

Exploring the Role of Mobile Games in a Blended Module of L2 Vocabulary Learning

Saeed Khazaie¹

*PhD in Applied Linguistics, Shahid Chamran University of
Ahvaz, Khuzestan, Iran*

Alireza Jalilifar

*Professor, Shahid Chamran University of Ahvaz, Khuzestan,
Iran*

Received on January 25, 2014

Accepted on June 20, 2014

Abstract

Mobile-game (m-game)-enhanced learning offers beneficiaries a modern and attractive window of opportunity to practice and learn language skills and subskills. Through grafting m-games onto the blended module of second language (L2) vocabulary teaching, this study endeavored to help educators to see if an m-game-enhanced didactic platform is applicable for L2 teaching and assessment. One-hundred fifty two females and males from four classes at Ahvaz Jundishpur University of Medicine were selected through purposive sampling to participate in this study. To embrace triangulation, data were aggregated via of formative and summative assessments of the participants' performance, attitude questionnaire, as well as a semi-structured interview. After participating in Vocabulary Levels Test (VLT), the participants answered the preresearch questionnaire. Afterwards, for a period of an 18-blended-session course, they were taught the materials by their instructors inside the classroom and then practiced them through the m-games individually or collectively in dyads in the extramural setting. At the end of the course, participants took part in summative tests of English Vocabulary Recognition and Recall (EVRR) followed by the interview with the instructors. For data analysis, descriptive statistics as well as Olkin and Finn,

¹Corresponding Author: saeed.khazaie@gmail.com

Spearman correlation, and paired samples t-test were run. Descriptive and inferential analyses showed the tendency of the participants and instructors towards collective m-game-enhanced practicing and the prediction power of formative assessment ensued from it on the one hand and accordingly the statistically significant effect of the manner of m-gaming on the anticipatory power of the formative assessment on the other. As a consequence, the findings underscored the urgent need for collaboration in the process of m-game-enhanced L2 vocabulary learning and assessment.

Keywords: blended method, collaborative and individual practicing, m-game, vocabulary learning

1. Introduction

Both formal and nonformal didactic settings as man-made learning situations can yield some serendipitous edges for language learning. However, the average level of English that students of English as a foreign language (EFL) learn in the conventional classroom is insufficient to enable them to function effectively. Mobile technology would give students more opportunities for exposure to second language (L2) input and interaction (Banados, 2006). In this fashion, L2 learning in nonformal settings of information technology (IT)-mediated instruction is consistent with full immersion in L2, by communicating only through the medium of foreign language (Czpielewski, 2011).

Johnson (2013) contends that "it is increasingly evident that policy makers and educators should not ignore the fact that mobile technology holds great potential for student engagement, continuous formative assessment, and authentic learning experiences" (p. 1). As a matter of fact, these properties allow leveraging mobile technology in the way of teaching and learning to augment the pedagogical process realized in the real world. Owing to the ubiquity of mobile-enhanced learning, over the course of the last twenty years, the application of situational judgment tests (SJTs) to human resource selection has become

an increasingly admired trend (Chao & Chao, 2011). SJTs can provide real working samples and have better predictive power when compared to knowledge-based tests, with regard to actual performance (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001). Furthermore, when assessments are utilized for formative purposes, they impart information directly to students to tell them about the result of their learning and performance, and to provide route for improvement of their subsequent actions (William & Black, 1996). Along these lines, Galmin (1995) argues that IT-enhanced education flourishes in two quite different directions of individual flexible teaching model, where students are allowed to start their progression at any time, and the extended classroom model which assumes that the students are organized into groups to meet each other in the virtual situation through the medium of technology. In a similar vein, Chapelle (2003) and Ellis (2000) propose that mobile-enhanced language learning tasks are generally developed in two types to engage students; that is, interpersonal communication and learner-mobile communication.

Zhu, Meng, Wang, and Zhang (2011) maintain that on the account of the fact that virtual learning circles provide an interactive learning environment for students, they can participate in the discussion of the circles in keeping with their area of interest, and exchange learning experience with their counterparts who have the same interest. For the meantime, receiving feedback from peer's part is important as it makes way for students to better gauge their own performance. It also encourages students to take a turn for better learning as they know that their work will be seen and evaluated by their partners (Tan, 2011). Squire (2003) is also of the opinion that "advances in assessment, such as peer-based assessment or performance-based assessment provide learners multiple sources of feedback based on their performance in authentic contexts" (p. 4). This way, incorporating a variety of rich interactions and

communications in the form of games and puzzles provides students a high level of interest and excitement in the pedagogical modules containing them (Pascall, 2013). Wierzchok (1994) emphasizes individual manner of practicing is a distorted view of reality. Mobile games (m-games), as a predominant mobile application (app) in conjunction with suitable instructional methods have pedagogical potential (Shih, Hou, & Wu, 2011).

M-games reinforce everyday learning opportunities and are powerful tools for fostering cordiality among students (Mehrotra, 2013). Sundre and Wise (2003) reported that when students are more motivated, assessment results are a better reflection of their ability. Moreover, because of the key characteristics of real-world phenomena in games many researchers have been encouraged and educators have become more interested in leveraging the potential of game in education. In didactic m-games, the ability to apprehend swift and faultless data about each student's performance opens new avenues to new modes of gauging students' progress and performance, in ways that reinforce engagement. This is accomplished by simultaneously assessing the students' performance.

This, in turn, helps educators to align assessment and practices with new standards. Belland (2012) states "in contrast with the classical games, many newer games are rooted in the social cognitive theories of scholars like Vygotsky (1962), according to which learning results from solving problems in collaboration with others" (p. 32). In fact, by virtue of m-games, assessment has become ongoing and students' awareness of their status is not summarized into the test time (Bajpai, 2013). Belland (2012) declares that "in assessments embedded in [m-]games, game developers can collect the data they need to optimize the didactic tasks and provide evidence of learning" (p. 40). Assessment procedures used during the instruction are then needed to be integrated as learning exercise, rather than an

exercise in memorization at the end of the course (Paige, Jorstad, Siaya, Klein, & Colbay, 2011). According to Schrader and McCreery (2012), in the game-enhanced language learning situations "assessment practices can be derived from the process of (a) evaluating curricular objectives, (b) identifying the affordances of games, and (c) drawing pedagogical connections among them" (p. 24).

