10-11-2018

Evaluation of a Re-Designed Framework for Embodied Cognition Math Games

Jonathan D.L Casano

Hannah Tee

Jenilyn Agapito

Ivon Arroyo

Ma. Mercedes T. Rodrigo

Evaluation of a Re-Designed Framework for Embodied Cognition Math Games

Jonathan DL. Casano

Ateneo de Manila University, Katipunan Ave, Quezon City, 1108 Metro Manila, Philippines

Ateneo de Naga University, Ateneo Ave, Naga, 4400 Camarines Sur, Philippines

Email: jonathancasano@gmail.com, jcasano@gbox.adnu.edu.ph

Hannah Tee

Ateneo de Manila University, Katipunan Ave, Quezon City, 1108 Metro Manila, Philippines

Email: hannah.tee@obf.ateneo.edu

Jenilyn L. Agapito

Ateneo de Manila University, Katipunan Ave, Quezon City, 1108 Metro Manila, Philippines

Ateneo de Naga University, Ateneo Ave, Naga, 4400 Camarines Sur, Philippines

Email: jen.agapito@gmail.com

Ivon Arroyo

Worcester Polytechnic Institute, 100 Institute Rd, Worcester, MA 01609, USA

Email: iarroyo@wpi.edu

Ma. Mercedes T. Rodrigo

Ateneo de Manila University, Katipunan Ave, Quezon City, 1108 Metro Manila, Philippines

Email: mrodrigo@ateneo.edu

Abstract Embodied cognition posits that the development of thinking skills is distributed among mind, senses, and the environment. Research in this field has resulted into the development of applications in different areas including mathematics. This paper reports one part of a larger series of studies on the design and implementation of embodied cognition-based mathematics educational systems. We describe the evaluation of a game called *Estimate It!*, a wearable-based game for teaching measurement estimation and geometry. Experts were invited to evaluate the game, resulting in a generally positive rating. The game's collaborative nature, its hands-on way of teaching estimation, and the incorporation of technology were seen as promising points. Infrastructure readiness, classroom control, and adjustment to the new technology were areas of concern.

1.1 Introduction

Cognition has traditionally been viewed through a narrow perspective in which the body has mostly sensory and motor functions, subordinate to central cognitive processing. In recent years though, the idea of cognition has broadened: It is now believed to be distributed among mind, senses, motor capabilities, and social interactions (Wilson and Foglia 2011). This theory of cognition, called embodied cognition, assumes that sensory perceptions, motor functions, and sociocultural contexts shape the structure and development of thinking skills including mathematical abilities and higher-ordered abstract reasoning, as well as sense-making in general (Hornecker and Buur 2006, Redish and Kuo 2015).

Research into embodied cognition has resulted in the development of applications in mathematics (Link et al. 2013), ecology and environment (Esteves 2012), problem solving (Malinverni & Burgues 2015), and other disciplines. Physically interacting with applications can involve physical tools that allow the manipulation of information and the incorporation of speech and gesture as inputs, among others. For example, Clavier is a walkable keyboard and audio device stalled along a foot path (Hornecker & Buur 2006) . As pedestrians walk along the path, they trigger colorful lights and drum beats.

Some of these studies have affirmed embodied cognition's pedagogical value, finding that children gain a better grasp of abstract representations such as number lines when asked to walk to a position instead of pointing to a value on a sheet of paper or on a blackboard (Link et al. 2013), for example. In the context of collaborative learning, embodied cognition has socio-affective impacts, creating increased perception of cooperation and reciprocal learning (Malinverni & Burgues. 2015).

Some researchers remain circumspect though. Despite the positive findings from the studies mentioned in the previous paragraphs, there is still uncertainty regarding what circumstances embodied cognition yields more intellectual or social profit than other pedagogical methods under (Esteves et al. 2013). Furthermore, the processes of designing and evaluating embodied cognition-based educational applications still lack guidelines (Esteves 2012, Schaper, et al. 2014). There is a dearth of literature providing designers with frameworks for building real-world interactions (Hornecker and Buur 2006).

In this paper, we report part of a larger study about the design and implementation of embodied cognition of mathematics educational systems. We describe *Estimate it!*, a wearable-based game for teaching size estimation. We then summarize the results of an expert review of the application. We attempt to answer the following questions:

1) In what ways is this method of teaching estimation superior (or inferior) to traditional ones?
2) What student characteristics (intellectual, social, and cultural) might make the application suitable (or not suitable) for teaching estimation?

