Simulation Marks (SimMarks) App: An Approach to Fix the Coursework Assessment Gap Problem Among Lecturers

**Muhammad Farid Muhammad Fathullah[1]*, Siti Isma Hani Ismail[2],
Zulfairul Zakariah[3], Farah Wahida Mohd Latib[4], Ainamardia Nazarudin[5],
Mohd Fairuz Bachok[6]**

[1]*Centre for Civil Engineering Studies, UiTM Cawangan Johor,
Campus Pasir Gudang, Jalan Purnama, Bandar Seri Alam,
81750 Masai, Johor Darul Ta'zim, Malaysia*
farid4472@uitm.edu.my

[2]*Centre for Civil Engineering Studies, UiTM Cawangan Pulau Pinang,
Campus Permatang Pauh, 13500 Permatang Pauh, Pulau Pinang, Malaysia*
ismahani.publication@gmail.com

[3]*Centre for Civil Engineering Studies, UiTM Cawangan Pulau Pinang,
Campus Permatang Pauh, 13500 Permatang Pauh, Pulau Pinang, Malaysia*
zulfairul@uitm.edu.my

[4]*Centre for Civil Engineering Studies, UiTM Cawangan Pahang,
Campus Jengka, 26400 Bandar Tun Abdul Razak Jengka, Pahang, Malaysia*
farahwahidaml@uitm.edu.my

[5]*Centre for Civil Engineering Studies, UiTM Cawangan Pahang,
Campus Jengka, 26400 Bandar Tun Abdul Razak Jengka, Pahang, Malaysia*
ainamardia0403@uitm.edu.my

[6]*Centre for Civil Engineering Studies, UiTM Cawangan Johor,
Campus Pasir Gudang, Jalan Purnama, Bandar Seri Alam,
81750 Masai, Johor Darul Ta'zim, Malaysia*
mohdfairuz@uitm.edu.my

Abstract: *Nowadays in China, the English proficiency of non-English majors in universities and colleges, in general, is not as satisfactory as expected. One of the reasons is supposed to be the sharp decrease of English input during classroom time, compared with that in junior and senior high school. The research investigates whether monolingual teaching is a possible answer to the problem and the role of the monolingual classroom in improving students' listening performance. The research is quantitative in*

nature. It adopts a quasi-experimental design, with two classes of freshmen as participants, one being the experimental group, receiving monolingual teaching while the other is the control group, which receives traditional bilingual teaching. In each class, there are 40 students. Before and after the implementation of 15 weeks' monolingual teaching, all the students have to take the pretest and posttest respectively. Besides, all the students have to take The Self-assessment Scales in China's Standards of English Language Ability before and after the implementation of monolingual teaching to see whether they themselves consider their listening ability as having improved. Together with the interview on some students, the research finds that students' listening performance in the experimental group does have a statistically significant improvement, compared with the control group and that L2 can help to change the class atmosphere and inspire students to learn English. The research serves as an empirical evidence for the effectiveness of monolingual teaching on students' listening ability, as well as an indication for the direction in how to help the reform of foreign language teaching, especially English teaching in universities and colleges in China in the future.

Keywords: *Bilingual teaching, Listening Ability, Monolingual Teaching*

1. INTRODUCTION

One of the most fundamental ways in which courses are evaluated is through course work, which is often completed throughout the semester . The final grade for a course is calculated by adding the grade for the coursework to the grade for the final exam. Grades range from 0.0 to 4.0, with the highest possible grade being an A+ or an A and the lowest possible grade being an F. Graduates will be placed into one of four categories based on their total grade point average: first, second upper, second lower, or third class. The graduate's academic transcript and degree scroll will reflect the achievement.

