

Article



# Assessing EFL Students' Performance and Self-Efficacy Using a Game-Based Learning Approach

Maram Meccawy <sup>1,\*</sup>, Asma Alzahrani <sup>1</sup>, Zinah Mattar <sup>1</sup>, Rawan Almohammadi <sup>1</sup>, Samah Alzahrani <sup>1</sup>, Ghadeer Aljizani <sup>1</sup>, and Zilal Meccawy <sup>2</sup>

- <sup>1</sup> Information Systems Department, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia; aattyahalzahrani@stu.kau.edu.sa (A.A.); zkhalilmatar@stu.kau.edu.sa (Z.M.); rghalialmohammadi@stu.kau.edu.sa (R.A.);
- smohammedalzahrani0009@stu.kau.edu.sa (S.A.); grahbanaljizani@stu.kau.edu.sa (G.A.)
- <sup>2</sup> English Language Institute, King Abdulaziz University, Jeddah 21589, Saudi Arabia; zmeccawy@kau.edu.sa
- Correspondence: mmeccawy@kau.edu.sa

Abstract: Educational technology plays a major role in today's learning as it offers significant advantages in delivering learning content, classroom communications, and assessing learners. Assessment is a cornerstone in modern and formal education and is particularly necessary for tracking progress, measuring knowledge or achievement, and planning future educational steps. This study investigates the effects of using an online game-based learning (GBL) approach in students' assessment in terms of their performance and self-efficacy in English as a foreign language (EFL) learning. A quasi-experiment was conducted on intermediate school-level students, who were divided into two groups, namely control and experimental. Each group underwent an English language test conducted through a Google Form and a role-playing game developed to mimic the functionality of an assessment tool. The results indicate that the performance of the experiment group was neither affected positively nor negatively using the GBL assessment tool. On the other hand, the results also show that using such a tool has positively affected students' self-efficacy levels.

check for updates

Citation: Meccawy, M.; Alzahrani, A.; Mattar, Z.; Almohammadi, R.; Alzahrani, S.; Aljizani, G.; Meccawy, Z. Assessing EFL Students' Performance and Self-Efficacy Using a Game-Based Learning Approach. *Educ. Sci.* 2023, *13*, 1228. https:// doi.org/10.3390/educsci13121228

Academic Editor: Lawrence Jun Zhang

Received: 15 November 2023 Revised: 4 December 2023 Accepted: 8 December 2023 Published: 11 December 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). **Keywords:** English language learning; assessment; educational technologies; game-based learning GBL; performance; self-efficacy; schools

# 1. Introduction

Assessment plays a crucial role in various aspects of education. It is vital for understanding learning progress, improving instruction, guiding individuals, and ensuring accountability in various educational contexts. Assessment is defined as the method(s) used to analyze and measure knowledge possessed by the education receiver [1], while the process of examination itself is defined as a single time-restricted focus point that occurs distinctively on one occasion [2]. The blurred distinction between both terms has progressively evolved to a more unified understanding, where examination has been upgraded to include multi-dimensional factors driven by technological advancements [3]. Black and William [4] argue that to raise standards, teachers need to gather information through assessments and use it to modify or change teaching and learning tasks to meet their learners' needs [5]. However, many educators approach assessment in the traditional written-text method closely related to the nature of "testing" with a unidimensional focus. This might be efficient in a traditional testing environment; however, when applied in online distance learning, this dominantly used approach, within non-higher education levels, introduces the risk of missing factors that school-based learning bypasses, such as the lack of in-person individual guidance and instruction within an immersive test environment [6].

As a result of the global COVID-19 lockdown, schools and educational institutions had to switch to online learning overnight [7], which also included online assessment. Online assessment is a method of evaluating knowledge, skills, and abilities through the use of

digital tools such as tests, quizzes, and surveys. It can be used in a variety of educational and professional settings to measure a wide range of competencies. One of the primary advantages of online assessment is its flexibility and convenience. Online assessments can be administered at any time and from any location with an internet connection, allowing for greater accessibility for learners. Moreover, online assessments can be easily automated and graded, reducing the workload for instructors, and providing immediate feedback to learners. However, online assessments also have some potential drawbacks. One challenge is ensuring the security and validity of online assessments, as test-takers may be susceptible to cheating by using outside resources during the assessment [8].