2. Prior Work

The work described in this paper is the continuation of work described in Arroyo 2014. The Cyberlearning Watch is a device that students wear on their wrist to receive clues that help them search for geometric pieces that may be hidden. As they complete the tasks, they receive accuracy feedback via the watch's display, buzzers, or lights, making the student's experience immersive and interesting.

The watches were originally made from Arduino Uno Microcontrollers (SparkFun) connected to a WiFly wireless module that was in communication with a central server. The server, located in-situ, selected questions for each student, kept score, and logged events for teachers and students to analyze later.

The main limitation of this implementation was that each watch had to be custom-built. Once the proof of concept was tested, the next step was to improve the devices by adding NFC scanners and considering commercially available cellular phones and wearables.

This paper describes the evaluation of the original Arduino-based watch migrated to Android-based and Tizen-based system. To demonstrate the system's functionality, we specifically migrated the game *Estimate It!* (Rountree 2015) and asked a group of grade school mathematics teachers to evaluate it. In *Estimate It!*, students receive clues such as *Find a rectangle 2' wide by 14' long* or *Find a grid with rectangles approximately 3.25 in by 2.25 in.* and have to traverse a physical space searching for objects, being provided with specific measurement tools such as a 12-inch dowel without markings (which encourages rounding when participants partition it or use it to measure larger objects). The participants are given 20 minutes to find as many objects as possible using the clues given, and can work in teams with watches semi-synchronized to each other.

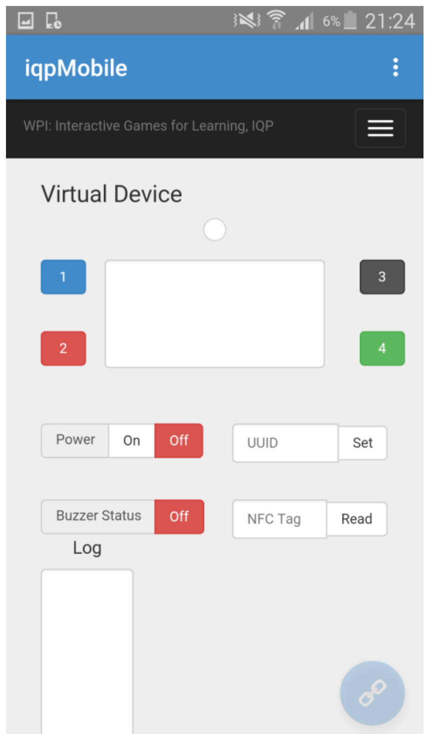

Fig 1. A screenshot of the "Estimate It!" game running on a Samsung Galaxy S5 (Android Lollipop)

3. Expert Evaluation

The goal of the expert evaluation was to obtain feedback from grade school mathematics teachers regarding the feasibility of use and implementation of *Estimate It!* in classrooms with real students. More specifically, we aimed to answer the research questions posed at the end of the first section.

3.1 Participants

A total of fourteen (14) Mathematics teachers from Ateneo de Manila University participated in the evaluation: eight females and six males. Eleven (11) of them handled grade school classes, one handled secondary school classes, while the other two were college instructors. They were of varying ages ranging from 23 to 64 years. Their teaching experiences were also diverse, with the youngest participant having taught for a year while the most experienced teacher had 27 years of experience.

3.2 Methods

The evaluation took place in a classroom equipped with wireless network connection. The server, which consisted of a laptop running on Linux that ran the web server software, was set up in a designated area in the venue. Objects that had been tagged with both NFC tags and sequences of color codes (Figure 2) were scattered around the classroom. The geometric pieces were everyday objects (e.g. a book - rectangular prism, a ball - sphere) that depicted the sought for objects, described by the clues given by the application. Two Samsung Galaxy S5 smartphones and a Samsung Gear S2 SmartWatch were used during the evaluation. These devices were pre-installed with the game and likewise made to connect to the wireless network.

Fig 2. Geometric objects used in the evaluation of 'Estimate It' (left). Color codes and NFC tags attached to geometric objects (right).

The experiment team consisted of two people who played the roles of facilitator and test monitor. Upon welcoming the participants, the facilitator described the game, how it worked, and how it might be implemented in their classes. The goals of the evaluation and the method by which it would be conducted were explained. The teachers were asked to fill out a demographics questionnaire before beginning their gameplay.