Gee (2003) opines that successful games and assessments are grounded on foundational principles of human learning. However, in view of the fact that current generation of students has grown with standardized tests and has the belief that these evaluations can accurately gauge their knowledge (Elan, Stratton, & Gibson, 2007), here the assessment of language learning is becoming more sophisticated, shifting from excessive dependence on pretest assessment throughout the learning experience and using alternative materials (Murphy, 1988). For that reason, it is necessary virtual games stick on the customs as a social activity (Huizinga, 1987).

To all intents and purposes, one major challenge for educators who want to weld m-games into the pedagogical process involves making justifiable inferences of anytime, anywhere, and at various levels of proficiency (Shute & Ke, 2012). In view of that, the main issue then is for educators and practitioners to properly lead the m-game-enhanced L2 learning projects to publish gainful settings on these virtual practicing platforms which add to the quality of their assessment of students' performance. Designing appropriate assessments is therefore predominant to designing m-games (Rupp, Gushta, Mislevy, & Shaffer, 2010), because simply assessing students' ability to perform tasks may not fully measure the intended impact of an instruction (Messick, 1994).

Because didactic game is intended to ease the students' access to the materials delivered inside the classrooms (Jantke, 2012), the purpose of this study was to investigate how constant

assessment of students' performance can be proved effective in the process of practicing L2 materials in the m-game-based blended L2 learning module. To that end, native and ready-made m-games were fully utilized in tandem with conventional face-to-face instruction in the blended module of L2 vocabulary learning. Investigating the impact of welding m-game in the blended module means, first, to identify an instance of learning as a result of practicing (individual vs. collective) materials via m-games, second, to extract students' attitudes towards the m-games, and third, to see whether the m-game type has important impact on the anticipation power of the assessment resulted from them or not.

2. Research Questions

This study intended to evaluate predictive power of formative assessment derived from different manners of practicing m-games (viz. m-gaming) to see if leveraging the accessibility of materials through m-games into the blended module of L2 learning can result in a better gauging of students' performance. Thus, the overarching research questions are as follows:

1. Under which scenario(s) students' formative test scores in the m-games of the blended module of L2 vocabulary learning indicate the amount of materials they would subsequently learn?
2. To what extent do the students' attitudes towards the manner of practicing accrue formative assessment results that can be matched with their subsequent assessment?

The second question can be broken down into a number of corollaries in order to capture the following interactional effects:

- I. Students' favorable attitudes towards using m-game-based L2 learning and their performance due to m-gaming;
- II. Students' favorable attitudes towards collective (vs. individual) practicing and their performance as a result of collective m-gaming;

III. Students' favorable attitudes towards collective practicing and the prediction power ensued from the formative assessment of their performance in the collective m-gaming.

3. Method

3.1 Participants

The participants in this study were 152 students of medicine (101 female & 51 male) within the age range of 19-25 from four classes enrolled in a mandatory course in the first semester of 2013 at *Ahvaz Jundishpur University of Medical Sciences*. In effect, they were purposively selected and homogenized through conducting VLT from among 160 students of medicine (see Instruments, VLT). To be exact, eight students, whose scores fall one standard deviation below the mean or one standard deviation more than the mean (Mean \pm 1 Standard Deviation) (42.6 ± 4.54), were kept out from taking part in the major phase of the study (mean = 42.61, Standard Deviation = 4.54). The details regarding the participants are tabulated in Table 1.

Table 1

Details about the Participants

Class	Male	Female	Number	Dyads
1	14	26	40	20
2	12	28	40	20
3	12	24	36	18
4	12	24	36	18

For the purpose of collective m-gaming participants were randomly further divided into dyads.

3.2 Instruments

This study made use of a triangulation research design consisting of an attitude questionnaire (preresearch survey questionnaire), and formative assessment of students' performance in the playgrounds during the time of conducting

the study, a summative test at the end of the course as well as a semi structured interview.

Vocabulary Levels Test (VLT): Nation and Waring (1997) believe that VLT can clearly demonstrate the students' correct starting point and provide a rationale for taking the most out of the learning efforts. The VLT thus prepared was administered to assess the students' knowledge of words with a view on excluding the words with which they are already familiar in the treatment phase and thus specifying appropriate materials for teaching in that phase. The word items for the VLT were selected from the list with 400 word items which was developed by virtue of submitting the assigned book, *Medicine 2* (McCarter, 2010) and its prerequisite, namely *Medicine 1* (McCarter, 2009), to a frequency analysis via the Word Frequency Counter application (<http://www.writewords.org.uk/word>). One word from every 10th words starting from the 10th item <headache> and continuing up to the 400th word item <groggy> was selected from the list. In carrying out the VLT, students were required to send through the medium of text messaging at least one Persian equivalent(s) of each word item.

English Vocabulary Recognition and Recall (EVRR) Tests: Taking into account what Chen, Hsieh, and Kinshuk (2008) proposed as a benchmark for gauging students' L2 vocabulary learning, two 40 items of multiple-choice and cloze tests were prepared by the researchers. They measured the participants' recognition and recall of the new vocabularies, respectively. They were hands-on tests used for gauging the students' performance at the end of the blended course of L2 pedagogy. A sample of EVRR is depicted in Figure 1.