Online assessments can take many forms, including multiple-choice questions, essay questions, simulations, and interactive activities. Some online assessment tools also use games or incorporate elements of gamification, such as rewards and badges, to promote engagement and motivation among learners.

Although online assessment became widely used during lock-down, it was available prior to that period and continues to offer its advantages as educators pursue better assessment alternatives. Efforts to investigate alternative assessments and environments should be maintained to offer the efficiency rate required for proper knowledge retention measurement [9]. Therefore, this research investigates an assessment approach using online games.

The use of online games has grown beyond entertainment or play; of the many contexts employing games, learning has been one field dominated by their implementation [10,11]. Three concepts have been cultivated in this aspect: gamification, game-based learning (GBL) [12], and serious games [13,14]. Playing games as part of the teaching and learning experience is believed to increase students' performance and motivation as well as encourage social and emotional development. These three concepts of educational games and their impact on learning will be explained in further detail in the context of digital learning in Section 2.

GBL has been found to have a positive impact on both students' learning. Research studies have shown that implementing game-based learning can increase students' self-efficacy, learning motivation, and learning performance [15]. Using digital games as a learning tool engages students in a stimulating and interactive environment. This immersive experience can enhance their motivation and interest in the subject matter leading in turn to improved learning outcomes [15]. When students feel more confident in their abilities and believe in their capacity to succeed, their self-efficacy increases. Game-based learning provides opportunities for students to experience a sense of achievement through overcoming challenges and progressing through levels, reinforcing their belief in their own abilities.

This study examines the possibility of introducing a GBL assessment environment as an alternative assessment approach. It aims to measure its effectiveness by asking two research questions:

- 1. Does using GBL as an assessment tool affect EFL students' performance?
- 2. Does using GBL as an assessment tool affect EFL students' self-efficacy?

Based on these questions, the first hypothesis (H1) posits that GBL as an assessment tool positively affects students' test performance, whereas the second hypothesis (H2) posits that GBL as an assessment tool positively affects students' self-efficacy.

A quasi-experiment was conducted on intermediate school-level students (12–14 years old), who were divided into two groups: control and experimental. Each group underwent an English language test conducted through a Google Form and a role-playing game developed to mimic the functionality of an assessment tool.

This paper proceeds as follows: the background and related work are presented in Section 2, followed by the study materials and methods in Section 3, the results in Section 4, the discussion in Section 5, the limitations alongside the future work in Section 6, and finally the conclusion in Section 7.

## 2. Background and Related Work

## 2.1. Assessment

Herman [1] defines assessment as a methodology applied to identify, measure, and quantify the level at which a learner has or can retain information, and Fulcher [5] further divides assessment into large-scale standardized testing (summative assessment) and classroom testing (formative assessment), with each having associated technologies, methodologies, and challenges. Traditionally, summative assessment aims to measure the actual outcome of the learning experience or how much is retained and applied after the learning experience. The students' level of achievement is measured using different evaluation techniques that primarily focus on students' academic abilities in the retention of the learned subject and memorization. For example, formal standardized quizzes, assignments, and exams consisting of true/false questions, multiple-choice questions, essays, and short-answer tests are used to evaluate and measure students' performance and the extent of their understanding [16]. On the other hand, formative assessment is defined as "encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged" [4]. Formative assessments are used to assist the ongoing learning journey with the objective of modifying teaching to achieve the desired outcome, being seen as an ongoing evaluation of knowledge, and providing continuous feedback to students and teachers, in which teachers establish educational goals or objectives, construct tasks that move learners towards these goals, and evaluate to what extent they have succeeded in doing so [5]. Both assessment types work dynamically to complement one another.

However, when assessing individual learners and their needs and abilities, teachers may consider alternative assessment methods [17], such as role play, posters, focus groups, and online assessments of all kinds. Some advantages of alternative assessments include improving students' problem-solving and decision-making skills, enhancing interactivity and engagement, and creating a more productive learning environment. Furthermore, alternative assessment measures students' decision-making and problem-solving skills, evaluates their performance, and discovers their skills in a more interactive, realistic, and feasible way to acquire information [3]. This approach also helps the teacher create a productive learning environment that meets each student's needs and skills.

## 2.2. The Use of Games in Education

Game employment in digital education over recent decades may be divided into classifications that are distinct in definition but overlap in implementation. These categories are gamification, GBL, and serious games [18]. Gamification is the application of game elements to non-game activities, whereas GBL provides an actual game as part of the educational process. On the other hand, serious games directly incorporate educational content into games without the entertainment element.