Coursework in the Civil Engineering Diploma program at UiTM, for instance, is evaluated using a combination of in-class and online quizzes, exams, lab reports, technical reports, fieldwork, industrial training, case study assignments, technical drawings, presentations, designs, models,

and mini-projects. In order to attain the desired objectives of the program (the "program outcome"), it is necessary to evaluate students' progress in multiple ways, including testing them across the cognitive, affective, and psychomotor domains as emphasized in the outcome-based education method ('outcome-based education'). Rather than basing evaluation solely on final exams (cognitive domain), outcome-based education incorporates qualitative assessment measures (coursework; affective and psychomotor domains) throughout the learning process (Aravind et al., 2008). A graduate's progress and competency in these areas can be inferred directly from the evaluations made during their academic careers. This is due to the fact that outcome-based education promotes higher-order cognitive skills such as analysis, synthesis, evaluation, and application (Spady and Marshall, 1991).

Assessment in outcome-based education, according to Jonathan (2017), will foster advanced cognitive abilities like critical thinking, sound decision making, and creative problem solving. In today's globalized economy, these "soft skills" acquired through outcome-based education are crucial. Higher education institutions today use a results-based education approach to teaching and learning, which Mohd Nor and Zaharim (2007) agreed would produce graduates in line with the needs of the industry and the needs of the country, provided the institution implemented the approach as outlined.

Therefore, if the review of course work does not adhere to the scoring system guidelines, the course will be inaccurately graded, whether it is graded more or less than it should be. This will not only have negative effects on the graduates, but also on the institution itself. This issue arises frequently because a study program not only has numerous courses, but also a variety of grading methods, and the majority of such programs are graded by a lecturer. Educators are the pillars of the educational system. The program and institution will fail if the lecturer cannot master the learning outcomes of the course and the objectives of outcome-based education, including teaching, learning, and assessment (Ahmad et al., 2012). According to Klien (2004), the fundamental domain of an instructor's professionalism consists of five areas of competence. The fourth domain focuses on the instructor's capacity to assess learning, student performance, and teaching effectiveness. Due to the fact that a course can only be offered during a specific semester and the same students do not take it every semester, the lecturer has no choice but to quickly acquire knowledge of the assessment domain.

Since outcome-based education was introduced to the Faculty of Civil Engineering at UiTM, scoring rubrics have been used to evaluate students in the affective and psychomotor domains using an assortment of tests, assignments, and laboratory activities, among others. Scoring rubrics are believed to be an evaluation tool that can make the assessment process appear more systematic and effective due to the clear criteria and processes involved. Additionally, rubrics can ensure that each evaluator assigns the same score for all on an equal footing. Stefl-Mabry (2004) concurred that the features of the scoring rubric were created in this manner to provide a clear, detailed, and evident basis and scoring standards. On the other hand, Pickford and Brown (2007) stated that rubrics are an excellent method to evaluate students' subjective tasks and work processes.

Examining the assessment of some coursework by comparing the evaluation of marks by lecturers and course coordinators reveals that, despite using the same marking rubric, the majority of courses differ by an average of 4 points, with the largest mark difference being 9 points. Course coordinators are selected from lecturers with extensive teaching experience and a high level of expertise. Multiple studies on the evaluation of student coursework have concluded that the lecturers' judgment applicability and the marking rubric's components are the two most important factors, as they have a significant impact on the precision and consistency of the assessment. This problem is exacerbated when the lecturers' perception of the subject assessment is minimal, and the grading rubric is poorly constructed. The survey revealed that the causes of this issue are lecturers who do not evaluate according to the scoring rubric leading to multiple lecturers evaluating a course, with an imprecise, difficult-to understand scoring rubric. The findings of this survey indicate that the lecturers and the scoring rubric in use are the primary contributors to this issue or problem.

The reputation of the educational establishment will be compromised if the level of academic performance does not adequately represent the real degree of accomplishment attained by the graduates. As a consequence of this, it is counterproductive to the purpose, vision, and overall goals of both the faculty and the university as a whole. The rate at which graduates of the afflicted school are able to find work will fall as a direct result of the

Simulation Marks (SimMarks) App: An Approach to Fix the Coursework
Assessment Gap Problem Among Lecturers

inadequate quality that is sought by employers. It is crucial that appropriate measures be taken to reduce or mitigate the effects of this problem, which are exacerbated by a number of factors, including the large number of courses offered, the variety of evaluation methods, the lecturers' competence in evaluating the students' performance, and the effectiveness of the rubrics themselves.

