Trends in EFL technology and educational coding: A case study of an evaluation application developed on LiveCode

Suwako Uehara

The University of Electro-Communications
uehara.suwako@bunka.uec.ac.jp

Edgar Josafat Martinez Noriega

The University of Electro-Communications
edgarjosaf@gmail.com

Forum

The availability of user-friendly coding software is increasing, yet teachers might hesitate to use this technology to develop for educational needs. This paper discusses studies related to technology for educational uses and introduces an evaluation application being developed. Through questionnaires by student users and open-ended discussion by the developers, the feasibility of educators to be involved in developing a new evaluation application is discussed. Results showed user satisfaction level of the application was high overall, although application updates incorporating feedback from users should be considered. Developing an evaluation application was achieved, and recommendations of design and software development are provided.[1]

Keywords: coding by educators, educational tools, evaluation criteria, academic presentations, application development

Introduction

In their guidelines for 21st century education, the Japanese Ministry of Education, Culture, Sports, Science and Technology (MEXT) (2010) encourages information communication technology (ICT) use in the classroom. Technology can be used to share resources, improve class management and make efficient use of time (Donnelly, McGarr, & O'Reilly, 2011). While Baker, Herman, and Gearhart (1996) emphasized the importance of developing a vision of how technology could be used to achieve educational goals in order to integrate technology in the classroom, integrating ICT in the classroom can be affected by first order-barriers (i.e., lack of funds), and second-order barriers related

to teachers beliefs which are deeply rooted and more difficult to change (Ertmer, 1999). Since 2000, MEXT has also encouraged curricula that develop presentation skills in classrooms, in part to provide students and future workers with skills thought necessary to keep pace with external pressures (i.e., globalization). With large class sizes, presentation evaluations are likely to be time-consuming; and language educators might feel challenged to effectively evaluate learners presentations in an L2. Thus, with the vision that technology could be used to more efficiently evaluate presentations, one language instructor and one student programmer collaborated to develop presentation evaluation software for educational purposes.

Literature review

EFL technology use

E-literacy is important for English language educators in order to adapt to the increasing technology in this digital world (Kumaravadivelu, 2013), and educators might need to be able to evaluate and correctly execute new programs (Robb, 2006; Warschauer, 2002). Classroom-based ICT might include learning management systems (e.g. Moodle, https://moodle.com/; Blackboard, blackboard.com), social network sites, cloud-based services (e.g., Google docs, https://www.google.com/docs/; Dropbox, www.dropbox.com). Social network systems might require setup to maintain a certain level of privacy for the users and thus might not be suitable in educational settings (Thongmak, 2013). According to Duncan and Chandler (2011), secure, alternative, and private social networks such as Edmodo (www.edmodo.com) are designed for educational use (del Val, Campos, & Garaizar, 2010). Thongmak (2013) investigated the technology acceptance model (Davis, 1989), namely TAM3 (Venkatesh & Bala, 2008) and found that the benefits of using ICT, educating learners on selected features, and encouraging online collaboration should be emphasized. Thus, one aspect of this current paper is to highlight the benefits of ICT use, including, if possible, learning to code for language educators.

Learn to code

Amiri (2000) argued that MA TESOL (Teachers of English to Speakers of Other Languages) courses should include training for language teachers both as consumers and producers of computer based materials. The University of Leeds and the University of Brighton provide such courses titled, *MA in Teaching English to Speakers of Other Languages and Information and Communication Technology* and *MA TESOL with ICT* respectively, however, such courses remain rare. Moreover, Chappelle and Hegelheimer (2004) indicated that teachers should have basic programming skills. Today, there is thought to be an increasing interest in programming and coding to benefit education in general. Coding allows for increased creativity, designing and remixing for the purpose of individual goals (Resnick, 2012). Several new free user-friendly applications for coding exist, and various application development platforms are accessible (Luterbach & Hubbell, 2015; Fox, 2013). Luterbach and Hubbell compared the cost, ease of use, stability and adaptability of 13 platforms and discussed the diversity and garment of applications on various online stores in a section on "Motivation for learning to develop Apps".

58   One trend for EFL instructors includes learning basic programming, and this trend is

boosted by the evolution of GUI friendly, online programming platforms such as Scratch (http://scratch.mit.edu/, Resnick, et al., 2009); Wix (www.wix.com), a desktop publishing-like website builder; and LiveCode (www.livecode.com), a programming software with English syntax. The commonalities of these platforms are that the sophisticated trial versions are accessible, free, user friendly, and include professional, creative templates and online instructional videos. Non-professionals can therefore develop products with minimal effort, tailored to the users' needs.

Presentations in the classroom

Educators incorporating learner presentations in the classroom must evaluate and provide feedback efficiently and effectively. In one study on providing customized feedback to learners for phonetic evaluation of recorded speech, it was noted that instructors face a challenge to focus on individual strengths and weaknesses manually for large class sizes, in a limited time; thus, the development of an application was suggested as a possible solution (Uehara, 2014 February). Furthermore, Uehara (2012) indicated that automated messages could reduce the teacher workload for repetitive tasks. Individual feedback often involves the replication of standard responses to the learners in different combinations; and a common form of evaluation is the use of rubrics. Rubrics provide uniform assessments, and are a constructive method to communicate expectations and performance to learners (Petkov & Petkova, 2006). Technology can transform assessment practices (Quellmalz & Pellegrino, 2009) and digitalized delivery of evaluations is increasingly popular (Buzzetto-More & Alade, 2006), thus, feedback provided to each individual learner could be achieved by compiling pre-determined automated messages in different combinations for each learner. The use of ICT could therefore greatly reduce the repetitive workload for language teachers, save time, and allow for tailored feedback for each individual learner. The technology should then allow instructors to expend their time and energy on other aspects of language teaching.

Rationale

With the increased accessibility of free real-time coding platforms, and an educational goal to provide tailored feedback to learners for presentation skills in a large EFL classes, how feasible would it be to develop an application for language education evaluation with the support of a graduate school student programmer? What precautions and advice can the developers provide to language educators who are considering to develop their own applications for the same or similar purpose? With these concepts in mind, the following research questions will be covered in this paper.

Research questions

RQ1) What presentation evaluation system can be developed by a non-programmer-EFL language instructor using a free real-time coding program with the aid of a graduate school programmer?

RQ2) How useful do students' perceive the evaluation software, and what are their opinions of it?