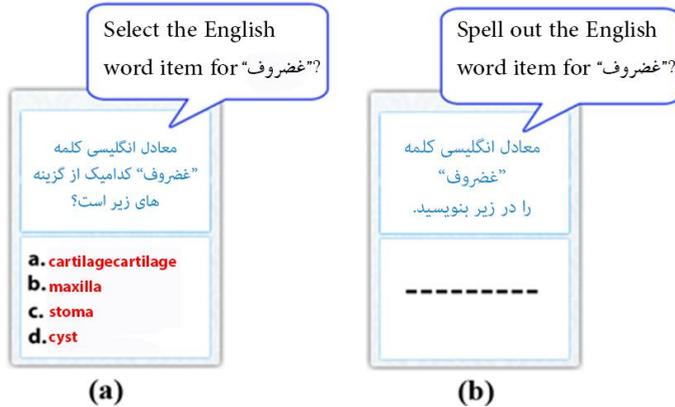


Figure 1. A sample of EVRR

The major aim was to determine if the m-gaming manner is an important agent in predicting the students' subsequent performance, namely to lend a hand to answer the first research question. Also, according to Jones (2004), these tests are usually used to check up the issue that they provide optimum conditions for gauging vocabulary learning ability both receptively and productively (Nation, 2001). EVRR tests' reliability was calculated through KR-21 method and it was .91. Their face and content validities were confirmed by five experts in TEFL.

Questionnaire (preresearch survey questionnaire): According to Hudson (1996), personal experience is a rich source of information on learning language in relation to didactic settings; that is, practicing progressed partly on students' own experience. To elicit students' preference regarding the application of m-games as an integral part of the blended learning experience and the prediction power of the formative assessment that results from practicing those m-games, namely to address the second research question and its corollaries, a text messaging questionnaire consisting 21 items was prepared and broken into two sections. To incorporate the pull mode of operation through which students' suggestion about

the suitable time and frequency of representing materials in the m-games can be surveyed (Kennedy & Levy, 2008), three multiple-choice questions at the start of the survey questionnaire were added (first section). Thus, game representation was scheduled for this time range and frequency. The second 17 question items were prepared using the 5-point Likert scales ranging from 5 (strongly agree) to 1 (strongly disagree), evaluating students' feedback and their overall attitudes toward the grafting of m-game onto the blended module of L2 vocabulary learning as well as the value of the formative assessment that is derived from different manners of m-gaming. The reliability of the questionnaire was computed .85 in a pilot study through Cronbach's alpha, which was administered to 40 students other than those who were to take part in the major study. As regards validity, content and face validity of the questions were investigated by five assessment and TEFL experts and were found to be satisfactory.

Semi-structured Interview: To effectively slip in the triangulation research design, three modified open-ended interview questions put forward by Al-Seghayer (2001) were employed to be asked from the instructors. The central focus of the items was shifted toward the possible effect of the practicing manner on the prediction power of students' formative assessment and instructors' subsequent decisions on assigning the didactic materials. The questions are like this:

Question 1: Which one of the practicing manners is best for helping you to select and develop the materials in this blended learning experiment?

Question 2: Which one of the manners of m-gaming can afford better prediction about the students' subsequent performance?

Question 3: What are the good features of m-games in this blended study that allow you, as a language instructor, to effectively implement the L2 pedagogy?

3.3 Materials

To deal with the research question of whether the practicing manner of m-games can be an agent with significant effect on the practice of formative assessment of students' performance in the m-games (viz. the first research question) some materials were selected as follows:

Software Package: Assessment means more than an objective test of how much information a student has learned in a given period of time (Gipps, 1994). For formative assessment of students' performance, their performance was recorded on the system log through the installation of the software package. The final product of designing such mobile solution that works seamlessly with students' mobile devices will be an affordable and easy to use toolkit for language providers and very accessible and engaging environment for L2 English students. This software package was developed by the third researcher of this study (registered at Information and Digital Media Development Center, Patent no. = 84883).

Instructional content: As Beaugrande (2006) asserts, the blended module of this study is the conflation of reality with hypothetical world as follows:

A. Textbook: The blended method of teaching L2 vocabulary in this study was dominated by the textbook, namely *Medicine 2* (McCarter, 2010), which is utilized extensively and determines the topics as well as the sequence of instruction and practice.

B. M-Games: The m-games were both ready-made and native games. The ready-made games, namely *Try This*, *Crocodile Board Games*, and *Sentence Monkey*, were those Havashki and Towhidi (2013) developed. However, for including such features as socialization, exchange of information, construction of knowledge, and application of knowledge practicing the Persian game of *Xane Bazi* (Rezvanfar, 2009) was considered an interactive manner of practicing, as Bowman (2001) described. *Xane Bazi* is a multi-user gaming environment.

The m-games in this study were examples of the epistemic game. An epistemic game is a unique game genre where players virtually experience the same things that professional practitioners do, as Shaffer (2007) defined. One example of an epistemic game is the Incidence of Disease. Students as individuals or in dyads attempt to find new information about the disease which is spread across the city to inform the citizens about its incidence and to take action to treat the affected people. Towards the end of the game, they reach a meaningful guideline regarding the disease and the prevention and the treatment.

Assessment of students' performance was intertwined to the fabrics of m-games. Evidence needed to assess the students' L2 vocabulary learning ability was thus provided by their interactions with the game itself or with their groupmates. In this respect, in the playgrounds, students responded to the questions as they completed the m-games and were given feedback accordingly either by their partners or through the system. It is of note that each turn of playing the m-games (both ready-made & Xane Bazi) had one score. Each m-game was comprised of eight turns and thus each had 8 scores. The applied m-games were confirmed by five experts in TEFL to be considered as learning supplementary materials in tertiary courses of medicine. The m-games for each session were 12-minutes long.

3.4 Procedures

This study lasted for 18 weeks of an academic semester. The steps set forth for conducting the study are as follows:

Step I: Introduction: The first week in the beginning of the course was assigned to students' familiarization with the blended course. In this step, students sat for VLT and thus the major corpus of new vocabulary items was specified. In other words, after students completed the VLT, it was found that the majority of them could get up to the 130th item <sensitive>. Therefore, the word items before the 150th word <negotiate> were omitted

from the major phase of learning, namely treatment (See Instruments, VLT). As a result, integrating the results of students' performance in the VLT and a topic-based curriculum culminated in the final polishing of the materials selected for carrying out the major study.

Also, to pave the way for addressing the second question of the study and before starting the treatment phase, in this step students were surveyed through the medium of the preresearch survey questionnaire (see Instruments, questionnaire). The survey was administered anonymously for all the participants and they were expected to complete the survey through the same channel of text messaging in 25 minutes.