Estimate It! was designed to be played with nine participants split into three teams. During the evaluation, the teachers acted as students of a team. As the game commenced, the participants awaited clues in their devices and moved around the room to find the objects described. The facilitator and the test monitor followed them around throughout the game, explaining their progress and answering questions as necessary. Once the game was completed, the teachers were asked to answer a debriefing questionnaire.

3.3 The Debriefing Questionnaire

The debriefing questionnaire used a five-point scale, with possible responses going from *Strongly Disagree* (1) to *Strongly Agree* (5). Questions were derived from the criteria described by Whitton 2009 for effective educational design of game-based learning applications. Items relevant to the purpose of the experiment were reconstructed into questions. Table 1 shows the items classified into four groups.

Follow-up questions asking about what aspects the evaluators liked the most and least about the game as well as their insights on its advantages and disadvantages over current teaching methods were also included.

Table 1 Effectiveness of educational design of game-based learning applications questionnaire, Whitton (2009).

Active Learning Support	
1	The game encourages exploration, problem-solving, and inquiry.
2	The game provides opportunities for collaboration.
3	The game provides opportunities to test ideas and gain feedback.
4	The game provides opportunities for practice and consolidation of knowledge.

5	Game goals are aligned with the subject's learning goals.
Engenders Engagement	
1	Goals of the game are clear and achievable.
2	The game supports a high level of interactivity.
3	The game stimulates curiosity and puzzlement.
4	The game establishes the application of estimation in the real world.
5	Game levels are appropriate and challenging.
Appropriateness	
1	Game goals are aligned with the curriculum.
2	Game goals are aligned with the subject's learning outcomes and assessment.
3	A game-based approach is applicable for teaching mathematics, specifically estimation.
4	The game could be played within the allotted class period.
Classroom Use	
1	I think the game will help students learn estimation.
2	I think my students will find this game fun.
3	I think I would use this game in my classes.
4	I would recommend this game to my colleagues.

3.4 Findings

In general, most of the experts gave highly positive evaluations of the game in the questionnaire. The aggregated frequencies of each question have been summarized in Figure 3. The dispersion of responses shows typically average to little variance except for Question 5 (Q5) in the *Engenders Engagement* group of questions (see Table 1), in which one of the evaluators strongly disagreed with the statement about the game levels being appropriate and challenging. According to her, the game design did not reflect cognitive levels that indicate mathematical understanding, particularly estimation. She suggested that the levels be created based on the different types of estimation (e.g. establishing a difference between eyeball estimation and computational estimation), since one type is higher than the other according to Bloom's Taxonomy. However, evaluations were generally positive, including the answers given in the follow-up questions which are summarized next.

The results shown in Figure 3 also reveal significant circumspection and noncommitance on the part of teachers in Question 3 (Q3) of the *Classroom Use* group of questions, which is about whether teachers would be likely to use the game in their classes. After digging into their comments, we found that the reason may be the perceived difficulty in organizing the game for a big class, as well as the extra time needed to set up the game. One participant specifically raised a concern about grade-school boys needing lengthier orientations to become acquainted with the activity and its interface. Thus, future implementations of wearable embodied games such as *Estimate it!* must consider the implementation aspect as well as the mechanisms that make it simpler for teachers to implement in their classrooms.

Participant teachers found the goals of the game clear and achievable (Q1:Engenders Engagement). All expressed their agreement that a game-based approach was applicable for teaching estimation (Q3:Appropriateness). Pointing out that the generation at present was very much into interactive technologies, the teachers concurred that kids would find the game fun (Q2:Classroom Use). They saw that the use of gadgets and games would entice the students more because they were growing up in the age of technology and were more visually oriented.