Gamification in a learning context means using game elements such as leader boards, achievements, points, badges, and levels to engage learners in various learning activities [19] and is used in teaching a variety of subjects, such as English [20] and Computer Programming [21]. Serious games are designed to improve learning outcomes without any entertainment or fun [22], "while gamification can be an effective way to educate, promote, and so forth by using various aspects of video games, it does not provide the solutions that serious games do" [14]. Both game types differ from GBL, which is an educational environment that utilizes the game experience as one of the learning techniques, including both entertainment and educational elements [23]. Researchers have investigated GBL in a broad range of subject areas, such as science [24,25], health [26], history [27], English language listening skills [28], English language writing [29], mathematics [30,31], and programming [32,33].

GBL has interesting storylines, clear goals, problem-solving, and challenges, providing students with a sense of achievement and quantifiable outcomes, inherently reflecting on the education environment to be more diverse and effectively increasing students' learning

interest and effectiveness [34]. GBL has shown the ability to motivate learners and positively impact their attitudes [35–38], increasing their engagement level for an extended period of time through a series of motivational features that create a high situational interest, including incentive structures [39–42]. Moreover, previous studies have used digital GBL to improve students' computational thinking skills, problem-solving strategies, and abstract thinking [25,32]. Additionally, socialization, communication skills, and emotional intelligence are positively affected [43].

To date, a considerable amount of literature and extensive research has been published on gamification, serious games, and GBL for different learning objectives [44–46]. Moreover, several research efforts have examined the use of games in education in terms of assessment. For example, one study used formative assessment-based gaming as a new approach to examining student behavior in the context of digital citizenship behaviors, with results showing that the proposed approach enhanced students' behavioral decisions with significant scores and promoted students' motivation and perceptions [47]. Another study recognized the potential benefit of using 3D educational computer games as an alternative method to traditional examination to reduce test anxiety and significantly improve students' performance [48].

## 2.3. The Use of Games in Language Education

A specifically designed application for learning French as a foreign language in a Norwegian secondary school was developed and used to measure student engagement and task control, revealing that not only were students engaged in the activities, but they were completing other language-related tasks of their own accord [49]. In another study on exploiting commercial adventure videogames for improving second-language vocabulary recall by designing Supplementary Material, an experiment was conducted on two student groups, finding that students who played the videogame outperformed the other group significantly in post-tests and had generally positive views towards using the game in language learning in general and vocabulary learning specifically [50]. Similarly, another multi-data study examined Hong Kong University students' perceptions of gamified flipped classrooms, in which game tasks were more effective than classroom discussions, contributed to a positive classroom atmosphere, helped review materials before exams, and reduced anxiety and hesitance surrounding using English [51]. All three studies suggest pedagogical implications for language education. Moreover, Bagunaid et al. [52] investigated students' self-assessment using a specifically developed gamified tool to help Saudi EFL students assess their own progress through social comparison with what other students achieved using the interactive visual representation of their language learning progress, which was positively received. Moreover, Socrative, a popular web-based online student response system (OSRS), was used to explore Saudi female EFL learners' attitudes towards mobile-based tests in English classes, demonstrating a preference for paper-based tests in the pre-experiment survey; however, after the experiment, students changed their preferences in favor of mobile-based tests [53]. A study conducted to determine the effectiveness of explicitly learning English vocabulary through a Charades game found that the game helped increase students' scores and enhance their academic performance [54].

# 2.4. Self-Efficacy

Self-efficacy is defined in education as a student's belief or expectations in successfully completing tasks or achieving specific objectives [55]. For example, a study found that different motivation and engagement techniques influenced students' self-efficacy in the learning environment and improved learning outcomes [56]. According to a survey conducted by Zheng et al. [57] following an experiment that used a digital game environment for junior high school students, the experimental group showed higher self-efficacy, confidence, and comfort using English compared to the control group, who only used traditional learning methods. Another study looking into digital game-based learning found an enhancement in students' self-efficacy when using badge mechanisms, including leaderboard ranking and learning practice with star icons as feedback in an English learning environment for third-grade elementary school students [58]. Another study used the social cognitive theory to measure school students' self-efficacy toward using GBL in three perception traits and using the Formosa Hope game, the results indicated that students' positive perceptions of learning and playing are fundamental in promoting self-efficacy in GBL [59]. Existing research has recognized the critical role played by students' self-efficacy as a factor in accepting new education technologies, particularly games, in the context of learning and students' performance.