2.   PROPOSED SOLUTIONS

The selection of proposed solutions is determined through group ideation. This selection begins by compiling the causes for the occurrence of this problem. Next, the causes of these issues are separated into primary causes and secondary causes. Each member of the group will provide a viable solution proposal by referencing and analyzing the primary and secondary causes. Each proposed solution is discussed, and then, using a why-why analysis approach, each proposed solution will have merits and cons listed. Then, the relationship between the causes of the problems and proposed solutions is established. Using specifics, the pros and cons, and relations to the root cause of the problems the best proposed solution is chosen. The best suggested solution is application development. This is due to the fact that the production of this application can address both the primary and secondary causes of the problem.

2.1 DEVELOPMENT OF THE PROPOSED SOLUTION

This proposed solution is intended to solve the problem of evaluating a coursework grades by lecturers who differ significantly. The phases of the ADDIE model (analysis, design, development, implementation, and evaluation) are applied to the development of the proposed solution. This model is utilized because it has been demonstrated to generate practical applications. The ADDIE model is able to develop a practical project because the model phases are interconnected and there will be a repetition phase due to ineffective implementation until the goal is achieved (Mohammed Nor Azhari et al., 2023).

2.1.1 ANALYSIS PHASE

This phase collects information that will be used as input for the subsequent phase of application development. The most important information to know is the causes of these problems, the requirements of the target audience, and the goals that must be attained after using this application. This information was obtained through a survey of the intended audience, in which twenty lecturers participated. After obtaining and analyzing this information, the best solution suggested by the target group is to develop an application capable of simulating marks based on the scoring rubric, including all the information that explains the scoring rubric in depth. Due to its pervasive use, the proposed application-based solution has a positive impact on the teaching and learning process because digital applications are invaluable resources that benefit both students and educators, as the software is able to provide information, direct access to knowledge resources, and the ability to communicate and interact (Farrah, 2011; Khaddage, 2012; Kizito, 2012; Mtega et al., 2012; Suwantarathip and Orawiwatnakul, 2015). Figure 1 depicts an analysis chart of the relationship between the cause of the problem and the requirements of the target audience, which is used as a development guide to ensure that this application can accomplish its goals and be used effectively. In this phase, the primary objective is also established: when the lecturer uses this application to evaluate course work, the discrepancy in marks between the lecturer and the course coordinator should be less than 3 marks. Based on the mark assessment system used at Universiti Teknologi MARA, a difference of more than 2 marks has the potential to affect the course grade.

2.1.2 DESIGN PHASE

During this phase, the application development framework encompasses the integration of interface display and planned operation. This includes the creation of a roadmap, wireframes, establishing connections between screens, project documentation, input storage, designing screen interfaces, and creating prototypes. The sources of information utilized in the development of the roadmap wireframes application include:

Simulation Marks (SimMarks) App: An Approach to Fix the Coursework
Assessment Gap Problem Among Lecturers

I. The course syllabus serves as a valuable resource for acquiring pertinent information regarding the assessment structure and allocation of grades for a given course.

II. The utilization of a scoring rubric serves as a valuable tool in the simulation of scores, aligning with the specific criterion domain being assessed within the context of coursework.

The utilization of a chart analysis to examine the relation between the root cause of the problem and the requirements of the target demographic serves as a framework for the development of applications, with the aim of attaining its intended goals.

Roots of Problem	Needs of Target Audience
The lecturer assigns grade without adhering to the prescribed distribution of grades based on the criteria outlined in the grading rubric.	The application capables to simulate scoring and execute calculations.
The criteria domain contains words that may cause misunderstandings.	For terms that are difficult to understand, include a word list with meanings in the application.
The lecturer failed to accurately calculate the total score.	
Lacks clarity regarding the assessment of grades.	Save sample evaluations of all types of coursework in the application.
Tests and assignments which to measure taxonomy skills, are done in a short time.	The application capables to display video recordings submitted by students for assessment of taxonomy skills.
Many students are evaluated at once as part of coursework such as laboratory activities observation to gauge taxonomy abilities.	

