenger on the first and third cycles. In the second cycle, they met on campus to discuss, therefore no data was collected for this cycle. The Japanese team used Line, which is the most commonly used SNS application in Japan. Line application has a similar functionality with other messengers. The International team decided to use Skype (video) and Skype text messenger. Skype messenger was used as a more precise communication tool to exchange information. Having members from 2 different countries, with a 5-hour time difference meant the teams needed to discuss and fix the timing of weekly discussions at the beginning.

To examine the relationships between the SUS scores and the types of communication used, a protocol analysis was used to categorize the content. As a way of analyzing the online communications, Erkens et al [16] introduced the idea of coding dialogues in the “computer-supported collaborative learning (CSCL)” environment.

2.2.3. System Usability Scales(SUS). Each team was evaluated by another team using system usability scales (SUS) as user feedback [17]. Based on the usability tests, implementation changes were made. The experimenter assigned the participants which (other) team they needed to evaluate; each team was evaluated by the two other teams. Before starting the development cycle, the participants evaluated the initial game as well as during the three following development cycles.

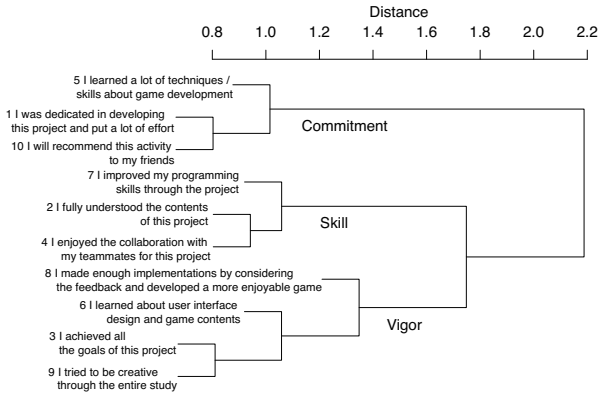

Figure 3. Clustering for responses of efficacy question items

2.2.4. Learning reflection. At the end of the study, individual participants were required to complete a questionnaire to measure their learning reflection. The survey consisted of 3 parts: 1. Self-efficacy (10 questions), 2. Satisfaction and achievement of learning, and 3. Difficulties (10 questions).

The efficacy was self-evaluated using a 5-point Likert scale. Ten (10) questions related to the learning efficacy were asked (see Figure 3) based on the motivational efficacy in previous work [18]. Note that the technical efficacy was added to measure students' understanding through the learning activities.

In addition, satisfaction and achievement of learning [6] [7] [8] were examined to measure the participants' level of emotional experience. Participants' responses included both their expectations at the beginning of the study (defined as initial) and post-development experience (defined as final) about the learning reflection. We examined these differences between the initial and final, which was affected by IPIP and/or IL.

In the online distance learning, students' distressing experiences due to communication breakdowns and technical difficulties have been observed in previous work [19]. In this study, the participants were also asked to answer questions about difficulties when they were proceeding with the project using the 5-point scale shown in Figure 4. The relationship between the learning efficacy through the participants' experiences and IPIP, IL, and the SUS scores are analyzed and discussed in Section 4.

3. Results

3.1. Characteristics and Skills

The results of IPIP and IL tests are summarized in table 1. These results shows that all teams have similar characteristics, in particular the level of programming skills and attitudes. The relationships between usability sores and these attributes will be discussed in Section 4.

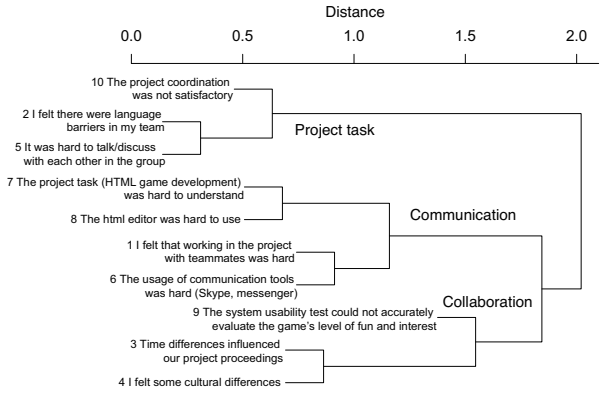

Figure 4. Clustering of responses for difficulty question items

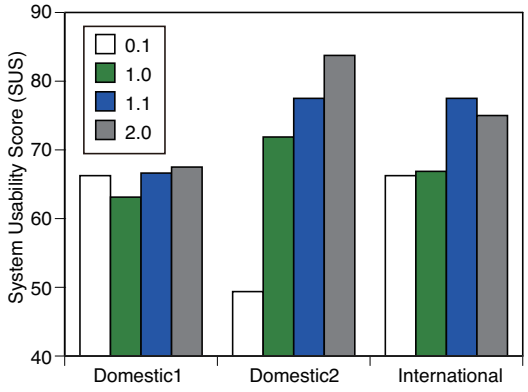

Figure 5. Comparison of SUS's changes across the three teams

TABLE 1. SUMMARY OF IPIP AND IL TESTS FOR EACH TEAM: MEAN(SD)