Step II: Teaching and Practicing (Treatment): To see if, m-games generally affects the students performance or not, the blended L2 learning in which learners were able to connect both inside and outside the classroom were employed. In this manner, L2 materials were provided in face-to-face and m-game-based teaching and practicing settings in 18 sessions of the academic course as follows:

A. Classroom Practicing: The students were generally expected to accomplish the classroom assignment, after the face-to-face teaching of the new content, in fairly conventional ways. In other words, they were required to fulfill the didactic activities with their classmates through paper and pencil before they were gathered by their instructors for the purpose of formative assessment. In the classroom use of L2, not only were the materials delivered in the face-to-face mode, but the practices were done in the same mode, too.

B. Extramural Practicing (M-Game Practicing): To address the first research question, this phase was appended. Ready-made and native games were the possible scenarios of game-enhanced L2 learning in the blended module of the study, namely nonformal part of the module. Through application of variegated m-games, students were able to indicate what they

need to do best. Thus, two 8-minute m-games were sent each session to the participants' cell phones in four classes at 6 p.m. In fact, they expected to play two to three m-games between 5 p.m. to 7 p.m. and be assessed continuously after each m-gaming regarding their answers to the items of the first part of the questionnaire. Also, to offset the effect of order of materials delivery a 2×2 Latin Square (LS) design was employed (Table 2).

Table 2

The percentage of the students' consent regarding time, frequency, and manner of their assessment

Items	Given alternatives	Frequency (%)
I. How many m-games do you think should be played each session?	One	38
	Two	44
	Three	12
	More than three	6
II. At what time do you prefer the m-games to be sent on your cellphone?	In the morning	17
	In the afternoon	59
	At night	24
III. How would you like the evaluation to be done in the playgrounds?	Each session after practicing the materials	81
	At the end of the course	5
	Every five session	6
	Others	8

On the basis of the results gained from gauging the participants' original vocabulary knowledge through VLT, it was scheduled that 16 word items to be presented to participants in every blended session. In this research project, by using 2×2 Latin Square (LS) the first eight words were delivered through one of the ready-made m-games to be played by first 76 students individually, and then eight word items were delivered through

the m-game of Xane Bazi to be played collectively in dyads. At the same time, eight new items were delivered via Xane Bazi to the second 76 students to be played collectively and then eight word items were delivered through the medium of canned m-games to be played individually by these students (Fig. 2).

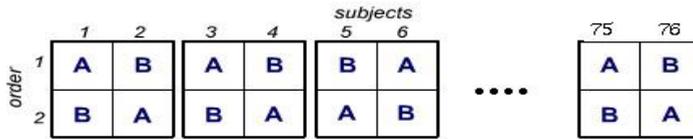


Figure 2. Two by two LS design of this study

In Xane Bazi students completed the tasks by engaging in the process of receiving immediate feedback to ascertain the accuracy of their performance. Using the 'Team Viewer' function from the software package, students were able to share their screens creations anywhere and anytime. However, in the case of individual practicing of the m-games, players individually direct the virtual environment of the games. Evidence coming from students' performance within the m-games composed raw data for formative assessment of students' performance. For the purpose of formative assessment, students' performance in the m-games continuously recorded.

Step III: Assessment: In pursuance of an answer to the first and second research questions this step was formed in two subphases as follows:

A. Terminal Test (EVRR Tests): At the end of the research period and after the 18th session, the EVRRs were administered through the application of text messaging (SMS). Students had 35 minutes to complete the tests.

B. Semi structured Interview with the Instructors: Following the course, semi structured interview was administered with the four instructors of the study. It was programmed that instructors self-report on the aspects of their experiences, interpretations, and views about the suitable

manner(s) of practicing materials through m-games, their effect on students' performance, and power of the ensued results in predicting students' subsequent performance in the blended module of L2 pedagogy in a 30-minute session. The interviews were recorded and transcribed.

4. Results

Details on the analysis of the gathered data are spelled out as follows:

4.1. Analysis of Students' Performance

The m-game as a nonclassroom part used in this study yielded an obvious improvement over the course on students' L2 vocabulary learning. Table 3 demonstrates the descriptive statistics, the mean score, and standard deviation of all the participants in the classroom test (paper-and-pencil test) and m-game-based assessment, respectively.

The mean scores obtained from the classroom tests ($m = 4.95$) differ from those resulted from assessment of students after practicing materials individually ($m = 5.82$) or collaboratively ($m = 7.43$) in the virtual playgrounds (viz. m-games), which in turn, can reveal if practicing materials through m-games contributed to students' L2 learning improvement.

Table 3

Descriptive Statistics of Students' Performance Inside and Outside the Classroom

Practicing	M	SD
Inside the classroom	4.95	.86
Individually outside the classroom	5.82	1.03
Collaboratively outside the classroom	7.43	.85
EVR score	68.82	7.55

Note. M = mean and SD = standard deviation.

To check if there are any significant differences between the scores obtained from the classroom tests and m-game assessments a paired-samples t-test was run for each manner of practicing. As shown in Table 4, statistically significant

differences were found between the mean scores of the students inside the classroom and outside it; that is, in the m-games ($p < .01$), which point to the usefulness of practicing m-games both individually and collectively.

Table 4

Paired Samples t-test Comparing the Mean Scores of Students Inside and Outside the Classroom

Differences	t	df	Sig.
Inside - Individual	-17.616	150	.000
Inside - Collaborative	-43.224	150	.000
Individual - Collaborative	-31.499	150	.000

Note. Inside = students' scores in paper-and-pencil test inside the classroom; Collaborative and individual = learners' scores resulted from collaborative and individual manners of practicing, respectively.