Fig. 1 The relationship between the root cause of the issue and the needs of the audience.

2.1.3 DEVELOPMENT PHASE

The development of applications in this phase relies on the input provided by the application development framework during the design phase. MIT App Inventor 2, an open-source software, serves as a platform for the development of applications. Additionally, Google Drive serves as a platform for document and video storage. The prescribed criteria for this application necessitates a user-friendly interface, readily accessible documents, and easily attainable objectives.

2.1.4 IMPLEMENTATION PHASE

During this phase, technical testing is conducted to verify the seamless operation of the application as programmed during the development phase. This includes testing various functionalities such as storage access, video playback, and command button functionality, among others. Two lecturers from the School of Computing, Informatics and Media have been designated as co-instructors. This test enables the identification of any failed operation commands that require fixing, as well as the evaluation of the suitability of the programmed operation commands.

2.1.5 EVALUATION PHASE

Once the testing phase is completed to verify the effective functioning of the application, the subsequent and concluding stage involves evaluating user acceptance and effectiveness in attaining the objective. This objective pertains to achieving a target difference of less than three evaluation marks between the course coordinator and the lecturer. In this phase, the sample consisted of 7 individuals from the target group, including 2 course coordinators and 5 lecturers from the Civil Engineering Study Center. This application is utilized for the evaluation of coursework in two Civil Engineering Diploma courses, specifically Water and Wastewater Engineering Laboratory, as well as Hydraulics. The criteria for assessing user acceptance are outlined below, and they are measured using a Likert Scale (Mohammad Fahmi, 2015):

I. Understandability, the ease of understanding the interface of each screen, including its features and functions.
II. The user-friendliness of a system refers to the attractiveness of the interface on each screen, which includes elements such as color schemes, layouts, images, and textual content.
III. The effectiveness of performance in application development is evident in its ability to achieve objectives and address challenges that require resolution.

The assessment of goal attainment is conducted through the comparison of the performance evaluations of various coursework forms, utilizing applications administered by instructors and course coordinators. Next, a comparison is made between the score evaluations of the lecturer and the course coordinator. A comprehensive assessment is conducted on a total of five samples for each variant of coursework. These two assessments will serve as indicators to determine whether this application is deemed suitable for use or requires further enhancements in order to meet the desired standards of the target audience.