		Domestic1	Domestic2	International
IPIP	1:Extraversion	3.5(0.3)	3.0(1.4)	3.2(1.0)
	2:Agreeableness	4.5(0.2)	3.5(0.7)	4.1(0.4)
	3:Conscientiousness	3.3(0.3)	2.9(0.4)	3.0(0.5)
	4:Emotional Stability	3.4(1.0)	2.6(1.0)	2.7(0.2)
	5:Intellect	3.9(0.8)	3.0(0.7)	3.5(0.6)
IL	IL-SF1:Skills	4.0(0.3)	4.2(0.3)	4.1(0.4)
	IL-SF2:Attitude	3.4(0.5)	3.2(0.2)	2.8(0.3)

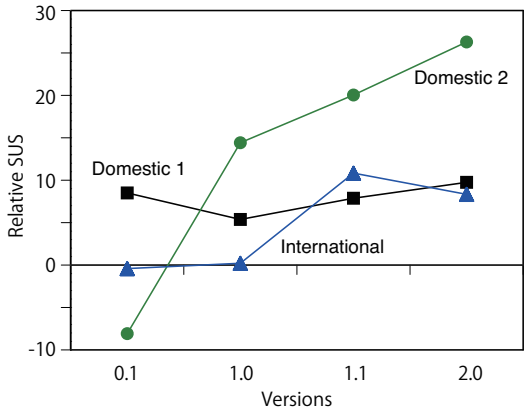

Figure 6. Changes of relative SUS scores along the stages of development

3.2. Categorizing communications

The provided SNS communications across all teams were 392 in total. The communications from all groups were divided into 2 categories; project related communications (PRC) and non-project related communications (non PRC). The purpose of this classification is to exclude first greetings, exchanging personal information, and time arrangements for the meetings.

The PRC were categorized into 4 different types of communication using protocol analysis [20]. Protocol analysis is often carried out to classify communication and dialogue [21] [22]. It is also used to evaluate social media communications [16] [19]. In this study, the social media communications were categorized as: Proposal, Permission, Encouragement, and Acknowledgment, with the following criteria:

- Proposal: Dialogues which include information about a new implementation idea.
- Permission: Acceptance of someone's proposal such as 'Okay' and 'I think so.'
- Encouragement: Communication where someone encourages other teammates, such as 'Good job' and 'Let's do our best.'
- Acknowledgment: Notification and acknowledgement when students reply to teammates' work such as 'Thank you' and 'I changed.' It also includes information about what a member did and thought based on specific discussions.

3.3. System Usability Scales (SUS)

The averages of the SUS scores are shown in Figure 5. The result of version 0.1 is the evaluation of the initial games (version 1.0 corresponds to usability after the first development cycle, 1.1 after the second development cycle, and 2.0 at the end of the development cycle). The Domestic Team 2 achieved a score of 49, whereas the other two teams achieved 66. Additionally, the overall usability scores did not improve significantly for the Domestic Team 1, unlike other teams where the usability scores have increased considerably. The most significant improvement from 0.1 to 2.0 was performed by the Domestic team 2.

Additionally, the relative SUS scores are calculated by subtracting the average of the evaluation scores for other teams from the SUS scores (summarized in Figure 6). These relative SUS scores indicate how much the participants were evaluated relatively compared to their own games.

Both the SUS and the relative SUS scores were not affected by the number of teammates, since Domestic team 1 which had four members performed the best in terms of the SUS scores although the Domestic team 2 had 3 members and the International team had five members.