4.2 The Analysis of the Prediction Power Difference between the Manners of M-Gaming

To determine whether or not there are significant differences established between the EVRR scores and those scores achieved from formative assessments of students' performance the Spearman correlation coefficient was used. The Spearman correlation coefficient made a comparison between the scores of students' formative assessment scores and their EVRRs to denote which manner of practicing bore fruit of more powerful anticipation of formative assessment. As illustrated in Table 5, significant differences were identified between the mean scores of the formative assessment achieved from collaborative m-gaming and EVRR scores. In effect, in the case of the collective manner of gaming, individuals' scores on the formative assessments correlated on the scores achieved as a result of their summative assessment. That is to say, if students scored high on the playground, they achieved high score on the summative test, too.

Table 5
Correlation between Formative Scores of Different Manners of Practicing and the EVRR Scores

	r SCP- SEVRR	Sig.	r SIP- SEVRR	Sig.	%95 CI	%95 Sig. of differences
rho	.781**	<0.001	.350*	0.025	(0.351692, 0.748831)	significant

Note. r = correlation; SCP = scores achieved from collective practicing; SIP = scores achieved from individual practicing; SEVRR = scores achieved from; Sig. = significance; CI = confidence interval.

Because the sample was large enough to consider it as a normally distributed sample, a %95 Confidence Interval (CI) for the differences between the two correlations was calculated (Cohen, Cohen, West, & Aiken, 2003). Also, 0 was not included in the %95 CI; therefore, on the basis of the explicit test of group differences (Olkin & Finn, 1995), it was inferred that the difference between the two correlations was significant; that is to say, the difference between anticipatory powers as a result of collaborative and individual practicing was significant. This way, it can be claimed that prediction power of formative assessment of students' performance was highly dependent on the manner of practicing to the extent that the predicting power of the tests intertwined to the way of practicing the didactic m-games.

Cooperative virtual playgrounds (m-games) also paved the way for students to showcase their L2 vocabulary learning talent. Appealing for or offering help in the process of collaboration for fulfilling the gaming task correlates with mastering of the materials and the prediction power of the m-game tasks. Feedback in the collaborative m-gaming provided such that it presented effective information in guiding students to solve their learning problem. As a result, the consequence of the decision students reached at in the collective m-game was deeper learning of new didactic materials (Table 6).

Table 6
Association Between the Number of Offering or Requesting for Feedback and the Manner of M-Gaming

Collaborative M-Gaming			
Offering/Requesting Feedback	Yes	No	Total
Yes	6.2	3.1	9.3
No	2.2	7.6	9.8
Total	8.4	10.7	19.1

Individual m-gaming imposed constraints on the inference of the students' performance. Therefore, the frequency of students' giving up in the case of ready-made games which were generally played individually was far more than the collective m-games (Table 7). Instead, in the contexts (viz. collective practicing contexts) where sharing of information within the m-games was enough that rendering assistance was possible, the odds were in favor of the students who practiced in those contexts and the prediction arisen from the formative assessment of the students' performance in those contexts.

Table 7
Association Between the Number of Giving up and the Manner of M-Gaming

Individual M-Gaming			
Giving up	Yes	No	Total
Yes	4	2	6
No	5	7	12
Total	9	9	18

Although playing both types of m-games showed an ascending pattern of L2 vocabulary learning, it was not easily applicable in the case of the results of learning word items played in the ready-made m-game. In other words, the fluctuating patterns of prediction ensued from the EVRR tests of the individually practiced words unfolded that a comprehensive design document describing predicted students' experiences never came up with. Interestingly, collectively played m-games

contents met the intentions; that is, nearly all the anticipation appeared true. In general, as the level of collaboration among students augmented, the accuracy of assessment also increased. What is of interest to be achieved from the m-game-enhanced learning very likely was unfolded under the conditions of the collaborative environment.

4.3 Analysis of Students' Attitudes towards M-Games

To answer the second question or to see if the things that students say in relation to the m-games and manner of practicing the m-games have any significant relationship with their performance and the predictive power of the m-games, their answers to the Likert-type questionnaire (viz. the second part of the prereserach survey questionnaire) were analyzed.

As the data indicate, although students agreed that pedagogically offered m-games were useful in facilitating L2 learning process, a great majority of them (81%) voted for the collective manner of practicing m-games. Fifty eight percent of the students were not into individual m-gaming. The students' rankings of practicing interest were in line with their test scores, that students liked the team working (viz. collective practicing) the best.

More than 83 percent of students believed that application of new contents in the engaging contexts of gaming prevents them from telegraphing their outcome, the term Shute and Ke (2012) applied, at the start of the game. They added the attractive edge of the m-game accelerates learning by further exploration in the playground.

Students asserted collective practicing in the native m-games allows them to consider their performance from multiple perspectives. Naturally, collective learning offers opportunities of reflection and discussion to the students. Students reported that collective manner of gaming encourages them to influence over gameplay; thus, they stated that this feature has visible effects on their L2 learning ability. They were of the opinion

that their collective effort for completing the m-games to reach a meaningful plot paves the way for deep processing of the materials. They claimed at the end of each turn of collective games they can take away a deep understanding of new materials. They added that in fact, completing the tasks through interaction during nonformal environment is a protective environment where they can feel free to express and evaluate themselves which in turn encourages them to make sense of L2 for themselves. Students claimed that nearly all of the learning problems can be addressed by virtue of their interaction with their partners while they are completing the gaming tasks. Students believed learning L2 through collaboration with confederates enhances feelings of self-confidence and self-esteem, because they have more experience with the materials and this, in turn, will increase their level of comfort dealing with new topics of L2. In this fashion, they said Xane Bazi does a better job of predicting college students' vocabulary learning ability. Seeing that students believed that cooperation supports inferences in the pedagogical situations to predict what they do next in other situations, students asserted, practicing new materials through native m-games gives the right evidence (see Appendix). The correlation between attitude towards collective gaming and its formative assessment achieved a correlation ($\rho = .804$) significance at .01 level ($p < .01$) (Table 8).

On the other hand, in the case of the individual manner of playing didactic m-games it is unclear how one can understand the new instructional materials. Students stated that in the individual practicing they make sense of the situation only through resorting to their a priori knowledge. They believed that individual manner of playing the didactic m-games makes the game much more difficult and thus all the more competitive, in the same way.