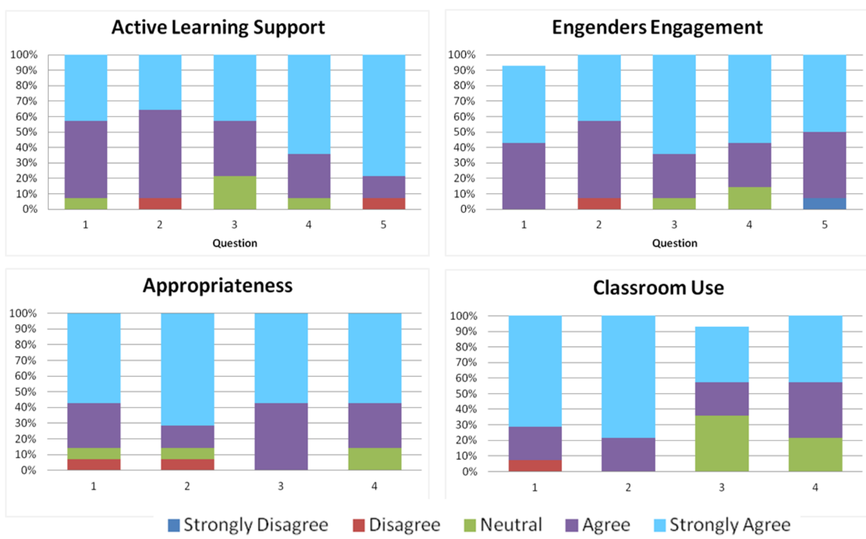

Fig 3. A graph presenting the aggregate percentages of responses from the evaluators

3.4.1 Most/Least Likeable Aspects of the Game

The evaluators liked the idea of allowing students to collaborate with one another and the opportunity provided for interaction not just with classmates but with the environment as well. The method provides a more concrete and pictorial way of presenting the concept of estimation. The use of common, ordinary household or office items was seen as a strong point because of their accessibility. Furthermore, this interaction with familiar surrounding objects was judged as being promising in its potential to allow students to physically see and associate measurements and lengths in the things that they saw daily.

On the other hand, the game's high dependence on technology was concerning for teachers. The need for a stable Internet connection and the presumption that students should already be familiar with smartphones and smartwatches led one of the evaluators to speculate that the game might be useful for private schools but was unlikely to work well in public schools in the Philippines. The installation process was likewise a concern, perceived as a possible way in which technology could impede the lesson instead of aiding it. One of the teachers disliked how the game did not require students to write their thinking processes (i.e. how they decided which object corresponded to the instruction). In her opinion, writing gives students something to reflect on regarding their performance and understanding of the matter. Another point of concern was how well the experience would work for big classes, as most teachers taught 40 or more students. Improvements to the user interface were also suggested. Some mentioned the instructions being too wordy for the small watch screen. Lastly, there was a recommendation that certain unclear terminology (e.g. a sphere which is *x* inches wide) be reviewed in the instructions. More mathematical terms (such as *circumference* to refer to a sphere's *width*) could be more appropriately used in the messages to avoid confusion.

3.4.2 Advantages/Disadvantages Over Current Teaching Methods

According to the evaluators, the use of technology would be attractive to students. Hence, they assumed that the technology and the *Estimate It!* game would be a motivator in the classroom. Some reported that the interactive nature of *Estimate It!* made it more fun and exciting than the usual classroom lecture. The game's stimulation of collaboration by allowing students to work within a team to achieve a competitive goal was a positive. Another positive aspect highlighted was that it encourages students to work and think fast within a given time frame. It was mentioned that it would help the audience see and appreciate math. The dynamic nature of the game was highlighted: The game allows students to *get their hands dirty* as opposed to passively sitting down during discussions.

The logistics and technological requirements, however, were gray areas. Respondents felt that connectivity, being a key factor in such a technological intervention, might hinder implementation since schools might not be ready for it infrastructure-wise. Further, the need to provide smartphones and/or smartwatches could be an issue affecting feasibility. It was mentioned that setting up may be tedious to some teachers and might take too much time. Also, some teachers were concerned about the need for greater classroom management than in the usual classroom setting. Lastly, there was concern that the gadgets might divert the attention of the student from learning the concept to playing with the device.

Fig 4. Math teacher being assisted by the facilitator in scanning Geometric objects around the classroom using a Samsung Galaxy S5 phone during the evaluation

3.5 Difficulties Encountered

Initially, we planned to conduct tests using the version of the game uploaded to the Heroku server. However, we discovered that the school network only had Port 80 open and did not allow the establishment of communication through other ports. As a result, we had to reschedule one of the test sessions and resorted to testing via Localhost (a server running locally in a laptop or computer in the classroom, serving WiFly). Also, since the game is highly dependent on an Internet connection, the game response was affected whenever the Wi-Fi connection in the venue was unstable. During one of the tests, we had problems with the connection and had to invite the participant to step outside so that we could get a good network signal.