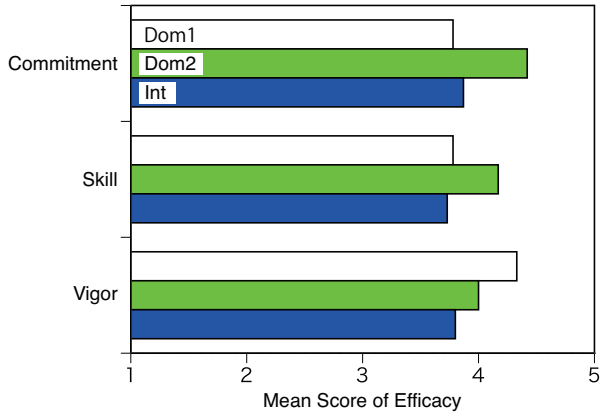

Figure 7. Mean scores of efficacy across the three teams

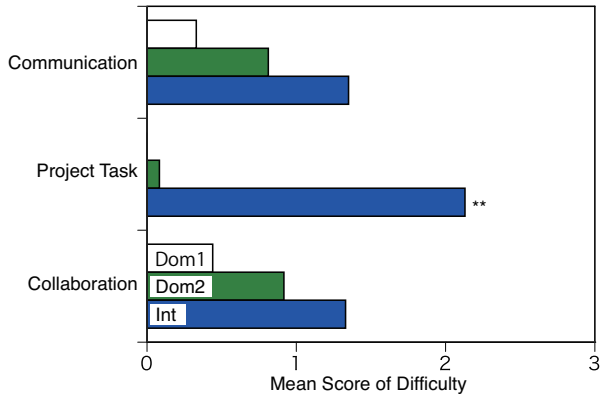

Figure 8. Mean scores of difficulty across the three teams

3.4. Learning reflection

3.4.1. Efficacy. The cluster analysis was carried out to categorize the participants' efficacies, which are illustrated in Figure 3. The efficacy variables are classified and summarized as Commitment, Skills, and Vigor. The results of the mean value of the clustered efficacy are shown in Figure 7. There were no statistical differences in efficacy between teams. Domestic Team 2 had the highest score in both commitment and skills, although Domestic Team 1 had the highest score in vigor.

3.4.2. Difficulties. The cluster analysis was conducted to classify the participants' difficulties shown in Figure 4. These variables are defined as: Communication, Project task, and Collaboration. The results of the mean value of the clustered difficulties are shown in Figure 8. The International team experienced highest difficulties among all three aspects, and compared to the other teams, they were in particular facing significant difficulties in terms of the project task (ANOVA; Tukey post hoc). It is considered that these difficulties affected their SUS scores, and the relationships are analyzed and discussed later in this section.

TABLE 2. CORRELATION COEFFICIENTS BETWEEN SUS AND INFORMATION LITERACY (N=12)

Ver.	Usability	Information literacy	
		Skills	Attitude
0.1	SUS	(0.49)	0.71
	r-SUS	0.53	0.56
1.0	SUS	(0.15)	(-.16)
	r-SUS	(0.08)	-.50
1.1	SUS	(0.26)	(-.36)
	r-SUS	(0.21)	-.61
2.0	SUS	(0.15)	-.56
	r-SUS	(0.10)	-.64

r-SUS: relative SUS
( ): not significant coefficient ($p > 0.10$)

4. Discussion

4.1. Relationships between SUS scores and attributes

To investigate the research questions in Section 1, the aforementioned results were analyzed. Correlational analyses were performed to examine any relationships between the SUS scores and the participants' attributes such as characteristics, skills, and attitudes. As a result, all factors in IPIP were not significantly correlated with the SUS scores, which means the participants' characteristics do not affect their SUS scores. Nevertheless, information skills and attitude factors from IL significantly contributed to the SUS and the relative SUS scores (see Table 2). The result shows that the learners' attitude is positively correlated with the SUS ($r = 0.71$, $p < 0.10$) as well as the relative SUS ($r = 0.56$, $p < 0.10$), and that skills are positively correlated with the relative SUS in version 0.1 ($r = 0.53$, $p < 0.10$). However, after the version 0.1, attitude is negatively correlated with the relative SUS scores ($r = -0.50$ in version 1.0; $r = -0.61$ in version 1.1, $p < 0.10$). In the cycle 2.0, attitude is also negatively correlated with the SUS scores ($r = -0.56$, $p < 0.10$) and the relative SUS scores ($r = -0.64$, $p < 0.10$). These results indicate that the participants who have high information literacy in terms of skills and attitudes lead better usability score at the beginning of the stage, but after version 0.1, high attitude causes smaller usability scores. It can be interpreted that the feedback of version 0.1 conveys stricter evaluations across teams. This interpretation appears to be supported by a sample participant survey comment:, "…the skill of my team was very high, so we could make our game enjoyable." This indicates that the level of the information skills directly affects teams' performance, and it is observable in the beginning of the implementation according to the data analysis also.