Method for research question one

Between September 2014 to June 2015, one EFL instructor at a national university in Japan with no previous programming experience (instructor, hence forth) and a doctorate candidate in the field of programming (programmer, hence forth) conducted weekly sessions to develop an EFL educational application. The instructor was in charge of content and design, and the programmer was in charge of design and coding using LiveCode Ltd. (n.d.). The instructor and programmer met multiple times during the week for approximately two hours each session to discuss the hardware and platform to be used, the design, content and functions, to program, and test the work in progress. The session aims were not independent of each other, and as new functions were added, design and content were modified. Learners beta-tested the application during class, in a manner that did not disturb the objectives of the course, and the application was subsequently updated for function, design and content accordingly. The total number of hours spent for the instructor and programmer to meet, discuss the design and to test the product was 169 hours; the programmer spent 81 hours to learn the programming language, and 338 hours for programming. The programmer and instructor spent approximately 100 hours to draft a manual describing the development objectives and usage. The total number of real-time testing in the classroom was 96 times to 96 participants, of which, 66 responded to surveys related to the application, a response rate of 68.75%.

The platform chosen was LiveCode, a free programming language used to develop cross-platform applications including common operating systems for mobile and desktop computers. The programing language has the ability to run and compile real-time, thus simplifying the coding process. LiveCode has a GUI interface with drag and drop functions with English-like language to enable simple coding and creation of original applications. For a code to be recognized by the computer's CPU, it must be converted from source (a high-level language) into machine code (a low-level language). This process is referred to as "compiling" the code. An Apple Inc. iPad was selected as the hardware due to its high specifications compared to other competitive products, and ease of interface with existing computer hardware, which both the instructor and programmer had prior to this study. In order to develop using LiveCode, the complete development software was downloaded from the LiveCode link: http://LiveCode.com/download (LiveCode Ltd., n.d.), and instructions on how to develop on an iPad platform were reviewed thoroughly (http://livecode/com/guides-documentation/). LiveCode support forum and related topics on free online videos were also referred to, and no other programmer was consulted in the development. The programmer had no knowledge of the LiveCode program prior to the inception of this study. The presentation evaluation application is described in the section on Results for research question one.

Method for research question two

Two groups of learners (*N* = 66) responded to an online questionnaire which described the purpose of this study, and contained a 5-point Likert scale with responses ranging from 1 (strongly disagree) to 5 (strongly agree) for 22 statements related to the design, content, speed, form and frequency, perception of presentations improvement, and understanding of the evaluation criteria. Additionally, there was one Likert scale item, with responses ranging from 1 (very dissatisfied) to 5 (strongly satisfied) for one statement on the overall

evaluation of the application, and finally, an open-ended question requesting comments to improve the application was included. The survey was bilingual (Japanese and English) and, students responded to the open-ended question in their L1 or L2. The first group of learners (Group 1) were second- and third-year science and engineering university students (n = 8), who attended a 15-week English course focusing on academic presentations skills during the spring or fall semester of 2015. Each student conducted one academic presentation between five to ten minutes with an additional two to five minutes for questions and answers, and the instructor evaluated the students based on the evaluation criteria described in the results for RQ1. Students assigned to be the timekeeper announced the length of each presentation and the instructor noted the time. The evaluation results were recorded using the application, and the results were sent to the students by email after each presentation. In November 2015, Group 1 responded to the online survey. A second group of learners (Group 2) were first-year students (n = 58) who attended a 15-week Academic spoken English course during the fall semester of 2015. Each student conducted a three-minute two-sided argument presentation. In January 2016, the instructor evaluated the students' based on the criteria described in Appendix A. Evaluations were recorded using the application, and results were forwarded by email immediately following each student's presentation. This email included a link to the same online survey received by the students in Group 1. A summary of the responses is described in the results for RQ2, and a survey sample including the average response value in parentheses is shown in Appendix B.

Results for research question one: About the application

The instructor and programmer developed a new learning application to support presentation evaluation. Tentatively named *GradeMe,* the application can send out tailored feedback and advice to multiple students over multiple classes. The complete source code and templates required to run this version of the evaluation software on LiveCode version 7.0.5 described in this paper and "*GradeMe* version α", an in-house document which provides detailed development instructions are available upon request, and the current version is described in the following sections.

The purpose of *GradeMe* is to evaluate the proficiency of EFL learners' academic presentations conducted in English. This application is operated using the touch panel on an iPad mini. A student presenter is selected from a pre-loaded list, then the evaluator scores the presenter on a scale of 0 to 6 following the definitions for six presentation evaluation criteria. The evaluation results are displayed on a pie chart and the positive results (the skills performed favorably) and negative results (skills requiring improvement) are displayed in a simple color-coded format. Evaluation criteria headings are displayed in green and red for the positive results and negative results respectively: and results can also be displayed with descriptive text. One photo can be taken during the evaluation. The evaluator can send by email the full descriptive evaluation result, an image of the pie chart, one photo taken during the presentation, and two points of advice. The email is created automatically, by retrieving selected pre-determined feedback scripts for each score and evaluation criteria systematically, thereby reducing the workload of the evaluator to compile individual feedback to the presenter, which in the past may have been done manually. The presenter can retrieve the feedback immediately following the presentation, if email is accessible, which in the past would have been received by the presenter in the next class session or at a later date. Feedback provided directly following a presentation would likely be useful for the **61**

presenter to recognize elements of presentation skills that are performed well or that the presenter can improve; and *GradeMe* enables the presenter to retain such feedback for future reference.

Setup

To start using the application, an iPad mini and a computer with Microsoft Excel were prepared. A spreadsheet (Template A) was prepared with the following details; the course name and semester, student number, name in Japanese, name in romaji, and email for each student (see Appendix C). A second spreadsheet (Template B) was prepared with empty fields for evaluation criteria, criteria definition, feedback and advice (see Appendix D). Then the texts in Appendix E1, E2, E3 and E4 were inserted in the relevant section of Template B and saved as an Excel file on Microsoft Excel. Due to proprietary rights, LiveCode can not load the file using the original format (e.g. .xls or .xlsx); therefore these files must be exported from Excel to a simple format for LiveCode such as .txt. Files are uploaded to the application (iPad mini) using iTunes, by drag and drop. In this way, users can upload class lists, feedback and advice as required, however one of the restrictions is the fixed matrix of six evaluation criteria and the 7-point descriptor scale.

Application user instructions

Login. Click the application icon to open Card #1 and click *Log In* button to open Card #2.
Note. User or ID, password input, Teachers and Student registration functions are currently not available. See Figure 1 and Figure 2.
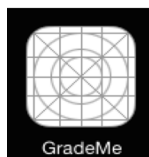

Figure 1. Application icon.

Figure 2. Card #1 login.