4. Student Evaluation

In addition to the expert evaluation, a student evaluation was also conducted to determine the usability of the smartwatch application *Estimate It!* from the perspective of student users. The usability was measured by three main factors: (1) the extent to which students learned by using the app, (2) the ease with which the game can be used in a classroom environment, and (3) the problems and issues with the game's interface design.

4.1 Participants

Seven students—consisting of four boys and three girls, ranging from 4th to 6th grade—participated in this study. All students were familiar with the use of touchscreen-based technology. The students were also fluent in English, minimizing misunderstandings caused by language barriers.

4.2 Method

Each student placed in a controlled environment with a set number of people. A facilitator briefed the student on the app they were testing and what to do during the study. Each student took a pre-test before using the app in order to determine the students' prior levels of proficiency in the domain of estimation. The test was paper-based, consisting of two multiple-choice items, a selection item, and one or two word problems. The content of the test was modelled after tests made for a previous usability study for the application.

The student then played the app on the smartwatch until they finished the game or chose to give up. Each question in the game constituted one task in the study. In order to complete the game, the student had to answer all five questions correctly. As three users are required for progress through the levels, an observer played as the student's other teammates. This portion of the study was filmed so that it may be reviewed at a later time.

During this time, a think-aloud protocol was practiced. Additionally, observers noted down the following data to investigate the last two goals of the study. For the goal of determining the app's feasibility within a classroom environment, two metrics were noted: whether the student was able to complete each task (task completion), and the time it took them to do so (time-on-task). The time per task measurements were added to get the total time it took to complete all tasks. For the app to be deemed feasible for class use, all students had to finish all tasks, and the resulting time could not exceed 30 minutes for any student.

For the goal of determining the ease of interaction, another four metrics were observed. The first was task completion. Similar to the previous study goal, all students had to complete all tasks for the app to be deemed easy to use.

Another metric observed was the errors made, testing whether the interface was easy to navigate. During a previous usability study on the app from the perspective of the instructor, participants noted that the application's interface was difficult to navigate, with buttons on the smartwatch being too small to push. This had led to accidental button pushes, producing wrong answers. Similar error instances were expected to mostly occur within the first half of playing the game.

Student behavior was also noted. Behavioral instances of students were observed in 15-second intervals, with the dominant emotion among engagement, confusion, boredom, and frustration being noted for each interval. The total instances per student were then subject to a one-way ANOVA. Lastly, any issues which the student encountered during play were noted.

The facilitator could step in to assist the student whenever one of the following occurred: (1) The student asked for help; (2) the student had gotten the question wrong thrice; and (3) the student had taken more than five minutes on a question.

After the activity, the student took another written test. Questions in this post-test were similar but not identical to those in the pre-test (i.e. did not have the same numbers). The results of the pre- and post-tests were subjected to a paired T-test (see Table 2 for a summarized list of the metrics) to determine if there was a significant difference correlating to the playing of the game, thus revealing if learning had occurred.

After the post-test, students were interviewed and asked a number of questions to get feedback on the app and the task they needed to complete (self-report).

4.3 Results

Although all students were affected by a faulty network connection which caused delayed responsiveness, Students 1 and 2 were the most affected, leading to them not finishing all five questions of the game. Out of all the students, four opted to stop using the smartwatch and continue playing the game on a smartphone instead.

During student playthroughs, the database had to be reset multiple times due to unresponsiveness. This occurred mostly during the progression to level 2. There were instances wherein, although all three users in the team successfully input the correct answer, the app did not allow them to move to the next level, instead getting stuck on the message: *Now take your shape to the starting area and wait for the rest of your team.* From this point onwards, the game became unplayable. Resetting the database fixed this, but returned the student to the beginning of the game. This led to frustration on the part of the student as some students had to keep answering the same question, only to get stuck immediately after.

4.3.1 Pre-test and Post-test results

As seen in Table 2, all students did worse in the post-test than they did in the pre-test. The T-test showed that there were significant differences in the results of the pre-test from the post-test, implying that not only was there minimal learning involved, but that the results of the post-test were significantly lower than the pre-test. This could have been because the game did not provide opportunities for students to improve on their understanding of word problems (the part of the test which most students were unable to answer correctly). It could also indicate the students' loss of interest on the study.