4.2. Relationships between SUS scores and Communications

The relationships between the usability scores and communications among the teams were examined to compre-

TABLE 3. CORRELATION COEFFICIENTS BETWEEN SUS AND COMMUNICATIONS (N=12)

Cycle	Usability	Communications				
		Total	Pro	Per	Enc	Ack
1	SUS	0.50	0.55	(0.32)	(0.11)	0.50
	r-SUS	(-.21)	(-.21)	(-.09)	(-.29)	(-.19)
2*	SUS	(-.11)	(-.16)	(-.16)	(-.09)	(-.06)
	r-SUS	(-.27)	(-.32)	(-.34)	(-.22)	(-.21)
3	SUS	(-.33)	(-.37)	(-.22)	(-.11)	(-.24)
	r-SUS	-.53	-.51	(-.36)	(-.09)	(-.45)

r-SUS: relative SUS; 2*, N=9
( ): not significant coefficient ($p > 0.10$)

TABLE 4. CORRELATION COEFFICIENTS BETWEEN SUS AND SCORES OF DIFFICULTY (N=12)

Ver.	Usability	Score of difficulty		
		Communication	Project Task	Collaboration
0.1	SUS	(-.48)	-.50	(-.07)
	rSUS	(-.30)	(-.13)	(0.20)
1.0	SUS	(-.37)	(-.12)	-.61
	r-SUS	(0.06)	0.55	(-.12)
1.1	SUS	(-.16)	(0.02)	-.80
	r-SUS	(0.11)	(0.48)	-.50
2.0	SUS	(0.23)	(0.33)	-.59
	r-SUS	(0.37)	0.59	(-.29)

r-SUS: relative SUS
( ): not significant coefficient ($p > 0.10$)

hend the process of the team activity (see Table 3). A correlation analysis was performed to examine the relationships between categorized communications and the SUS scores. In cycle 1, the SUS score was correlated to Proposal (r = 0.55, p $<$ 0.10) and Acknowledgment (r = 0.50, p $<$ 0.10) as well as the total number of communications (r = 0.50, p $<$ 0.10). In cycle 3, the relative SUS score was negatively correlated to the total number of communications (r = -0.53, p $<$ 0.10) and Proposal communication (r = -0.51, p $<$ 0.10). These results indicate that communication, especially Proposal and Acknowledgment, leads to better usability scores at the beginning of the project. To share new ideas and proposals, the brainstorming style discussion might be helpful in the first cycle. Nevertheless, Proposal communication in cycle 3 causes the adverse effect to the score, and it is considered because the third cycle is the last implementation stage, and it might be too late to discuss new ideas.

4.3. Relationships between SUS scores and Reflection

The correlation analysis was carried out to examine the relationships between the usability scores and the learning reflections. Although there are no significant relationships between the SUS scores and learners' efficacy, the relationships between the SUS scores and clustered difficulties were examined (shown in Table 4); there are significant relationships between the SUS scores and project task, the SUS scores and collaboration.

Project task had a negative relationship with the SUS score in the version 0.1, but had a positive relationship with the relative SUS score at the end. It is considered that project tasks were hard at the beginning, where the usability score became low. However, once the teams started to implement the games' usability, the task difficulty seems to be a contributory factor. This is because the participants struggled with the task as much as they could. This appears to be supported by the typical comment from the international team: " $\cdots$ I had a very enjoyable experience$\cdots$, it was fun to work with people on the other side of the world. " Although the international team felt the difficulty the most, in terms of the project task (shown in Figure 8), it shows also that they were able to overcome this difficulty in the end.

Moreover, we examined the negative relationships between the SUS scores and overall collaboration. When the participants felt difficulty in collaboration work, the usability scores decreased as this project was conducted using team based learning (TBL). In fact, a typical participant comment was: " $\cdots$I enjoyed the project through the teamwork. At the same time, I could improve my programming skill. " This tends to indicate that teamwork does indeed enhance participants' skill and enrich their learning experience.

There is not any significant relationships between the SUS scores and communication difficulty, and it is considered that most of the participants were familiar with using the social media communication tools and did not feel any difficulty in communication.

5. Conclusion

In this study, the contributing factors to the usability scores in the game development activity were investigated using cross-cultural TBL. Based on the usability scores, the relationships with the learners' reflections, communications and attributes were analyzed. 1) The examination between SUS scores and the participants' reflection indicated that the usability scores decreased when the participants feel the difficulties in their collaboration. In order to yield better result, teamwork is the most significant factor in TBL. 2) Based on the relationship between SUS scores and communications, Proposal and Acknowledgment communications contribute to the usability scores in the early stage of the development cycles, in order to convey better usability result. We should encourage the student to have such communications among the teams for better discussions and team activities. 3) According to the relationship between the SUS scores and the participants' attributes, the information skills were related to the implementation performance as well as SUS scores. These results show that the collaboration work and communications are the most influential factors to their performance in cross-cultural TBL. Our findings can be useful for the educational method and instructional design in HCI learning environment.