Class and student selection. Select *Spring* or *Fall,* then *Class* from the pull down menu and select a *student.* When the student information is highlighted light blue, click the *Grade* button to evaluate the selected student. See Figure 3.
Note. Currently, photos can only be loaded manually using the LiveCode editor, as a server connection to a host data needs to be implemented to enable such function.
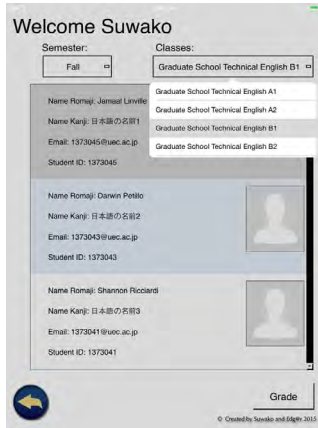
Figure 3. Card #2 student selection.

Student evaluation. Evaluate the selected student by clicking on an evaluation criterion, and select a value from 0 to 6 for each criterion. When an evaluation criterion is selected, the evaluation feedback is displayed in the window on the top left hand side, and the definitions of the criteria are displayed in the window below. The user can click on *Take Picture* to record a photographic visual during evaluations. Click the *Grade Now* button to display evaluation results. See Figure 4 where values selected are Opening Statement = 6, Structure and Originality = 2, Language = 4, Interaction =5, Confidence = 5, and Time Limit = 6.
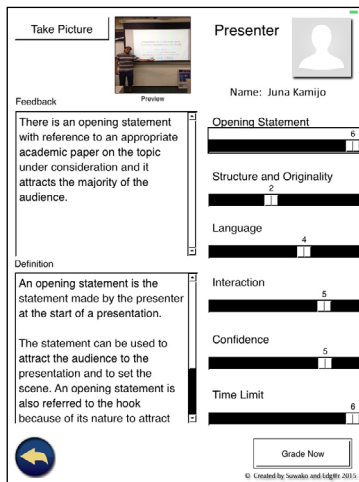

Figure 4. Card #3 evaluation.

Display evaluation results. Evaluation results are displayed as a percentage on the top **63**

left of the screen; and a value is calculated as a percentage of the total mark out of 36. If all evaluation values are 6, then final score is 100 percent. If values are those selected in Figure 4, then final score is 77%. The percentage value can be disabled by clicking on the *Disable Evaluation Percentage* box. The pie is divided into six sections, and these sections are colored more with higher scores, and less with lower scores. See Figure 5. Variations of the feedback are: Highlight criteria title of good results in green and poor results in red by clicking on *Well done* or *Try again* button. See Figure 6; Display all feedback text in one window by clicking on the *Feedback* button. See Figure 7; Display feedback text for selected results only. See Figure 8.
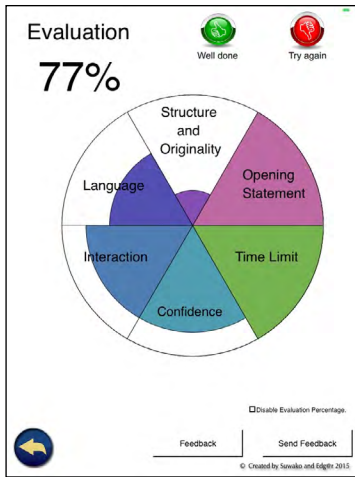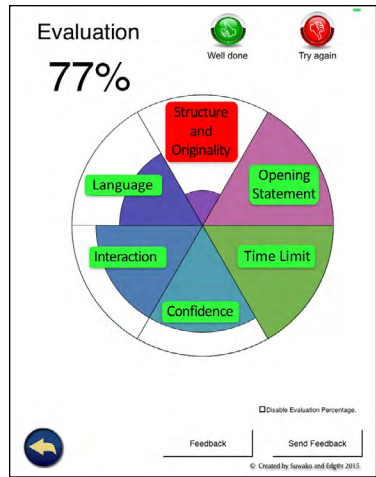

Figure 5. Card #4 results.
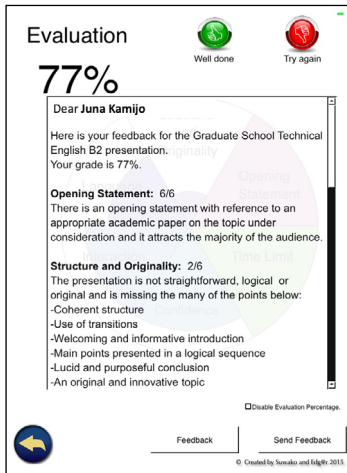

Figure 6. Card #4 Highlight results.

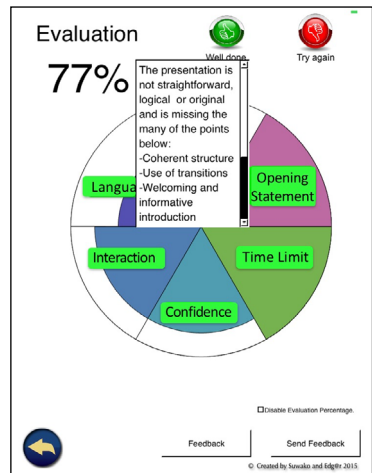Figure 7. Card #4 feedback text (all).

Figure 8. Card #4 feedback text (selected).

Send results by email. Click on "*Send Feedback*" in Card #4 to open an email. *GradeMe* opens an email and creates an email message to the student selected in Card #2; email information and name is retrieved from the pre-set class list. The automated message and can be edited before sending if required, and could also include an image of the pie chart in Card #4 with or without the green or red highlights. It is recommended to send an image highlighting the good points or bad points depending on the focus at that time. The email information and name is retrieved from the pre-set class list. A sample email message when a presenter is evaluated with score for Opening statement 6, Structure and Originality 3, Language 4, Interaction 5, Confidence 5, and Time limit 6 is shown in Figure 9.
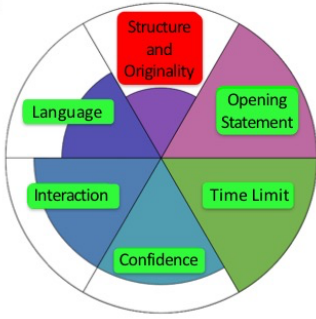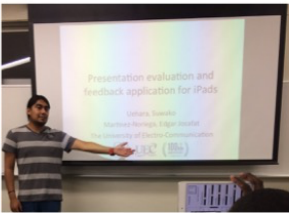

Figure 9. Sample feedback email message.

Evaluation criteria, definitions, descriptors and advice. Evaluation criteria, definition, descriptors and advice for the academic presentations were developed by Uehara (2015, February) and included feedback concepts adopted in the IELTS (International English Language Testing System, https://www.ielts.org/) grading scale where standardized descriptor feedback is provided. It also cites sections of the Common European Framework of Reference (CEFR) for "spoken fluency" and "grammatical accuracy" A1, A2, B1, B2 (Council of Europe, 2001). From these, six criteria items were identified as necessary for academic presentations in an EFL class for science and engineering students.