Overall, on the basis of the students' responses to the items of the preresearch survey questionnaire while there seemed to be

promise in relation to the application of m-games in the blended module of L2 vocabulary learning, there were low-points with some types of m-games and manner of m-gaming, namely ready-made games which were to be played individually rather than collectively.

As Table 8 shows, the correlation coefficients between students' attitudes towards application of m-games and the collective m-gaming with their performance were significant. In fact, the correlation between attitudes and m-game application as well as its manner of practicing achieved significance at .01 levels:

- Attitude to m-game — rho = 0.574, $p < .01$; and
- Attitude to collective m-gaming — rho = 0.771, $p < .01$.

Table 8

The Correlation between Students' Attitudes Towards M-Game, Manner of Practicing, and Affordability of the Resulted Formative Assessment and their Performance

Variables	rho	p
r between attitude towards m-game and performance	0.574**	0.009
r between attitude towards collaborative m-gaming and performance	0.771**	0.001
r between derived formative assessment that arisen from collaborative m-gaming and subsequent performance	0.804**	0.000

Note. r = correlation.

** $p < .01$

4.4 Analyzing Instructors' Interviews

The first instructor reported that "collective m-gaming had the affordability to accurately assess the actual learning that took place during the course; thus, given that I was better able to tune subsequent materials to optimize the students' performance".

The instructor of the second classroom believed that "social features should be embedded in the process of practicing the content to better support the teachers and students. In this vein, closely aligning game-related activities toward the cooperation

of students in the instructional games allows for appropriate learning by students as they are completing the tasks embedded within the games". Moreover, she said that "collaborative practicing of materials paved the way for collecting suitable data on student learning from mobile playgrounds or m-games. From her perspectives, because playing of ready-made m-games had little in common with the practices targeted at satisfying real-life needs, assessment derived from practicing these m-games could not show a comprehensive prospect of students' performance".

The third instructor stated that "actually in individual practicing of instructional content via m-games students were at a disadvantage trying to understand and express themselves in L2. He believed that there were clear hints of the ambiguity and confusion regarding the nature of the individual manner of practicing m-games. He concluded that "thus, teachers can gain better awareness of students' beliefs and practices when students are practicing in collaboration with each other".

The fourth instructor opined that "functionality in the commercially available m-games was limited to player; however, in the Persian game of Xane Bazi, functionality included player-to-player interactions, team interactions, and control of game and player features. In other words, as a team, students derived inspiration from a number of sources or people (viz. their group mates or confederates). He claimed that "I was able to plan more effectively and design better didactic content when I planned on the basis of the assessment which resulted from collaborative practicing scenario. In reality, students' cooperation in the playground enables educators and materials developers to expand their repertoire of strategic options and hence become more flexible practitioners".

5. Discussion and Conclusions

As the results throughout this blended study showed, collaborative m-gaming appeared to support L2 vocabulary learning. In regard to the manner of practicing and m-game type,

collective playing with new materials in the nonclassroom context of m-games (viz. extramural learning context) set up welcome opportunities for L2 students for knowledge retrieval immediately as they were able to practice the newly acquired items in correspondence with their confederates. In a similar vein, didactic Xane Bazi portrayed students' growth in L2 vocabulary learning in a more comprehensive framework. Hence, through utilizing the collaborative manner of practicing, educators and materials developers can provide practical teaching ideas in which L2 teachers can integrate indigenous artifacts and L2 teaching in a nonformal mode of teaching and assessing learners. By and large, the results thus obtained seem to stand as testimony to the claims that the collective manner of playing didactic m-games guaranteed students' outperformance. This implies that students can be more likely to succeed in learning through the two-way or collective model which is comprised of students' collaboration in comparison with the learning that resulted from interactions between students and the game, namely individual manner of practicing (Kiili, 2007). Therefore, as Salen and Zimmerman (2005) argue, the interpretation of knowledge and skills as the products of learning cannot be isolated from the context, and neither should assessment (as cited in Shute & Ke, 2012, p. 52).

Another important finding in the present research study is that the SJTs can take advantage of the collaborative manner of practicing m-games (collective vs. individual practicing). The representativeness of assessment in the collaborative didactic playground can be attributable to students' deep learning as a result of collective learning throughout the process. So, it can be claimed that the accurate nature of assessment is also a social process. Social consequences refer to the appropriateness of decisions made on the basis of the students' collaboration toward completing the gaming tasks. On the other hand, individual m-gaming underrepresented students' L2 learning ability. Indeed,

paying insufficient attention to the relationship between a priori knowledge and new didactic materials precludes deep learning which results in underrepresentation of the achieved scores in predicting students' subsequent performance. Therefore, the possibilities of taking advantage of individual ready-made m-games are much lower when compared to that of collective native m-games.

Although m-games make ways for educators to continually assess the performances of their students, the prediction power of them somewhat varies from the manner of practicing m-games with high prediction power to the ones with the lower. In the absence of a correct connection between didactic contents and learned knowledge, namely a priori knowledge, assessment of students' performance has an ad hoc quality, the results of which tending not to be very informative. The estimated values of the performance on the individual games were not matched with their performance on the end-of-the-semester (viz. terminal tests) tests. Thus, an assessment of students' performance is tied to their actions within virtual playgrounds (viz. m-games) (Shute, 2011).

This study presented another general pattern emerging in the obtained data. The findings unraveled that students adopted favorable attitudes toward application of collective m-gaming for the purpose of L2 vocabulary learning. Due to convergence of students' friendly attitudes towards efficacy of the collectivity in practicing L2 didactic materials and assessment and their outperformance as a result of practicing the didactic contents in the collaborative m-games; consequently, students' attitudes towards m-games types can be attributable to their experience with those m-games at the time of practicing materials. As regards the familiarity of the context and students' experience, Slator and Associates (2006) assert that "it seems necessary that educational technology capitalize on the natural human propensity for cooperative learning" (p. 4).