Table 2 Pre- and Post-test results

Participant	Pre-test	Post-test
1	63.64%	50.00%
2	90.91%	70.00%
3	100.00%	50.00%
4	63.64%	50.00%
5	100.00%	100.00%
6	81.82%	60.00%
7	81.82%	60.00%

4.3.2 Completion and Time on Task

As seen in Table 3, Students 1 and 2 were unable to complete the game, with Student 1 only reaching the second question and Student 2 reaching the fourth. This was due mostly to the connectivity issues of the application, as the application would sometimes stop responding and the team data would not synchronize at times. After multiple attempts at getting a network connection stable enough to continue the game, the two students gave up attempting to finish answering all the questions. This implies that the game in its current condition cannot be reliably played within a regular classroom setting, and that it must be improved to enable students to use it.

Moreover, while the final five students were able to complete the game within the 30-minute target, it should be noted that their recorded times do not include the time spent on getting the application to respond. The total length of each playthrough (including the downtime caused by connectivity issues) ranged from 40 minutes to an hour. As such, a full game may not be finished within the 30-minute limit given the current build of the application.

Table 3 Task Completion and Time on Task

	Task					
Participant	1	2	3	4	5	Total Time
1	7:57:00	3:26:00	DNF	DNF	DNF	11:23:00
2	6:17:00	6:27:00	0:39:00	0:55:00	DNF	14:18:00
3	8:16:00	6:47:00	3:52:00	3:25:00	1:01:00	23:21:00
4	14:01:00	2:03:00	2:42:00	4:50:00	4:21:00	27:57:00
5	5:20:00	7:06:00	1:19:00	2:04:00	1:54:00	17:43:00
6 & 7	0:50:00	1:22:00	0:35:00	1:05:00	0:55:00	4:47:00

4.3.3 Errors

For this metric, the only instances of error recorded were when the student verbally indicated that they meant to press another button.

As opposed to the anticipated result from the previous usability test, few errors occurred while the students played, with some students expressing no errors at all. This means that button mispresses are not as grave of an issue as originally thought. Moreover, all errors occurred within the first half of the playing time, implying that any errors caused by button presses can be overcome as time passes. Despite this, some students did complain about small buttons and suggested having larger ones during the self-report.

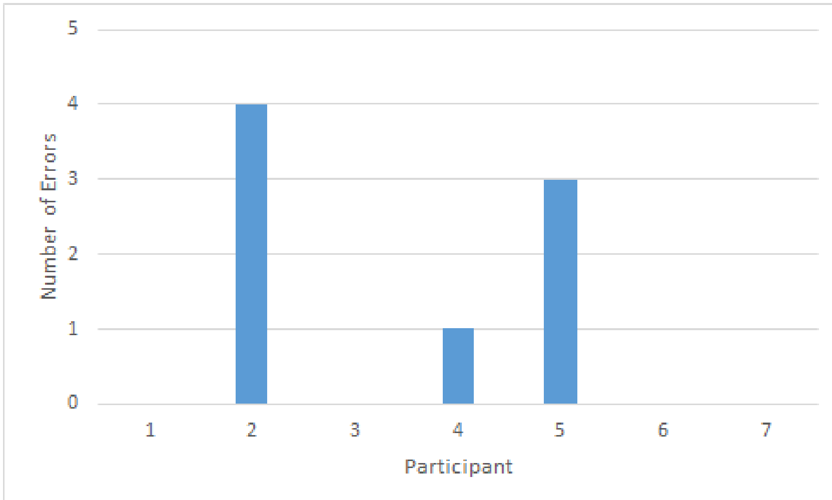

Fig 5. Number of Errors

4.3.4 Behavior Indicators

Unlike the data for Task Completion and Time-on-Task, data on student behavior included the downtimes caused by app unresponsiveness and unstable network connection, to take into consideration students' reaction to unexpected events. Additionally, while the students played for 40 minutes to an hour, student behavior was only observed during the first 15 minutes of each playthrough.

From Table 4, it can be seen that the most commonly occurring behavior is the feeling of being engaged within the game. A one-way ANOVA shows that there is a significant difference among the behaviors ($p < 0.0001$). A further Tukey mean separation test indicates that it is only the engaged behavior which is statistically significant. This implies that students were able to engage with the application despite the issues they encountered.