With *GradeMe*, a score of 0–3 on any particular criteria is considered negative: meaning the presenter requires more training and practice. A score of 4–5 means the presenter is doing well but there is room for improvement. Comparative adjectives are used to specify competence range: 0 = not been achieved; 1 = greatly lacking the described criterion; 2 = lacking the described criterion; 3 = slightly lacking the criterion; 4 = achieved the criterion but room for improvement; 5 = achieved the criterion with a slight room for improvement; Score 6 = achieved the criterion. In addition to the score description, two points of advice are provided as feedback. The grammatical structure of the advice is in the imperative form to indicate what skills should be maintained and what skills the speaker should improve. The application retrieves one piece of feedback advice from the highest scoring criterion, and one from the lowest scoring criterion. If the values are equal, the application prioritizes using the following order: Language, Interaction, Confidence, Structure & Originality, Time limit, and finally, Opening Statement. The grading score competence description for these six criteria are shown in Appendix E.

Results for research question two: Student surveys

On a 5-point Likert scale, the participant response averages are indicated in parentheses where applicable (see Appendix B). Student comments were randomly marked S1, S2, etc. to mean student 1, student 2, and so forth, where S1 to S8 were from Group 1 and S9 to S66 were from Group 2. Responses for open-ended statements by the respondents were in English or Japanese. The translated version into English marked with L1 will be shown here where applicable. Overall, for the 23 Likert-scale items, students endorsed *GradeMe* very positively (4.21). The participant responses to the statements related to the overall design showed positive results (4.29), and found the evaluation score was easy to understand (4.29), on the positive side about whether the photo aided where to stand (3.91), and whether it helped to check their own facial expressions (3.88) and gestures during their own presentations (3.86).

The students evaluated for presentations were attending seminars where the aim was to teach English through English. These concepts were indicated at the start of the course, however, multiple students still wished to receive feedback in Japanese; "I wanted to see the evaluations in Japanese" (S16 L1). The participants did however understand the evaluation and advice (4.30) and while there was the desire for the feedback to be in their L1 (3.71), they were equally positive with the feedback in English (3.71). Most participants were positive about receiving evaluations and advice instantly (3.95), were very content that the results were sent to their own email (4.52) that they were the only ones able to view the results (4.55) and they strongly endorsed a desire for a similar feedback on more than one occasion (4.24). For issues related to the participants' perceived improvement in presentation skills, while participants responded that they knew how to improve presentation skills without

receiving some sort of evaluation and advice (4.08) as a results of receiving the feedback, the participants' perception was that they knew how to improve their own presentations skills (4.30), knew what skills to work on (4.20), knew what skills they were doing well (4.03) and that this would help improve their presentation skills (4.30) Responses showed that receiving the feedback instantly helped for presentations improvement (4.03). One student said, "It was good to receive evaluation results instantly and the description rather than scores was refreshing…" (S18 L1).

Overall evaluation and open-ended responses

There were 54 statements from the participants to the open-ended question, and four categories emerged relating to; design, technical, content, and positive comments. If one statement included comments to multiple categories, they were each counted once for each category. For example, S34 wrote, "I was able to see what I was lacking (for academic presentations)" in L1, hence this statement was counted as 1 for content and 1 for positive comment. Another respondent wrote, "It was helpful to received feedback in English. I would appreciate to receive demo videos" (S59 L1), hence this statement was counted as 1 for content, 1 for design, 1 for technical, and 1 for positive comments. There were 14 comments related to the application design, eight comments related to technical issues, 12 related to content, 33 related to positive comments, and 11 that were interpreted as having no comment.

For design, one respondent found it difficult to understand the graph (S14). There were overlapping for statements related to both design and technical issues, where a couple learners thought the traditional radar chart would be an improvement (S58). One respondent wished to see a class average (S31), note though that would not be feasible when sending individualized evaluations following each presentation; and another wanted to view scores on a longitudinal scale (S3). Three students desired for either a demo video or a video of their own presentation (S15, S17, S59); and two mentioned they did not receive comments immediately (S41, S47). It was later found that this was due to either network issues or an incorrect email address. One respondent said that the photo was not viewable on his smart phone (S26). There were split opinions related to the content where one sub-group was satisfied with the clarity of the automated feedback and that it helped them to understand what they lacked in presentation skills (S12, S18, S25, S27, S34, S59, S60, S64). Another sub-group wanted the feedback to be in their L1 (S4, S16, S50, S57). Despite the goal of instructor and programmer to create general, automated feedback, two learners felt personalization was also a topic of concern, and requested for feedback that was directed more specifically to the student (S18, S33). Finally, one respondent was overwhelmed with the length of the feedback and wrote, "please more conpact" (*sic*) (S9). Out of the 33 positive comments, 24 respondents simply wrote that they thought the application was good, easy to understand, innovative, or felt technologically advanced; while others were more descriptive in their statements noting that it was a good application because "the previous systems which use paper will take more time than this one" (S19), "evaluation is very detail where I should improve. I think that point is important and want them to continue" (*sic*) (S60), and others wanted to receive this type of feedback more frequently (S39, S40) noting, "This was a good system. I wish it was used in the spring semester as well." (S39 L1). Finally, the overall evaluation from the participants of the study was very positive (4.21), and one response was, "I thought this application is highly polished" (S30), while another

was, "it was easy to see what changes need to be made, and so (I think) it is a very useful system" (S27 L1).

Discussion

The applications itself is in working condition as shown in RQ1, and it is apparent from the survey results in RQ2 that the majority of the participants are satisfied with the application; however, the current application design remains limited. The evaluation criteria and evaluation parameters are not variable, and the weight of each evaluation criterion is all equal. Structure and Originality should ideally be in a separate criteria; however, due to the limitations of the freeware in LiveCode, and to program efficiently, six evaluation parameters were a feasible solution considering the time and funds dedicated to this project. The current evaluation criterion description was constructively criticized that lower level EFL learners might not understand the language, and a simpler form could be useful. To overcome this issue, users can incorporate their own evaluation criteria and of course class list, so customization of the application is feasible within the fixed template, however not outside the template. The developers resolved the issue of the time lag in learners receiving feedback (Uehara, 2012), and the next step would be to consider solutions to store and retrieve information for multiple evaluations per presenter to enable longitudinal feedback. In addition, student observers or the presenter do not see the evaluator selection in real-time, the benefits of a digitally shared viewing of the evaluation process and evaluation results should be considered. The implementation of a server connection to a host data will enable record keeping and hosting real-time updates that can be shared amongst the class near synchronously, and this is considered for the future. *GradeMe* should also be further tested in class and studies on not only student users, but teacher users should be compiled in order to consider future updates.