The studies presented in the literature demonstrate a range of advantages of using games in educational fields. However, with the promising results in recent experimental research using GBL as an assessment method, further research is required to illustrate its effectiveness as a reliable alternative assessment to examine student's level of achievement in different subjects as well as its impact on students' self-efficacy. Therefore, this study investigates the influence of using GBL, as an alternative assessment, on students' performance in learning English as a foreign language, and addresses whether GBL affects students' self-efficacy while undergoing English exams.

The two hypotheses that were influenced by the literature and tested in this study are as follows: (H1), which hypothesizes that GBL as an assessment tool positively affects students' test performance, whereas the second hypothesis (H2) predicts that GBL as an assessment tool positively affects students' self-efficacy. Alternatively, the null hypothesis [60], which is a statistical test that proposes that no statistical significance exists in a set of given observations, will be accepted.

# 3. Materials and Methods

# 3.1. Research Design

This research falls into a quasi-experimental design [61] as it measures differences in English exam performance and students' self-efficacy between two groups after using a GBL assessment tool. The experiment used two instruments, as will be described in the following sections.

# 3.2. Instruments

The first instrument is an English language test designed using Google Forms by an English teacher with 15 years of experience teaching intermediate school students and was reviewed by three experts who majored in English. It contained 24 multiple-choice questions in five categories (vocabulary = 6, grammar = 4, reading = 2, pronunciation = 6, listening = 6).

The second instrument was designed to measure the students' self-efficacy after using the GBL as an assessment tool and is a five-point Likert-scale questionnaire with 16 items divided into two dimensions. The first eight items represented the dependent dimensions of students' self-efficacy, and the later eight items represented the independent dimensions of GBL as an assessment tool. The instrument questions were adapted from [58,62] and slightly modified to suit this study's purpose. Only two of the questions were adopted from [62] to evaluate participants' self-efficacy toward English language examinations, and the remaining questions were adopted from [58] and used to measure the effect of features such as feedback and rewards of the game on students' self-efficacy. Nine questions were modified to expressly reflect features implemented in the GBL assessment tool, such as achieving higher levels, collecting rewards, and competing with the English masters, enemies, and main characters. Furthermore, five of the questions developed from one instrument to measure the effects of each enemy were created to challenge the student's English skills, as represented by Vocabulary Master, Listening Master, Reading Master, Grammar Master, and Pronunciation Master. The questionnaire items' rating scale was from 5 for "Strongly agree" to 1 for "Strongly disagree". The Cronbach's coefficient of the study questionnaire is 0.89, representing consistent reliability, see Table 1 for variables (self-efficacy and GBL).

First Dimension	Self-Efficacy	Mean	Std. Deviation
SE1.	I can answer the questions on the provided English language test.	4.67	0.516
SE2.	I do well on the provided English language test using the game.	4.17	0.753
SE3.	I do well on the English language test using the characters in the game, though answering English tests is usually difficult for me.	3.67	0.816
SE4.	I can answer the most difficult English language test questions by trying to reach higher levels in the game.	4.17	0.753
SE5.	I can solve the most difficult English exam questions by trying to collect the electronic awards and points in the game.	3.83	1.169
SE6.	I can finish my English exam using the game.	4.33	0.816
SE7.	I cannot answer the English language test questions without the game.	2.00	0.632
SE8.	I am not capable of excelling in English.	1.50	0.837
Second Dimension	GBL as an Assessment Tool	Mean	Std. Deviation
GE1.	Competing with opponents inspires me to continue with the English language test.	4.17	1.169
GE2.	I can finish the listening questions by having a conversation with the 'listening opponent' in the game.	4.67	0.516
GE3.	I can finish the reading questions by having a conversation with the 'reading opponent' in the game.	4.00	1.095
GE4	I can finish the grammar questions by having a conversation with the 'grammar opponent' in the game.	4.00	1.095
GE5.	I can finish the pronunciation questions by having a conversation with the 'pronunciation opponent' in the game.	4.17	1.169
GE6.	I can finish the vocabulary questions by having a conversation with the 'vocabulary opponent' in the game.	4.00	1.095
GE7.	I would take more English language tests if I could compete against different opponents in the game.	4.00	1.095
GE8.	I can complete choosing the correct answers on the English language test by competing with opponents in the game.	4.00	0.894