Table 4 Behavior of Students

Behavior	Participants						Average
	1	2	3	4	5	6&7	
Engaged	37	43	38	45	34	38	39.1666667
Bored	8	3	6	0	1	13	5.1666667
Frustrated	4	6	9	2	18	7	7.6666667
Confused	11	8	7	13	7	2	8

4.3.5 Issues

Table 5 Usability Issues

	Usability Issues		
Severity	Faulty network connection	Hard to press buttons	Takes too long to load; app stalls
Low	1	0	2
Medium	2	1	2
High	5	2	3

Most issues occurred with the faulty network connection. Measures were taken at the beginning of the study to reduce network connection issues by having three different networks to connect to, but all three networks still caused the application to load slowly and become unresponsive. This led to long playing times and the inability to complete tasks for some students.

Some students also had difficulty pressing the buttons. Specifically, the buttons were so small that the students could not see the icons while their fingers hovered over the buttons. Consequently, students were unsure whether the button was actually being pressed.

Finally, the long loading times and random stalling of the application might have been due to the network connectivity. The long loading time caused students to press buttons multiple times under the assumption that the watch had not detected their finger when they pressed the button. This led to double presses which made the students' answers wrong.

4.3.6 Self-Report

The ratings of the students were generally positive. They found the application fun and wanted to use it in the classroom with their classmates. The only dissenters were the final two participants, who stated that *[their] class was noisy enough without it.* This is an important point to consider during classroom implementation. Some students also commented that the directions were easy to follow, allowing them to get a better grasp of the app despite having never tried it beforehand.

Another point to note is the students' responses when asked about their preferred device. None of the participants preferred the use of the smartwatch, mostly citing the larger screen size of other devices as a reason. Should this project continue to be deployed on a smartwatch, some care must be taken in the design of the interface so that the screen size is not too much of a point of contention due to the limited space it provides.

5. Conclusion and Future Work

This paper presented the evaluation of *Estimate It!* as part of a larger study about the design and implementation of embodied cognition of mathematics educational systems.

Originally designed to run on Arduino Uno Microcontroller Cyberlearning Watches (Arroyo et al., in press), *Estimate It!* has now been successfully migrated to both AndroidOS and TizenOS. Specifically, the game can now be played on phones running on Android 2.3 (Gingerbread) or higher and on watches running on Tizen 2.3.1. The phone version is now capable of playing the game via NFC. Server code has also been refactored, deployed, and tested to run in a public domain.

Evaluation objectives were two-fold. This study aimed to investigate how the game compared with traditional instruction methods and what cultural, social and intellectual characteristics made the application suited (or not) to the teaching of estimation. Math teachers were invited to appraise the game's value in terms of the objectives. Most of the experts gave positive evaluations of the game.

In general, the evaluator teachers liked the collaborative nature of the technology and enjoyed how the game allows students to interact both with their classmates and their environment. There was appreciation for how the overall teaching method provided a more concrete and hands-on way of presenting estimation concepts as well as how the method allowed the teacher to monitor and move around the classroom together with the students, allowing for active participation and communication between students and teachers and among students themselves. The ability to include common household objects in the game had its appeal as it made the game seem accessible and easily customizable. The evaluators estimated that the game's competitive nature would make the method *more fun and exciting* than the usual classroom lecture. It was a common guess among the evaluator teachers that the injection of technology would be seen by students as attractive and motivating to them. These strong points made *Estimate It!* seem like a promising aid for the teaching of estimation concepts as part of the curriculum.

At the same time, there were areas of concern raised by the evaluators that the authors of this study consider *cultural* concerns. There was a common worry that implementation of the game in schools might be hindered because schools may not be financially ready for the required technology infrastructure. In particular, the need for a stable Internet connection was repeatedly mentioned. Acquiring devices (smartwatches or smartphones) for students to use could already be difficult. Other evaluators worried that setting up the game might be too tedious for some teachers, while one particularly disliked how playing the game removed the need for students to write down their mathematical working. One evaluator urged the researchers to reflect on how the game could accommodate classes of 40 students. Evaluators also dwelled on the propensity of students to misbehave while playing and on the possibility of students going off-task because of other interesting applications on the device.

In summary, evaluators seemed open to the idea of adopting a game-based reinforcement in their math classes. As one respondent put it, *when executed well, the advantages of* Estimate It! *will outweigh the disadvantages*.