Conclusion

This project was a collaboration between a language teacher with no programming skills (in charge of content and design) and a doctoral student programmer (in charge of programing and design) with visions to create an application that will enable customizable and concise feedback that can be sent instantly to the learners. The recommendations for future teacher coders are to scaffold your own learning goals for coding and if working with a programmer, communicate your visions, draw images, consider screen real estate, and expect to spend much time and start up funds on your first project. Competent teacher-coders could also create the first version, and then have a more tailored version made by professional programmers. Test frequently and keep a record of the development in the form of a development and user manual, and the educational goals will become clearer in this way.

Coding might become the norm for some language instructors, where these instructors might code their own instructional tools. Teaching methods might likely change in the same way teaching methods have changed with the use of personal computers and desktop publishing software. The growth of educational technology and user-friendly coding technology is creating an environment for instructors to befriend technology inside and outside the classroom; whether it be to use existing technology, or by creating their own, as has been demonstrated in this paper. Furthermore, the aim of having the knowledge in coding is for educators to make better judgments, and not necessarily for them to

become professional programmers (Godwin-Jones, 2015). In the ESL context, this might mean efficient integration of CALL resources. From a broader prospective, understanding code, implies one has the understanding of the internalization of methods, the analysis and the critical thought process, which is the foundation of programing, and this way of thinking might help language educators to apply these skills to learn or integrate new technology for instructional use. However, "technologies, by themselves, will not give us a place in heaven" (de Val, Campos, & Garaizar, 2010), and to understand the basic concepts of technology and coding might bring light to a possibly more seamless, or less bumpy journey of technology use in education.

Note

1. Earlier versions of this application were presented at the 17th Temple University Applied Linguistics Colloquium on the 7th February, 2015 where the primary aim was to introduce *GradeMe* in the development stage (Uehara & Martinez Noriega, 2015, February), and the updated version of the application was presented as a Show & Tell presentation at JALTCALL 2015 on 5–7 June, 2015 (Uehara & Martinez Noriega, 2015, June). Finally, learner perceptions of receiving feedback via *GradeMe* were discussed at the 18th Temple University Applied Linguistic Colloquium on the 7th February, 2016 (Uehara & Martinez Noriega, 2016 February).

Acknowledgements

References

Amiri, F. (2000). IT-Literacy for language teachers: Should it include computer programming? *System, 28*(1), 77–84.

Baker, E. L., Herman, J. L., & Gearhart, M. (1996). Does technology work in schools? Why evaluation cannot tell the full story. In C. Fisher, D. Dwyer, & K. Yocam (Eds.), *Education and technology: Reflections on computing in classrooms* (pp. 185–202). San Francisco, CA: Apple Press.

Buzzetto-More, N. A., & Alade, A. J. (2006). Best practices in e-assessment. *Journal of Information Technology Education*, 5(1), 251–269.

Chappelle, C., & Hegelheimer, V. (2004). The English language teacher in the 21st century. In S. Fotos & C. Browne (Eds.), *Reclaiming the local in language policy and practice.* (pp. 3–24). Hahwah, NJ: Lawrence Erlbaum.

Council of Europe (2001). Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge: Cambridge University Press.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.

Donnelly, D., McGarr, O., & O'Reilly, J. (2011). A framework for teachers' integration of ICT into their classroom practice. *Computers & Education, 57*, 1469–1483.

Duncan, J.C., & Chandler, P.D. (2011). A Community of practice for early career biology teachers: Social networking and digital technologies. *Contemporary Approaches to Research in Mathematics, Science, Health and Environmental Education,* 1–9.

Ertmer, P.A. (1999). Addressing first- and second-order barriers to change: Strategies for technological integration. *Educational Technology Research and Development, 47*(4), 47–71.

Fox, Z. (2013, March 13). *Learn to code for free with these 10 online resources.* Mashable. Retrieved from http://mashable.com/2013/03/13/learn-to-code-free/

Godwin-Jones, R. (2015). The evolving roles of language teachers: Trained coders, local researchers, global citizens. *Language Learning & Technology, 19*(1), 10–22.

Kumaravadivelu (2013). Afterword: Rethinking global perspective and local initiative in language teaching. In S. Ben Said & L.J. Xhang (Eds.), *Language Teachers and Teaching: Global Perspectives and Local Initiatives* (pp. 317-323). New York: Routledge.

LiveCode Ltd., (n.d.). LiveCode. Retrieved from http://livecode.com/app-in-a- day/livecode/

Luterbach, K.J., & Hubbell, K.R. (2015). Capitalizing on app development tools and technologies. *TechTrends, 59*(4), 62–70.

MEXT. (2000). グローバル時代に求められる高等教育の在り方について（答申）平成１２年１１月２２日大学審議会）[Demands in the global era: The way for high school education (Report of an inquiry) 22nd November 2000, University Council.] Retrieved from http://www.mext.go.jp

MEXT. (2010). 学校教育の情報化に関する懇談会：これまでの主な意見（第１回～第６回）21世紀にふさわしい学校や学び、学校教育の情報化が果たす役割等について、生涯学習政策局情報教育課）[Meeting to discuss the use of Information Communication and Technology (ICT) in school education: Summary of opinion from the 1st to the 6th meeting. Ideal school and study for the 21st century, the role of ICT in school education. Lifelong education policy information education department]. Retrieved from http://www.mext.go.jp

Quellmalz, E.S., & Pellegrino, J.W. (2009). Technology and Testing. *Science, 323*(5910), 75–79.

Petkov, D., & Petkova, O. (2006). Development of scoring rubrics for IS projects as an assessment tool. *Issues in Informing Science and Information Technology, 3*, 499–510.

Resnick, M. (2012). *Let's teach kids to code*, filmed at TEDxBeaconstreet. Retrieved from www.ted.com/talks/mitch_resnick_ let_s_teach_kids_to_code

Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., Millner, A., Rosenbaum, E., Silver, J., Silverman, B., & Kafai, Y. (2009). Scratch: programming for all. *Communications of the ACM, 52*(11), 60–67.

Robb, T.N. (2006). Helping teachers to help themselves. In P. Hubbard & M. Levy (Eds.). *Teacher Education in CALL.* (pp. 335–347). Amsterdam, NL: John Benjamins.

Thongmak, M. (2013). Social Network System in Classroom: Antecedents of Edmodo© Adoption. *Journal of e-Learning and Higher Education, 2013.* doi:10.5171/2013.657749.

Uehara, S. (2012). Building an online framework for self-regulated, goal directed learning. In S. Uehara, J.P. Richard & D. Beglar (Eds.). *Proceedings of the 14th Annual Temple University Japan Applied Linguistics Colloquium*, 124–134. Tokyo, Japan: Temple University Japan.