#### Table 1. Second instrument variables.

#### 3.3. The GBL Assessment Tool

Evidence from the literature suggests that feedback and rewarding students are among the most critical factors influencing students' self-efficacy. Therefore, both factors were considered when developing the GBL assessment tool in this study. This study uses a game created through RPG Maker MZ [63] to test the assumption of a positive relation between quantitatively measured student test performance and the GBL test environment. The RPG was selected after studying the test segment of intermediate-level students (12–14 years old), in which RPG can employ two-dimensional, non-overly obstructive designs.

The GBL assessment tool integrates several elements into the assessment process, starting at the platform Hub or village central courtyard where the student can enter different interactive game segments, including an "Exam Centre" that contains a teleportation or access point to each added exam. The student can interact with non-player characters (NPCs) that help guide and introduce the different functionalities of the platform. This experiment's focus point resides within the "exam dungeons", a map area comparable to entering a conventional LMS-based exam. Within the limits of this research, the character was instructed to head to the examination center and then travel to the English Exam. Upon dungeon entry, the platform enables an automatic NPC introduction to the exam with the promise of a reward, further integrating the concept of testing with playing a game. The English exam features five languages, "Masters", each testing the student on a different English language skill: vocabulary, pronunciation, reading, writing, and grammar. The challenge contains 24 question elements personified in multiple choice, true/false, and matching questions.

The test tool introduces gaming elements that play a role in affecting student perception and behavior. The integrated functionalities include multi-dimensional feedback and a reward/incentive function. Audio, kinetic, visual, and text-form character response feedback is provided when an answer is correct or incorrect. The reward system introduces "Experience Points" obtained when answering one question correctly. It can then be gathered for character status level up and in the form of "Reward Items" such as armor, weapons, potions, and accessories when answering exam questions and exiting the arena. The assessment tool environment shown in Figure 1 displays these elements, highlighting the design levels used within the bounds of the experiment tool development; each level is



segmented to the main elements that affect the user experience, while Figures 2–5 display different aspects of the GBL assessment tool.

Figure 1. GBL assessment tool environment.



Figure 2. Exam question formation.



Figure 3. False text-form feedback.



Figure 4. Summative result function.



Figure 5. Achievement level reward.

## 3.4. Participants

The study was conducted on 12 intermediate-level school students (12–14 years old) from different schools across public and private sectors in Saudi Arabia, and data were collected using convenience sampling, a non-random sampling technique. This sampling technique was chosen due to the nature of targeted population selection appropriation that meets the research criteria, such as intermediate school students willing to participate in the experience during COVID-19 precautionary measures. One of the concerns researchers face when using the convenience technique is possible biased sampling [64]. Therefore, in order to prevent bias, a pre-experiment test was conducted on 14 participants (35% male, 64% female) with different intermediate school levels (first grade = 6, second grade = 4, third grade = 4) with an online pre-experiment test consisting of 15 points developed by an experienced English teacher (15 years teaching intermediate school students) and reviewed by three unrelated English language experts holding a bachelor's degree in English language and literature. Informed consent was obtained from all individual participants in the study.

# 3.5. Procedure

The control group took the English test using a traditional online medium (Google Forms), whereas the experimental group took the English test on the GBL assessment tool using the researchers' laptop devices. In order to validate that the experimental group

fulfilled the online test criteria, the researchers visited the participants in their homes. They provided them with a laptop with the game downloaded and ready to use. The researchers and test moderators did not have an immediate effect on the test run but were available to provide support for any technical issues, providing the participant with the environment of having an independent online test. The experimental group test was recorded, and it was found that, on average, participants spent 22 min performing the exam (minimum 11 min and maximum 31 min). The test consisted of two sections: the first consisted of multiple-choice questions in grammar, and the second was a fill-in-the-blanks form of reading questions and forming English sentences. Two participants were excluded: one was due to the test subject's incapability to continue the experiment, and another was due to not meeting the pre-test completion condition. The participants were filtered down to 12 participants (male 33% and female 66%) with different intermediate school levels (first grade = 5, second grade = 3, third grade = 4). After assigning the participants into two groups, with the experiment group and the control group containing six students each, the two groups' distribution was normally distributed, meaning that the academic performance of the two groups was similar. The number of participants was low due to COVID-19 social distancing measures when the tests were conducted.