As for the game's feasibility within a classroom setting, the application is still not ready to be deployed in its current state. From the student evaluation, it is possible that issues such as network connectivity will hinder students from efficiently completing tasks and cause large delays, making the game go on longer than it should be played.

Lastly, regarding its ease of use from the point of view of the students, the network connectivity once again might prevent students from completing the tasks. Other issues reported by and observed from students included small buttons, difficult-to-navigate screens, and slow responsiveness. Moreover, while errors in button pressing were not as frequent as previously feared, some improvements to the interface can still be made, as per the suggestions of the participants. However, the behavior of students indicated that they were able to engage with the application and not be completely frustrated by the issues they faced. Participants also reported that the game was generally fun, though they would rather play on a device with a larger screen such as a smartphone or a computer.

Moving forward, an iteration of the game that takes into consideration the findings of this study may be created and retested with evaluators and students to assess how they respond to this new way of teaching estimation. Given that previous evaluations had respondents that came from the United States, perhaps an evaluation of how this new method may affect the performance and understanding of estimation of Filipino students can be conducted. Results can then be compared to discover similarities and differences.

Acknowledgments

We would like to thank the teacher experts and students for responding to the invitation to evaluate the game. We thank the Ateneo Laboratory for the Learning Sciences (ALLS) for the acquisition of items needed for the migration and the experiments (Samsung Gear S2, respondent tokens, toys and NFC tags etc.). We thank the Department of Information Systems and Computer Science (DISCS) for allowing us to use computer laboratories in order to conduct the evaluation.

References

Whitton, N. (2009). Learning and teaching with computer games in higher education. Game-based Learning Advancements for Multi-Sensory Human Computer Interfaces: Techniques and Effective Practices. Thomas Connally, Mark Stansfield and Liz Boyle editors. Information Science Reference, Hershey-New York.

Esteves, A. (2012, February). Designing tangible interaction for embodied facilitation. In Proceedings of the Sixth International Conference on Tangible, Embedded and Embodied Interaction (pp. 395-396). ACM.

Rountree, W. L. (2015) Redesigning Traditional Children's Games to Teach Number Sense and Reinforce Measurement Estimation Skills Using Wearable Technology (Master's Thesis).

Hornecker, E., & Buur, J. (2006, April). Getting a grip on tangible interaction: a framework on physical space and social interaction. In Proceedings of the SIGCHI conference on Human Factors in computing systems (pp. 437-446). ACM.

Wilson, R. A., & Foglia, L. (2011). Embodied cognition. Accessed 26 April 2016 from http://stanford.library.usyd.edu.au/entries/embodied-cognition/http://stanford.library.usyd.edu.au/entries/embodied-cognition/

http://stanford.library.usyd.edu.au/entries/embodied-cognition/
http://stanford.library.usyd.edu.au/entries/embodied-cognition/

Malinverni, L., & Burguès, N. P. (2015, June). The medium matters: the impact of full-body interaction on the socio-affective aspects of collaboration. In Proceedings of the 14th International Conference on Interaction Design and Children (pp. 89-98). ACM.

Redish, E. F., & Kuo, E. (2015). Language of physics, language of math: Disciplinary culture and dynamic epistemology. Science & Education, 24(5-6), 561-590.

Esteves, A., van den Hoven, E., & Oakley, I. (2013, February). Physical games or digital games?: comparing support for mental projection in tangible and virtual representations of a problem-solving task. In Proceedings of the 7th International Conference on Tangible, Embedded and Embodied Interaction (pp. 167-174). ACM.

Schaper, M. M., Malinverni, L., & Pares, N. (2014, November). Participatory design methods to define educational goals for full-body interaction. InProceedings of the 11th Conference on Advances in Computer Entertainment Technology (p. 50). ACM.

Link, T., Moeller, K., Huber, S., Fischer, U., & Nuerk, H. C. (2013). Walk the number line–An embodied training of numerical concepts. Trends in Neuroscience and Education, 2(2), 74-84.

Android NFC and Mobile Development Stacks for Web. Accessed 20 May 2016 from https://developer.android.com/reference/android/nfc/package-summary.html

Tizen native application development tutorials and developers forum. Accessed Feb 1, 2016 from https://developer.tizen.org/forums/native-application-development/active