This study demonstrated that the collective fulfilling of native m-games helps us use the familiar context of practicing for the purpose of assessments more effectively. Accordingly, the gleaned information from the collective virtual playgrounds (viz. m-games) can be better used for adjusting the subsequent support for students. Indeed, embedding social properties in the assessment settings is considered of great consequence for eliciting the skills and subskills the educators care about. In the view of that, the same idea of assessment through m-games can be applied to almost any L2 English skill or subskill being taught at Iranian medical school or at any tertiary institution. One of the most effective ways of doing this can be accomplished with collective fulfilling of the gaming tasks.

In short, the manner of m-gaming was considered as a moderator concerning the impact of m-games on learners' performance and the inference derived from it. As a result, the real challenge for educators is to be acquainted with new pedagogical tools and be taught how to use them in their everyday activities (Czepielewski, Christodoloplu, Kleiner, Mirinaviciute, & Valencia, 2011).

However, it is important not to be limited to collaboration as a feasible feature for powerful prediction. In other words, cooperative playing as the only feature which can result in precious features is akin to transforming the collaborative act into a learning objective.

References

- Al-Seghayer, K. (2001). The effect of multimedia annotation modes on L2 vocabulary acquisition: A comparative study. *Language Learning & Technology*, 5(1), 202-232.
- Bajpai, S. (2013). Impact of game-based learning on education. Retrieved from <http://www.edtechreview.in/index.php/news/trends-insight/399-game-based-learning-impacting-education>
- Banados, E. (2006). A blended-learning pedagogical model for teaching and learning EFL successfully through an online

- interactive multimedia environment. *CALICO Journal*, 23(3), 533-550.
- Beaugrande, R. (2006). Critical discourse analysis: History, ideology, and methodology. *Studies in Language and Capitalism*, 1, 29-56.
- Belland, B. R. (2012). The role of construct definition in the creation of formative assessments in game-based Learning. In D. Ifenthaler, Eseryel, D., & X. Ge (Eds.), *Assessment in game-based learning* (pp. 29-42). London: Springer.
- Bowman, L. (2001, November). Interaction in the online classroom. *The teachers' net Gazette*, 2(7). ERIC Document Reproduction Service No. ED464319. Retrieved from <http://www.eric.ed.gov/>
- Chao, K.-H., & Chao, Z.-Y. (2011). The construction of text-based and game-based teacher career aptitude tests and validity comparisons. In M., Chang, W.-Y. Hwang, M.-P., Chen, & W. Müller (Eds.), *Edutainment technologies* (pp. 527-532). London: Springer.
- Chapelle, C. A. (2003). *English Language Learning and Technology*. Amsterdam/Philadelphia: John Benjamins.
- Chen, N. S., Hsieh, S. W., & Kinshuk (2008). The effects of short-term memory and content representation type on mobile language learning. *Journal of Learning and Technology*, 12, 93-113.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple/regression correlation analysis for the behavioral sciences*. Hillsdale, NJ: Erlbaum.
- Czpielewski, S. (2011). The virtual world of second life in foreign language learning. In S. Czpielewski (Ed.), *Learning a language in virtual worlds: A Review of Innovation and ICT in Language Teaching Methodology* (pp. 15-24). Warsaw, Warsaw Academy of Computer Science, Management and Administration.
- Czpielewski, S., Christodoloplou, C., Kleiner, J., Mirinaviciute, W., & Valencia, E. (2011). Virtual 3D tools in online language learning. In S. Czpielewski (Ed.), *Learning a Language in Virtual Worlds: A Review of Innovation and ICT in Language Teaching Methodology* (pp. 7-15). Warsaw, Warsaw Academy of Computer Science, Management and Administration.

- Ellis, R. (2000) Task-based research and language pedagogy. *Language Teaching Research*, 4(3), 193–220.
- Gamlin, M. (1995, September). *Distance learning in transition: The Impact of Technology: A New Zealand perspective*. Paper presented at the 1995 EDEN Conference on the Open Classroom, Distance Learning, and New Technologies in School Level Education and Training, Oslo, Norway.
- Gee, J. P. (2003). *What video games have to teach us about learning and literacy*. New York: Palgrave/Macmillan.
- Gipps, C. V. (1994). *Beyond testing: Towards a theory of educational assessment*. London: The Falmer Press.
- Havashki, H. & Towhidi, H. (2013). English Games (version 1) [software]. Tehran: Khate Sefid Press.
- Huizinga, J. (1987). *Homo Ludens. Vom Ursprung der Kultur Im Spiel*. [A study of a play element in culture]. Boston: Beacon Press.
- Jantke, K. P. (2012). Patterns of game playing behavior as indicators of mastery. In D. Ifenthaler, Eseryel, D., & X. Ge (Eds.), *Assessment in game-based learning* (pp. 85-103). London: Springer.
- Johnson, P. (2013). *Mobile tech offers authentic learning potential*. Retrieved from <http://eschoolnews.com/2013/09/18/mobile-learning-potential-120>
- Jones, L., (2004). Testing L2 vocabulary recognition and recall. *Learning and Technology*, 8(3), 122-143.
- Kennedy, C., & Levy, M. (2008). Using SMS to support beginners' language learning. *Recall*, 20 (3), 315-330.
- Kiili, K. (2007). Foundation for problem-based gaming. *British Journal of Educational Technology*, 38, 394-404.
- Maue, T. P. (2012). Writing as a tool for thinking: An interactive exercise. In J. A. Chambers (Ed.), *Selected Papers from the 22nd International Conference on College Teaching and Learning* (pp. 119-136). Florida: Center for the Advancement of Teaching and Learning.
- McCarter, S. (2009). *Medicine 1*. Oxford: Oxford University Press.
- McCarter, S. (2010). *Medicine 2*. Oxford: Oxford University Press.
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Champion, M. A., & Braverman, E. P. (2001). Predicting job performance using