## 4. Results

The SPSS statistical program was used to analyze the collected data for the study's first instrument (English language test) and the second instrument (self-efficacy questionnaire). The study performed the Mann–Whitney statistical test, a non-parametric test performed on the data due to the non-random sampling technique with a sample size lower than 50 [65], to investigate differences between the experimental and control groups. The statistical test results, as presented in Tables 2 and 3, indicate no difference in English test results between the experimental and control groups; therefore, H1 is rejected, and the null hypothesis is accepted.

Table 2. Mann–Whitney results.

Null Hypotheses	Test	Sig. <sup>a,b</sup>	Decision
1. The median of the post-test is the same across categories of group type.	Independent-sample median test	1.000 <sup>c</sup>	Retain the null hypothesis
2. The distribution of the post-test represents the same across categories of group type.	Independent-sample Mann-Whitney U test	0.937 <sup>c</sup>	Retain the null hypothesis

<sup>a</sup> The significance level is 0.050. <sup>b</sup> Asymptotic significance is displayed. <sup>c</sup> Exact significance is displayed for the test.

 Table 3. Descriptive analysis of the quasi-experimental post-test only.

Group Type	Ν	Mean	Std. Deviation
Control group	6	19.667	2.58199
Experiment group	6	19.6667	2.65832

The correlation coefficient test [66] used in the study to analyze the relationship between the independent variable, the GBL assessment tool, and the dependent variable, students' self-efficacy, is the Spearman correlation coefficient, which is used in ordinal or ranked data and appropriate for analyzing the results of the Likert-scale questionnaire. As represented in Table 4, the results of the Spearman correlation coefficient value of 0.886 indicate a strong correlation between the independent variable GBL assessment tool and the students' self-efficacy of the experiment group. Therefore, it is accepted that the study's second hypothesis (H2) suggests that using the game as an examination tool positively affects students' self-efficacy. The opposite null hypothesis, which indicates that there is not a positive effect based on the (Sig. or *p*-value), is rejected.

		<b>GBL</b> Assessment Tool	Self-Efficacy
Spearman's Rho	Correlation coefficient	1.000	0.886 *
	Sig. (two-tailed)	-	0.019
	N	6	6

Table 4. Descriptive analysis of Spearman correlation coefficient results.

\* Correlation is significant at 0.05 level (two-tailed).

## 5. Discussion

This research was formulated as an experiment using a GBL assessment approach to English language testing for intermediate school students and its influence on students' performance and self-efficacy. Quantitative student knowledge retention rates were measured using students' scores, where H1 hypothesized that the GBL, as an assessment tool, positively affects the students' performance. However, it was rejected based on the statistical results, which indicate no differences in students' performance between the experimental and control groups. These results contradict similar studies in which researchers have focused on students' performance using the game as an educational intervention, influencing their knowledge level, and improving test scores. For example, ref. [54] investigated a game's effectiveness in English, and the measurement used to evaluate the participant was a test score of a vocabulary English exam, where, after using the developed tool eight times, the results showed a significant improvement in the post-test results.

Another experimental study used a quiz game as an alternative assessment for 30 postgraduate students, in which the results indicate that students' performance was significantly improved when participants used the GBL assessment tool rather than the traditional assessment method (pen and paper), and students had less anxiety during the exam [46]. Moreover, the study presented in [47] proposed a formative assessment-based contextual gaming, in which a digital game was designed to enhance students' decisions and learning processes, and the results of the quasi-experimental study showed an enhancement in student behaviors, motivation, and perceptions.

Although the research hypothesis (H1) in this study was rejected, the game did not negatively impact the test scores of students as the results indicate that both groups had the same means and same median of the test scores, suggesting that the game can be valid as an examination tool, and each group performance is similar.