Acknowledgments

This work was partially supported by an International Collaboration Research Promotion Fund, School of Engineering, Tokyo Institute of Technology.

References

[1] L. K. Michaelsen, A. B. Knight, and L. D. Fink, *Team-based learning: A transformative use of small groups*. Greenwood publishing group, 2002.

[2] E. A. Gomez, D. Wu, and K. Passerini, "Computer-supported team-based learning: The impact of motivation, enjoyment and team contributions on learning outcomes," *Computers & Education*, vol. 55, no. 1, pp. 378–390, 2010.

[3] M. Nakayama, "Current topics in the design of hci courses with computer science curricula," in *Information Visualisation (iV), 2015 19th International Conference on*. IEEE, 2015, pp. 255–258.

[4] E. Prasolova-Førland, T. Wyeld, and T.-W. Chang, "Constructing a virtual tower of babel: a case study in cross-cultural collaboration across three continents," in *International Conference on Virtual Systems and Multimedia*. Springer, 2007, pp. 143–153.

[5] T. Wyeld, "Working with cultural differences: a case study in multicultural teamwork using a 3dcve," *Educational Games: Design, Learning and Applications. Nova Science Pub Inc., Bombay*, pp. 137–149, 2010.

[6] M. G. Moore and G. Kearsley, *Distance education: A systems view of online learning*. Cengage Learning, 2011.

[7] I. Jung, S. Choi, C. Lim, and J. Leem, "Effects of different types of interaction on learning achievement, satisfaction and participation in web-based instruction," *Innovations in education and teaching international*, vol. 39, no. 2, pp. 153–162, 2002.

[8] M. Paechter, B. Maier, and D. Macher, "Students ' expectations of, and experiences in e-learning: Their relation to learning achievements and course satisfaction," *Computers & education*, vol. 54, no. 1, pp. 222–229, 2010.

[9] R. M. Marra, J. L. Moore, and A. K. Klimczak, "Content analysis of online discussion forums: A comparative analysis of protocols," *Educational Technology Research and Development*, vol. 52, no. 2, p. 23, 2004.

[10] R. R. V. Rasiah, "Transformative higher education teaching and learning: Using social media in a team-based learning environment," *Procedia-Social and Behavioral Sciences*, vol. 123, pp. 369–379, 2014.

[11] S.-J. Chen and E. J. Caropreso, "Influence of personality on online discussion," *Journal of Interactive Online Learning*, vol. 3, no. 2, pp. 1–17, 2004.

[12] M. Ogot and G. E. Okudan, "The five-factor model personality assessment for improved student design team performance," *European Journal of Engineering Education*, vol. 31, no. 5, pp. 517–529, 2006.

[13] M. Nakayama, H. Yamamoto, and R. Santiago, "Impact of information literacy and learner characteristics on learning behavior of japanese students in on line courses," *International Journal of Case Method Research & Application*, vol. 20, no. 4, pp. 403–415, 2008.

[14] 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.

[15] 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(4), pp. 387–395, 2007.

[16] G. Erkens and J. Janssen, "Automatic coding of dialogue acts in collaboration protocols," *International journal of computer-supported collaborative learning*, vol. 3, no. 4, pp. 447–470, 2008.

[17] J. Brooke *et al.*, "Sus-a quick and dirty usability scale," *Usability evaluation in industry*, vol. 189, no. 194, pp. 4–7, 1996.

[18] M. Nakayama, K. Mutsuura, and H. Yamamoto, "Analysis of student's reflections upon the efficacy of their note-taking activities in a blended learning course," in *Information Technology Based Higher Education and Training (ITHET), 2015 International Conference on*. IEEE, 2015, pp. 1–6.

[19] N. Hara, C. J. Bonk, and C. Angeli, "Content analysis of online discussion in an applied educational psychology course," *Instructional science*, vol. 28, no. 2, pp. 115–152, 2000.

[20] A. Teranishi, M. Nakayama, T. Wyeld, and M. Eid, "Effectiveness of social media communication in game development study using team-based learning," *4th European Conference on Social M*, 2017, in press.

[21] J. A. Daly, D. J. Weber, A. L. Vangelisti, M. Maxwell, and H. Neel, "Concurrent cognitions during conversations: Protocol analysis as a means of exploring conversations," *Discourse Processes*, vol. 12, no. 2, pp. 227–244, 1989.

[22] C. G. Aleman and A. L. Vangelisti, "Protocol analysis as a method for analyzing conversational data." 1994.