- situational judgment tests: A clarification of the literature. *Journal of Applied Psychology*, 86(4), 730-740.
- Mehrotra, D. (2013). Quality education: The ultimate desire by parents. Retrieved from <http://www.edtechreview.in/news/trends-insights>
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23 (2), 13–23.
- Murphy, E. (1988). The cultural dimension in foreign language teaching: Four models. *Language, Culture, and Curriculum*, 1(2), 147-163.
- Nation, I. (2001). *Learning vocabulary in another language*. Cambridge: Cambridge University Press.
- Nation, P., & Waring, R. (1997). Vocabulary size, text coverage and word lists. Retrieved from <http://fltr.ucl.ac.be/fltr/GERM/ETAN/bibs/vocab/cup.html>
- National Survey of Student Engagement (2010). Retrieved from http://nsse.iub.edu/html/toolsand_services.cfm.
- Olkin, L., & Finn, J. D. (1995). Correlations redux. *Psychological Bulletin*, 118, 155-164.
- Paige, R. M., Jorstad, J., Siaya, L., Klein, F., & Colby, J. (2003). Culture learning in language education: A review of the literature. In D. Lange, & R. M. Paige (Eds.), *Culture as the Core: Integrating Culture into the language education* (pp. 173-236). Greenwich, CT: Information Age.
- Pascall, W. (2013). E-learning: where do serious games make most sense? Retrieved from <http://www.slideshare.net/InteractiveLearning/webinar-driving-elearning-engagement-and-interactivity-using-raptivity-tips-and-techniques>
- Rezvanfar, M. (2011). *Iranian native games*. Tehran: Nashre BoomAbad.
- Rupp, A. A., Gushta, M., Mislevy, R. J., & Shaffer, D. W. (2010). Evidence-centered design of epistemic games: Measurement principles for complex learning environments. *Journal of Technology, Learning, and Assessment*, 8(4), 4–47.
- Salen, K., & Zimmerman, E. (2005). Game design and meaningful play. In J. Raessens & J. Goldstein (Eds.), *Handbook of computer game studies* (pp. 59–80). Cambridge, MA: MIT Press.

- Schrader, P. G., & McCreery, M. (2012). Are all games the same? In D. Ifenthaler, Eseryel, D., & X. Ge (Eds.), *Assessment in game-based learning* (pp. 11-28). London: Springer.
- Shaffer, D. W. (2007). *How computer games help children learn*. New York: Palgrave.
- Shih, Y.-H., Hou, H.-T., & Wu, Y.-T. (2011). A review on the concepts and instructional methods of mni digital physics games of PHYSICSGAMES.NET. In M., Chang, W.-Y. Hwang, M.-P., Chen, & W. Müller (Eds.), *Edutainment technologies* (pp. 517-521). London: Springer.
- Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. In S. Tobias & J. D. Fletcher (Eds.), *Computer games and instruction* (pp. 503–524). Charlotte, NC: Information Age.
- Shute, V. J., & Ke, F. (2012). Games, learning, and assessment. In D. Ifenthaler, Eseryel, D., & X. Ge (Eds.), *Assessment in game-based learning* (pp. 43-58). London: Springer.
- Slator, B. M. & Associates (2006). *Electric worlds in the classroom*. New York and London: Teachers College.
- Squire, K. (2003). Changing the game: What happens when video games enter the classroom? *Innovate*, 1 (6) [online report]. Retrieved from <http://www.innovateonline.info/index.php?view=article&id=82>.
- Sundre, D. L., & Wise, S. L. (2003). *Motivation filtering: An exploration of the impact of low examinee motivation on the psychometric quality of tests*. Paper presented at the annual meeting of the National Council on Measurement in Education, Chicago.
- Tan, E. C. C. (2011). Encouraging student learning through online e-portfolio development. In R. Kwan, C. McNaught, P. Tsang, F. LeeWang, & K. C. Li (Eds.), *Enhancing learning through technology* (pp. 9-21). London: Springer.
- Vygotsky, L. S. (1962). *Thought and language*. Cambridge, MA: MIT Press.
- Wieczorek, J. A. (1994). The concept of French in foreign language texts. *Foreign Language Annals*, 27(4), 487-97.
- Wiliam, D., & Black, P. (1996). Meanings and consequences: A basis for distinguishing formative and summative functions of

assessment? *British Educational Research Journal*, 22 (5), 537–548.

Zhu, Ch., Meng, W., Wang, Y., & Zhang, X. (2011). Cage-based tree deformation. In M., Chang, W.-Y. Hwang, M.-P., Chen, & W. Müller (Eds.), *Edutainment technologies* (pp. 409-413). London: Springer.

Appendix: The Mean Scores and Standard Deviation of the 17 Items in the

Attitude towards	Items (n= 17)	Description of the items	Mean	SD	t-test
MG	1	M-games can be helpful to students' learning.	3.31	0.770	4.89
	2	M-games can motivate students to learn L2 English.	3.93	0.819	14.05
	3	M-games can help students to improve their L2 vocabulary learning.	3.65	0.848	9.47
	4	M-games pave the way for learning anytime and anywhere.	4.14	0.737	19.03
	5	M-games can cover all the L2 learning problems.	3.11	0.773	1.67
	6	Educators should develop variegated m-game tasks for teaching L2.	3.19	0.886	2.60
	7	Practicing m-games makes way for deep learning	3.45	0.779	6.94
	8	M-games help students develop a sense of participation.	4.29	0.752	20.70
	9	Collective practicing of m-games increase the interaction between members of the groups.	4.06	0.939	13.78
	10	Collaboration for fulfilling m-games enhance students feelings of self-confidence	3.97	.918	12.84
CP	11	Collaborative practicing of m-games leads into to meaningful learning	3.97	0.945	12.61
	12	Collaborative gaming reduces sense of giving up	3.68	0.857	9.84
	13	Collective practicing allows students to consider their performance from multiple perspectives	3.38	0.718	6.55
	14	Collaboration in m-gaming lets students' many further in-depth explorations	3.81	0.968	10.23
FA	15	Collective m-game practicing is fair to assess students' performance.	3.99	0.806	15.10
	16	M-games gives right evidence	3.52	0.956	6.70
	17	The predefined time intervals for playing m-games is favorable	2.6	0.859	-5.70

Note. Scoring: 5 = strongly agree, 1 = strongly disagree; MG = M-Game; CP = Collective practicing; FA: Formative assessment derived from collective gaming.