H2 posited that GBL as an assessment tool positively affected the students' selfefficacy, and the hypothesis is accepted as the experimental group showed significant differences. The students expressed greater confidence in performing the English language test because of the game features, such as feedback on answered questions and various rewards (level-up, sounds, and color effects) and especially the dialogue with the English Masters or opponents in the game (i.e., Vocabulary Master, Listening Master, Reading Master, Grammar Master, and Pronunciation Master). This result is consistent with the findings of the experimental study by Hung, Huang, and Hwang [2], which used a GBL environment for children in mathematics and measured their self-efficacy, identifying an improvement in students' self-efficacy, motivation, and achievement. The work in [58] also reports an enhancement in students' self-efficacy and learning performance after adopting the badge mechanism of rewarding and achievement boards in a digital game-based English learning environment.

The results obtained from this research have aligned with those from previous studies in terms of the positive effects GBL online assessment tools have on students' self-efficacy. As for students' performance, the study did not show any positive or negative effects, which does not align with the majority of results obtained from the literature. Several factors could have contributed to this matter, such as the limited number of participants in our study, the limited period of usage of the GBL tool, for example, using it once vs. using it eight times as reported in study [54], the test settings, or the game's usability aspects. Moreover, positive results of educational technology interventions are more likely to be reported and published in the literature. Nevertheless, since no negative impact has been detected, this GBL assessment tool could still be considered a valid assessment option, especially in situations where learning styles, such as verbal-visual, are being applied. A student with a verbal learning style could take the traditional text-based test, while another student with a visual learning style might be more suited for tests using the GBL assessment tool.

## 6. Limitations and Future Work

The scope of this research was limited in terms of the small sample size and testing time limitation, as well as a lack of ability to communicate directly with students, all due to the conditions of the COVID-19 pandemic, which lasted for almost two years in Saudi Arabia (2020–2022). The pandemic restrictions challenged the process of acquiring a larger sample size due to the issues related to direct communication and the nature of the test requiring a closely monitored environment. A more significant number of participants may cause a difference in results and allow different testing and sampling techniques to be implemented. For example, it is possible to expand the range to include the entire school or more than one school in the same city, and for the GBL tool to be used repetitively over a fixed period of time.

# 7. Conclusions and Recommendations

In summary, game-based learning GBL as an assessment tool has been shown to have a positive impact on students' self-efficacy. It enhances students' belief in their own abilities, leading to improved learning outcomes. On the other hand, this study has shown that there has been neither a positive nor a negative impact on students' performance with GBL as an assessment tool. However, these results may only apply to some students, particularly those whose environment, skills, and abilities differ. Moreover, the limited number of participants and the limited test sessions, due to COVID-19 social distancing rules at the time, may have affected the performance results. Nevertheless, the work presented in this study offers a framework for exploring methods that may contribute to evolving educational progress in the field of online assessment using alternative approaches such as GBL.

Teachers applying GBL as an assessment tool in foreign language teaching are recommended to consider the following:

- 1. The alignment between the game and learning goals and objectives that are being assessed. The game should be assigned in a way to allow students to use and show their newly gained language skills effectively.
- Providing ongoing feedback to students throughout the game is essential to support their learning during this assessment activity and encourage them to progress further in the game. It is considered a type of formative assessment that aids the overall learning process.
- 3. Students should be made aware before engaging in the game about what is being assessed, for example, which skill, as well as how they are going to be assessed, and according to which criteria or rubric? This knowledge would help students focus on what they are expected to achieve while using the game.
- 4. If possible, assessment games to teach foreign languages should be designed to encourage collaborative learning, where students can interact and communicate with each other, engage in meaningful conversations, share their knowledge, and help each other. This would be particularly useful in immersive online learning environments, such as those provided by virtual reality (VR) games.

By following these recommendations, foreign language teachers can utilize gamebased learning GBL as an engaging and meaningful assessment tool in their classrooms. Author Contributions: Conceptualization, all team members; methodology, all team members; software, A.A., Z.M. (Zinah Mattar), R.A., S.A. and G.A.; validation, all team members; formal analysis, A.A., Z.M. (Zinah Mattar), R.A., S.A. and G.A.; investigation, all team members; resources, all team members; writing—first draft, all team members; writing—review and editing, M.M. and Z.M. (Zilal Meccawy); supervision, M.M.; project administration, A.A. and Z.M. (Zinah Mattar); All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: This study did not require an ethical approval.

**Informed Consent Statement:** Informed consent was obtained from the parents/guardians of all subjects involved in the study.

Data Availability Statement: Available upon reasonable request.

**Acknowledgments:** The authors thank all students and teachers who agreed to participate in this study for their time and insights.