2.2 SIMULATION MARKS (SIMMARKS) APP

The development of the SimMarks App aimed to address the issue of discrepancies in grading among different instructors for a given course. The assessment method utilized in the application is domain-oriented rather than mark-oriented. This entails the selection of a specific domain for the assessment, followed by the simulation of marks based on the chosen domain. The primary objective is to mitigate discrepancies in grading among instructors for a given course by advocating for a just and efficient assessment system that eradicates bias and emotional influence. The additional functionalities encompass serving as a comprehensive resource for instructors to disseminate course-related information, while also facilitating the hassle-free distribution of recorded grades to both individual students and class cohorts. The application is compatible with Education Revolution 5.0, Industrial Revolution 4.0, and Sustainable Development Goal (SDG) #4, which pertains to the attainment of Quality Education. Figure 2 displays a collection of screen samples pertaining to the application. In the present context, Figure 3 illustrates the case diagram of the SimMarks App.
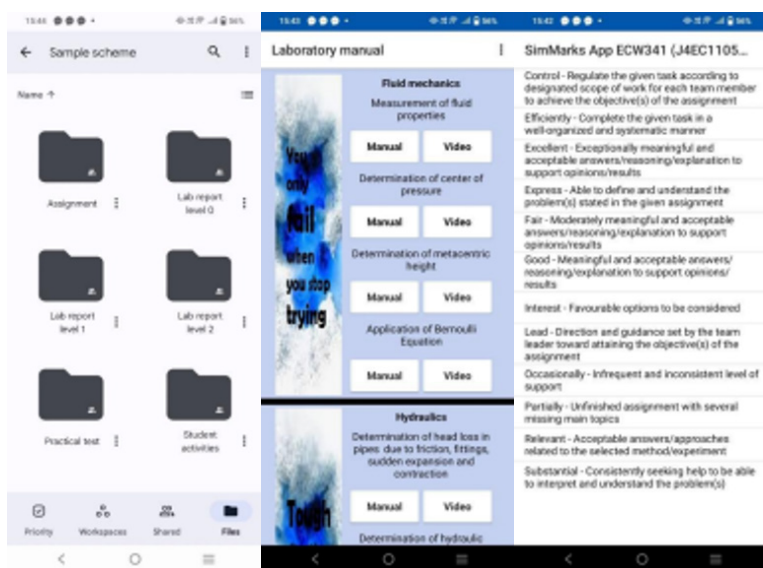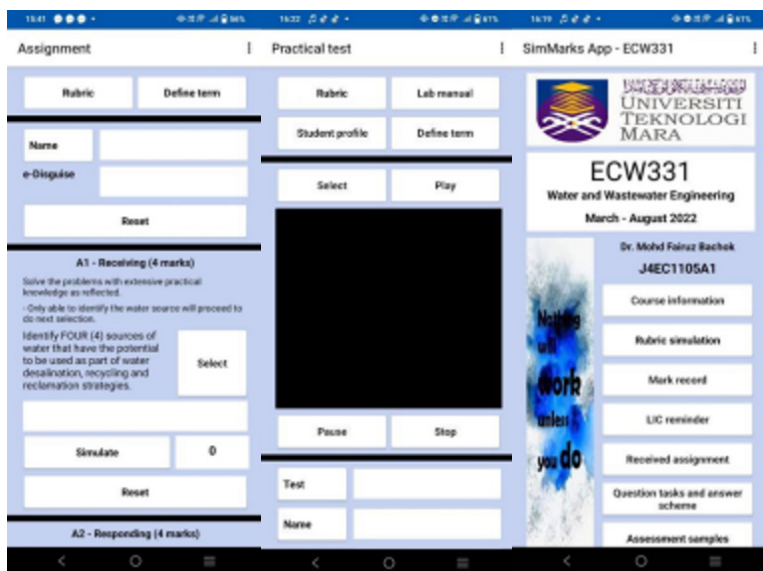
User (Lecturer)	Assessment samples	Manual
Course information	Practical test	Demonstration video Student
Rubric simulation	Lab observation	
LIC reminder	Assignment	
Laboratory manual	Lab report (Level 0) Lab report	Rubric
Question tasks and answer scheme	(Level 1) Lab report (Level 2)	Student profile Define term

Fig 2 Samples of SimMarks App Screen

3. DISCUSSION

The goal of what??? this study? Please specify… was met as efficiency and user-friendliness were both prioritized throughout the creation of the SimMarks App, which was based on the ADDIE Model and went through 5 stages. The efficiency of this application is purportedly high, as evidenced by the test results obtained during the implementation phase. These results indicate that the majority of operation commands are fully functional, with only a small portion experiencing failures. Notably, the failure rate is recorded to be as low as 20% (Table 1). Table 2 provides evidence supporting the assertion that this application is tailored to the user-friendly requirements of the target demographic. This conclusion is drawn from a survey conducted among members of the target group, wherein respondents consistently rated the application as either good or very good. Despite the efficient and user-friendly operation of the application, its utility is rendered futile if it fails to fulfill its intended developmental purpose. Nevertheless, the stated objective of this application has been successfully attained, as evidenced by the average discrepancy of 1.1 marks in coursework grades, which falls below the predetermined target for mark differences (refer to Table 3). This statement suggests that the utilization of the ADDIE model in application development can result in the creation of efficient and user-friendly applications that align with customer preferences and successfully achieve their intended developmental goals.

In addition to employing the ADDIE Model approach, which serves as the primary contributing factor in the practical development of this application, other factors were also considered. These include identifying the root causes of the customer's problems, requirements of roadmaps wireframes, conducting tests on the application's functionality, attracting the intended audience, and achieving the established objectives. Furthermore, it is imperative to enhance the application until it attains a level of satisfaction. Hence, when developing practical applications, it is essential to take into account these factors, particularly in the context of applications utilized in educational settings.