Uehara, S. (2014, February). *Feedback methods of phonetic evaluation for Japanese university science and engineering majors*. Poster session presented at the International Symposium on Innovative Research and Teaching in ESP, Tokyo, Japan.

Uehara, S., & Martinez Noriega, E. J. (2015, February). Developing an iPad based evaluation and synchronous feedback application for Class-based presentations. Paper presented at the *17th Temple University Applied Linguistics Colloquium*, Temple University Japan, Osaka Campus, Japan.

Uehara, S., & Martinez Noriega, E. J. (2015, June). Presentation and feedback application for iPads. Paper presented at the *JALTCALL 2015 Conference*, Kyushu Sangyo University, Fukuoka, Japan.

Uehara, S., & Martinez Noriega, E. J. (2016, February). Learner prospective of the synchronous iPad based evaluation and feedback application. Paper presented at the *18th Temple University Applied Linguistics Colloquium*, Temple University Japan, Tokyo Campus, Japan.

del Val, J. L., Campos, A., & Garaizar, P. (2010). LMS and Web 2.0 tools for e-learning: University of Deusto's experience taking advantage of both. *Education Engineering (EDUCON), 2010 IEEE*, 1751–1757.

Venkatesh, V., & Bala, H. (2008). Technology acceptance model 3 and a research agenda on interventions. *Decision sciences, 39*(2), 273–315.

Warschauer, M. (2002). A developmental perspective on technology in language education. *TESOL Quarterly, 36*(3), 453–475.

Author biodata

Suwako Uehara is an associate professor at the University of Electro-Communications teaching English to science and engineering undergraduate and graduate students. Her research interests are in the areas of the use of technology, evaluation methods, future use of English, motivation, autonomy, and creativity for EFL learners.

Edgar Josafat Martinez Noriega is a PhD candidate at the University of Electro-Communications. His research interests are high performance computing, real time scientific data visualization and parallel computing. Recently, he has been using a new high level cross platform programming language (LiveCode) to develop a mobile application for English presentation evaluation.

Appendix A

Evaluation criteria for a two-sided argument presentation

Criteria 1. Two-sided Argument

A two-sided argument presentation should indicate knowledge of both arguments, and indicate you are for or against the topic sentence. The speech should be organized with an introduction, main and conclusion. Each main point should have examples. The speaker should use transitions to guide the listener.

Criteria 2. Academic Reference

Academic presentations should include evidence of academic knowledge. The presenter **71**

can do this by citing and referencing academic papers which are relevant to the topic being presented. The presenter should cite and reference correctly to avoid plagiarism.

Criteria 3. Language
Language refers to vocabulary choice, grammar, and clear pronunciation.

Criteria 4. Delivery
Delivery refers to the performance of the presentation in terms of having an attractive opening statement, fluency, and interaction. The speaker should not rely strongly on printed notes or what is written on their visual aids such as PowerPoint.

Criteria 5. Time Limit
The speaker should speak within the time limit and not over the specified time.

Criteria 6. QA.
The speaker should respond in full sentences in a manner that answers the questions from the audience in a fluent manner.
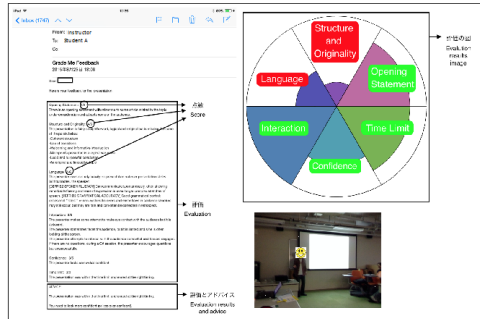
Appendix B

Sample student survey

(Average response per statement from 66 learners are shown in brackets)

**Please read statement
1 to 22 and select your
response on the following
scale.**

1 = Strongly disagree
2 = Disagree
3 = Neutral
4 = Agree
5 = Strongly agree

Image

About the design: Look at the image above.
1   Overall design is good. (4.29)
2   The evaluation score was easy to understand. (4.29)
3   The evaluation result image was easy to understand. (4.12)
4   The photo attached to the email helped me to know where I should stand. (3.91)
5   The photo helped me to check my expression when I present. (3.88)
6   The photo helped me to check my gesture when I present. (3.86)

Content
7   I understood the evaluation and advice. (4.30)
8   Evaluation and advice in English were good. (3.76)
9   I want to received evaluation and advice in Japanese. (3.71)
10 I want to receive a description of the evaluation criteria. (4.15)

Speed
11 I received the evaluation and advice instantly. (3.95)

Functions and Form
12 I want to see a video of my presentation. (3.32)
13 I want to view a good example of the presentation. (3.27)
14 I'm glad to receive the evaluation and advice directly to my own email. (4.52)
15 I'm glad I am the only one to see my own evaluation and advice. (4.55)
16 I want to receive evaluation and advice more than once. (4.24)

Presentation Improvement
17 I know how to improve my presentation skills without receiving any evaluation and advice. (4.08)
18 I know how to improve my presentation skills since I received the evaluation and advice. (3.39)
19 I know what skills I need to work on because I received the evaluation and advice. (4.20)
20 I know what skills I'm doing well because I received the evaluation and advice. (4.03)
21 By receiving the evaluation and advice, it will help me improve my presentation skills. (4.30)
22 I received the evaluation and advice instantly, so this was helpful for presentation improvement. (4.03)

Open-ended question
23 Please provide any comments that may help to improve the "GradeMe" application. No word limit.

24 Finally, what is your overall evaluation of the "GradeMe" application? Answer on the following scale.
1 = very dissatisfied, 2 = dissatisfied, 3 = Neutral, 4 = Satisfied, 5 = Strongly satisfied. (4.21)