4.   CONCLUSION

A high level of competence is required of lecturers in order to demonstrate professionalism in the teaching and learning process. One of the skills that has to be acquired at a high level of competence is the evaluation of student course work. Inadequate and inconsistent evaluations can yield adverse consequences for learners, lecturers, and the institution.. The issue becomes particularly pronounced in study programs that   consist of numerous courses with a substantial weightage assigned to coursework assessment, while simultaneously requiring evaluation by multiple instructors across distinct student cohorts. It signifies that the evaluation of course work is subjective and wholly dependent on the lecturer's evaluation, specifically the evaluation of the psychomotor domain. Therefore, it is imperative to establish a mechanism that effectively mitigates any substantial disparity in the assessment of grades among lecturers. It is anticipated that the development of the SimMarks App will effectively address this issue. Therefore, it is anticipated that in future periods, this application will be utilized within UiTM campuses and potentially expanded to include all other tertiary educational institutions.

5.   REFERENCES

Ahmad, E., Razali, H., Jamaludin, H. dan Mohd Yusop, H. (2012). Peranan UTHM dalam Melahirkan Pendidik Berketrampilan. Cabaran Pendidikan Teknik dan Vokasional (PTV) di Malaysia. 1. 1 – 13. Aravind C. V., Rajparthiban and Gibert Thio (2008). Industrial Placements Through Internet Based Cooperative System. Proceedings of Fourth International Conference on University Learning and Teaching. pp. 72 – 83.

Farrah, M. (2011). Online Communication and Enhancing Language Skills, Motivation and Cultural Understanding. AUC TESOL Journal. 2. 128 – 140.

Jonathan, V. M. (2017). Implementing Outcome-Based Education (OBE) Framework: Implications for Assessment of Student' Performance. Educational Measurement and Evaluation Review. 8(1). 1 – 10.

Khaddage, F., Christoph, L. and Bray, E. (2012). *Mobile Apps Integration for Teaching and Learning. Are Instructors Ready to Re-blend. Proceedings of Society for Information Technology and Teacher Education International Conference. pp. 2545 – 2552.*

Kizito, N. (2012). *Pre-testing Mathematical Concepts with the Mobile Phone: Implications for Curriculum Design. International Review of Research in Open and Distributed Learning. 13(1). 38 – 54. Mohamed Nor Azhari, A., Syakirah, R., Zahidah, A. and Hendri, P. (2023). The Development of Mobile Application Software MyNutrient in Home Science Subject. Asian Journal of University Education. 19(1). 28 – 38.*

Mohammad Fahmi, M. Y. (2015). *MyHomePharmacy: Mobile Android Application. B. Tech. thesis. Universiti Teknologi Petronas.*

Mohd Nor, M. J. dan Zaharim, A. (2007). *Perbandingan Pendekatan Pendidikan Sepadu dalam Pendidikan Berasaskan Hasil (OBE) dengan Model Pendidikan Islam. Prosiding Seminar Pendidikan Kejuruteraan dan Alam Bina. pp. 55 – 62.*

Mtega, W., Bernard, R., Msungu, A. and Sanare, R. (2012). *Using Mobile Phones for Teaching and Learning Purposes in Higher Learning Institutions: The Case of Sokoine University of Agriculture in Tanzania. Proceedings of Fifth UbuntuNet Alliance Annual Conference. pp. 118-129.*

Pickford, R. and Brown, S. (2007). *Assessing Skills and Practice. London and New York: Routledge Taylor and Francis Group.*

Spady, W. G. dan Marshall, K. J. (1991). *Beyond Traditional Outcome-Based Education Transformational Outcome-Based Education Gives Schools a Profoundly Different Means of Restructuring Themselves. Educational Leadership. 49(2). 67 – 72.*

Stefl-Mabry, J. (2004). *Knowledge Quest. 32(5). 21 – 25.*

Suwantarathip, O. and Orawiwatnakul, W. (2015). *Using Mobile-assisted Exercises to Support Students' Vocabulary Skill Development. Turkish Online Journal of Educational Technology. 14(1). 163 -171.*