Appendix C

Template A for class name, semester and student information

	A	B	C	D	E
	Class name and Semester	Student No.	Name in Japanese	Name in English	email
	Graduate School Technical English A2	12620710	ミズカワ サンラク	Sanraku Mizukawa	12620710@uuec.ac.jp
	Spring	12628350	ネモト カタ	Kata Nemoto	12628350@uuec.ac.jp
		12696050	オオイシ ヒサユキ	Hisayuki Oishi	12696050@uuec.ac.jp
		12666950	トクドメ アリノリ	Arinori Tokudome	12666950@uuec.ac.jp
		12664530	シムラ コウヨウ	Koyo Shimura	12664530@uuec.ac.jp
		12656480	カミジョウ ジュナ	Juna Kamijo	12656480@uuec.ac.jp
		12684290	タクドメ ギョウクショウ	Gyokusho Takudome	12684290@uuec.ac.jp
		12605620	シマザキ ケイシ	Kacey Shimazki	12605620@uuec.ac.jp
		12621220	カタブチ ウメゾウ	Umezou Katabuchi	12621220@uuec.ac.jp
		12655920	ノザキ アリノリ	Arinori Nozaki	12655920@uuec.ac.jp
		12633080	オカサワ タクミ	Takumi Okasawa	12633080@uuec.ac.jp
		12627740	ミズカワ ケンシン	Kenshin Mizukawa	12627740@uuec.ac.jp
		12621110	タカスギ ウメコ	Umeko Takasugi	12621110@uuec.ac.jp
		12690990	スギモト ソウホウ	Soho Sugimoto	12690990@uuec.ac.jp
		12678270	ノザラ テイジ	Teiji Nozara	12678270@uuec.ac.jp
		12625640	キタ クミ	Kumi Kita	12625640@uuec.ac.jp
		12606880	ナカツカ ユウノスケ	Yunosuke Nakatsuka	12606880@uuec.ac.jp
		12685060	ナカエ チコ	Chiko Nakae	12685060@uuec.ac.jp
		12646600	アオキ ウド	Udo Aoki	12646600@uuec.ac.jp
		12643090	イワクラ キヨノブ	Kiyonobu Iwakura	12643090@uuec.ac.jp
		12669320	イシグロ セイリョウ	Seiryo Ishiguro	12669302@uuec.ac.jp
		12639900	ウエシタ トモユキ	Tomoyuki Ueshima	12639900@uuec.ac.jp
		12612070	ウエハラ ニイナ	Nina Uehara	12612070@uuec.ac.jp
		12633050	ナイトウ リコ	Riko Naito	12633050@uuec.ac.jp
		12685490	タケハラ カズアキ	Kazuaki Takehara	12685490@uuec.ac.in

| ▶ | Graduate B1 | Graduate B2 | Graduate A1 | Graduate A2 | + |

Appendix D

Template B with empty fields for evaluation criteria, criteria definition, feedback and advice

			FEEDBACK							ADVICE	
			NEGATIVE				POSITIVE			SCORE VALUE	
SCORE	EVALUATION CRITERION	CRITERION DEFINITION	0	1	2	3	4	5	6	If 0, 1, 2, 3 , 4 or 5	If 6
1											
2											
3											
4											
5											
6											

Appendix E1

Descriptions for evaluation criterion, criterion definition, feedback and advice

Evaluation Criterion: Opening Statement
Criterion Definition: An opening statement is the statement made by the presenter at the start of a presentation. The statement can be used to attract the audience to the presentation and to set the scene. An opening statement is also referred to the hook because of its nature to attract the audience. The statement may include a citation to an academic paper or information demonstrating background knowledge.

Score	Feedback and Advice for Opening Statement
	Feedback
Score 0	There is no opening statement.
Score 1	There is an opening statement, but does not attract the audience.
Score 2	There is an opening statement, but attracts only a few of the audience.
Score 3	There is an opening statement, and attracts some of the audience.
Score 4	There is an opening statement with reference to some article related to the topic under consideration and attracts some of the audience.
Score 5	There is an opening statement with reference to an academic paper on the topic under consideration and attracts the majority of the audience.
Score 6	There is an opening statement with reference to an appropriate academic paper on the topic under consideration and it attracts the majority of the audience.
	Advice
Score 0, 1, 2, 3, 4, or 5	You need an opening statement with an appropriate reference to an academic paper on the topic under consideration. The statement should be used to attract the audience to the presentation and to set the scene. An opening statement is also referred to the hook because of its nature to attract the audience. The statement may include a citation to an academic paper or information demonstrating background knowledge.
Score 6	You have an opening statement with an appropriate reference to an academic paper on the topic under consideration and have been able to attract the majority of the audience.

Evaluation Criterion: Structure and Originality
Criterion Definition: A structure in a presentation should be straightforward and logical and the topic should be original. It should contain the following elements: A welcoming and informative introduction; a coherent series of main points presented in a logical sequence; a lucid and purposeful conclusion. The speaker uses transitions to aid the listeners. Originality in a presentation refers to innovative, creative and new ideas presented by the speaker, or ideas at the very least that are not plagiarized.

Score	Feedback and Advice for Structure and Originality
	Feedback
Score 0	There is no structure. The idea is plagiarized.
Score 1	The presentation is not straightforward, logical or original and is missing the majority of the points below: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.
Score 2	The presentation is not straightforward, logical or original and is missing the many of the points below: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.
Score 3	The presentation is not straightforward, logical or original and is missing the some of the points below: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.
Score 4	The presentation is fairly straightforward, logical and original but is missing the some of the points below: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.
Score 5	The presentation is straightforward, logical and original but is missing a few of the points below: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.
Score 6	The presentation is straightforward, logical and original and covers all the points below: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.
	Advice
Score 0, 1, 2, 3, 4, or 5	The presentation should be straightforward, logical and original and cover all the points below. Check which one you are missing and add: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.
Score 6	Your presentation is straightforward, logical and original and covers all the points below: Coherent structure, use of transitions, welcoming and informative introduction, main points presented in a logical sequence, lucid and purposeful conclusion, an original and innovative topic.

Appendix E2

Descriptions for evaluation criterion, criterion definition, feedback and advice

Evaluation Criterion: Language
Criterion Definition: Language refers to the level of reading presentation notes or presentation slides. EFL learners should not rely strongly on printed notes or what is written on their visual aids such as PowerPoint. Language also refers to the manipulation level of grammar and speaking fluency based on CEFR recommendations without relying on notes. Language covers how fluently the presentation proceeds, and the complexity level of grammatical expressions. The feedback text marked CEFR are text taken directly from the Council of Europe (2001) framework.

Score	Feedback and Advice for Language
	Feedback
Score 0	There is no presentation.
Score 1	The presenter is reading the presentation notes from start to finish. Without notes, the speaker: [CEFR A1 SPOKEN FLUENCY] Can manage very short isolated, mainly pre-packaged utterances, with much pausing to search for expressions, to articulate less familiar words and to repair communication. [CEFR A1 GRAMMATICAL ACCURACY] Shows only limited control of a few simple grammatical structures and sentence patterns in a learnt repertoire.
Score 2	The presenter frequently reads the presentation notes or presentation slides. Without notes, the speaker: [CEFR A2 SPOKEN FLUENCY] Can make him/herself understood in short contributions, even though pauses, false starts and reformulation are very evident. CEFR A2 GRAMMATICAL ACCURACY] Uses some simple structures correctly, but still systematically makes basic mistakes - for example tends to mix up tenses and forget to mark agreement, nevertheless, it is usually clear what s/he is trying to say.
Score 3	The presenter reads the presentation notes or presentation slides from time to time. Without notes, the speaker: [CEFR B1 SPOKEN FLUENCY] Can express him/herself with relative ease. Despite some problems with formulation resulting in pauses and "cul-de-sacs", s/he is able to keep going effectively without help. [CEFR B1 GRAMMATICAL ACCURACY] Communicates with reasonable accuracy in familiar contexts; generally good control though with noticeable mother tongue influence. Errors occur, but it is clear what s/he is trying to express.
Score 4	The presenter does not rely heavily on presentation notes or presentation slides. Without notes, the speaker: [CEFR B2 SPOKEN FLUENCY] Can communicate spontaneously, often showing remarkable fluency and ease of expression in even longer complex stretches of speech. [CEFR B2 GRAMMATICAL ACCURACY] Good grammatical control; occasional "slips" or non-systematic errors and minor flaws in sentence structure may still occur, but they are rare and can often be corrected in retrospect.
Score 5	The presenter rarely relies on presentation notes or presentation slides. Without notes, the speaker: [CEFR C1 SPOKEN FLUENCY] Can express him/herself fluently and spontaneously, almost effortlessly. Only a conceptually difficult subject can hinder a natural, smooth flow of language. [CEFR C1 GRAMMATICAL ACCURACY] Consistently maintains a high degree of grammatical accuracy; errors are rare and difficult to spot.
Score 6	The presenter does not rely on presentation notes and paraphrases clearly and concisely. Without notes, the speaker: [CEFR C2 SPOKEN FLUENCY] Can express him/herself at length with a natural, effortless, unhesitating flow. Pauses only to reflect on precisely the right words to express his/her thoughts or to find an appropriate example or explanation. [CEFR C2 GRAMMATICAL ACCURACY] Maintains consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).
	Advice
Score 0, 1, 2, 3, 4, or 5	The presenter should not rely on presentation notes and should paraphrase clearly and concisely when using presentation slides. Without notes: [CEFR C2 SPOKEN FLUENCY] Your aim is to express yourself at length with a natural, effortless, unhesitating flow. Pausing only to reflect on precisely the right words to express your thoughts or to find an appropriate example or explanation. [CEFR C2 GRAMMATICAL ACCURACY] You should maintain consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).
Score 6	You do not rely on presentation notes and you paraphrase clearly and concisely when necessary. Without notes, you: [CEFR C2 SPOKEN FLUENCY] Can express yourself at length with a natural, effortless, unhesitating flow. Pausing only to reflect on precisely the right words to express your thoughts or to find an appropriate example or explanation. [CEFR C2 GRAMMATICAL ACCURACY] Maintain consistent grammatical control of complex language, even while attention is otherwise engaged (e.g. in forward planning, in monitoring others' reactions).

Appendix E3

Descriptions for evaluation criterion, criterion definition, feedback and advice

Evaluation Criterion: Interaction
Criterion Definition: Interaction in a presentation refers to the level of speaker and audience contact. At the basic level it includes making eye contact with the audience. At advanced level it refers to the speaker interacting with the audience by for example asking questions, conducting polls, and other participation opportunities which encourages audience engagement.

Score	Feedback and Advice for Interaction
	Feedback
Score 0	There is no eye contact or interaction with the audience.
Score 1	The presenter makes very little attempt to make eye contact with the audience.
	The presenter is constantly facing away from the audience, or is constantly looking at the screen.
	The presenter does not attempt to interact with the audience and the audience is not engaged.
	If there are no questions, during a QA session, the presenter does not try to encourage questions in any way.
Score 2	The presenter makes little attempt to make eye contact with the audience.
	The presenter is often facing away from the audience, or is often looking at the screen.
	The presenter attempts to interact with the audience somewhat but the audience is not engaged. If there are no questions, during a QA session, the presenter may try to encourage questions but unsuccessfully.
Score 3	The presenter makes some attempt to make eye contact but it is awkward.
	The presenter sometimes faces the audience, but it is limited and s/he is often looking at the screen. The presenter attempts to interact with the audience somewhat and few are engaged.
	If there are no questions, during a QA session, the presenter encourages questions but unsuccessfully.
Score 4	There is eye contact with the audience. The presenter faces the audience, but usually in one direction. The presenter interacts with the audience for example by asking questions or polls, and some respond and some are engaged.
Score 5	There is much eye contact with the audience. The presenter faces the audience often and scans the room. The presenter interacts with the audience for example by asking questions or polls, and many respond, and many are engaged.
Score 6	There is consistent eye contact with the audience. The presenter continuously faces the audience and scans the room naturally. The presenter interacts with the audience on many occasions for example by asking questions or polls, and most of the audience are fully engaged.
	Advice
Score 0, 1, 2, 3, 4, or 5	You need to increase eye contact with the audience.
	You must face the audience and scans the room naturally.
	You should try to interact with the audience by, for example asking relevant questions or polls, and engage the audience.
Score 6	You have consistent eye contact with the audience.
	You continuously face the audience and scan the room naturally.
	You interact with the audience on many occasions for example by asking questions or polls, and the majority of the audience are fully engaged.

Evaluation Criterion: Confidence
Criterion Definition: Confidence in a presentation refers to the level of the speaker looking confident in the presentation. At the basic level it includes making eye contact with the audience with a certain comfort level. At an advanced level the speaker looks confident (not arrogant or presumptuous) about what the speaker is presenting. A confident appearance during the presentation is necessary to convince the audience that the new and innovative idea is what it is worth.

Score	Feedback and Advice for Confidence
	Feedback
Score 0	The presenter looks very unconfident or over confident.
Score 1	The presenter looks unconfident.
Score 2	The presenter occasionally looks confident.
Score 3	The presenter looks somewhat confident.
Score 4	The presenter looks fairly confident.
Score 5	The presenter looks confident.
Score 6	The presenter looks very confident.
	Advice
Score 0, 1, 2, 3, 4, or 5	You need to look more confident (or less over confident).
Score 6	You look very confident.

Appendix E4

Descriptions for evaluation criterion, criterion definition, feedback and advice

Evaluation Criterion: Time limit
Criterion Definition: The speaker should speak within the time limit and not over the specified time.

Score	Feedback and Advice for Time limit
	Feedback
Score 0	There is no speech.
Score 1	The presentation is far too short (far under the time limit).
Score 2	The presentation is too short (under the time limit).
Score 3	The presentation is slightly short or too long (slightly under or completely over the time limit).
Score 4	The presentation is within the time limit, but there is some time to expand and provide depth.
Score 5	The presentation is within the time limit, but there is some more time to expand and provide depth.
Score 6	The presentation was within the time limit and ended at the right timing.
	Advice
Score 0, 1, 2, 3, 4, or 5	You need to continue the presentation up to the time limit but not over.
	Try to time the presentation by expanding and providing some more depth.
Score 6	The presentation was within the time limit and ended at the right timing.