Conflicts of Interest: The authors declare no conflict of interest.

# References

- 1. Herman, J.L. What Research Tells Us about Good Assessment. Educ. Leadersh. 1992, 49, 74–78.
- 2. Hung, C.; Huang, I.; Hwang, G. Effects of digital game-based learning on students' self-efficacy, motivation, anxiety, and achievements in learning mathematics. *J. Comput. Educ.* **2014**, *1*, 151–166. [CrossRef]
- 3. Dikli, S. Assessment at a distance: Traditional vs alternative assessments. Turk. Online J. Educ. Technol. TOJET 2003, 2, 13–19.
- 4. Black, P.; Wiliam, D. Assessment and classroom learning. Assess. Educ. Princ. Policy Pract. 1998, 5, 7–74. [CrossRef]
- 5. Fulcher, G. Practical Language Testing; Routledge: London, UK, 2013. [CrossRef]
- 6. Orlov, G.; McKee, D.; Berry, J.; Boyle, A.; DiCiccio, T.J.; Ransom, T.; Rees-Jones, A.; Stoye, J. Learning during the COVID-19 pandemic: It is not who you teach, but how you teach. *Econ. Lett.* **2021**, *202*, 109812. [CrossRef]
- 7. Meccawy, M.; Meccawy, Z.; Alsobhi, A. Teaching and learning in survival mode: Students and faculty perceptions of distance education during the COVID-19 lockdown. *Sustainability* **2021**, *13*, 8053. [CrossRef]
- 8. Meccawy, Z.; Meccawy, M.; Alsobhi, A. Assessment in 'survival mode': Student and faculty perceptions of online assessment practices in HE during Covid-19 pandemic. *Int. J. Educ. Integr.* **2021**, *17*, 16. [CrossRef]
- Putz, L.; Schmidt-Kraepelin, M.; Treiblmaier, H.; Sunyaev, A. The influence of gamified workshops on students' knowledge retention. *GamiFIN Conf.* 2018, 2018, 40–47.
- 10. Gómez, R.L.; Suárez, A.M. Gaming to succeed in college: Protocol for a scoping review of quantitative studies on the design and use of serious games for enhancing teaching and learning in higher education. *Int. J. Educ. Res. Open* **2021**, *2*, 100021. [CrossRef]
- 11. Nguyen, T.D.; Melcer, E.F.; Canossa, A.; Isbister, K.; El-Nasr, M.S. Seagull: A bird's-eye view of the evolution of technical games research. *Entertain. Comput.* **2018**, *26*, 88–104. [CrossRef]
- Al-Azawi, R.; Al-Faliti, F.; Al-Blushi, M. Educational gamification vs. game-based learning: Comparative study. Int. J. Innov. Manag. Technol. 2016, 7, 131–136. [CrossRef]
- Brackenbury, T.; Kopf, L.M. Serious Games and Gamification: Game-Based Learning in Communication Sciences and Disorders. Perspect. ASHA Spec. Interest Groups 2022, 7, 482–498. [CrossRef]
- 14. Karagiorgas, D.N.; Niemann, S. Gamification and Game-Based Learning. J. Educ. Technol. Syst. 2017, 45, 499–519. [CrossRef]
- 15. Chen, C.C.; Tu, H.Y. The effect of digital game-based learning on learning motivation and performance under social cognitive theory and entrepreneurial thinking. *Front. Psychol.* **2021**, *12*, 12750711. [CrossRef] [PubMed]
- 16. Çalışkan, H.; Kaşıkçı, Y. The application of traditional and alternative assessment and evaluation tools by teachers in social studies. *Procedia Soc. Behav. Sci.* 2010, 2, 4152–4156. [CrossRef]
- 17. Mateo, J.; Sangrá, A. Designing online learning assessment through alternative approaches: Facing the concerns. *Eur. J. Open Distance E-Learn.* **2007**, *10*, 42.
- Huang, B.G.; Yang, J.C.; Chen, S.Y. An investigation of the approaches for integrating learning materials and digital games: A prior ability perspective. *Univers. Access Inf. Soc.* 2020, 20, 57–68. [CrossRef]
- 19. Sailer, M.; Homner, L. The Gamification of Learning: A Meta-analysis. Educ. Psychol. Rev. 2019, 32, 77–112. [CrossRef]
- Bagunaid, W.A.; Meccawy, M.; Allinjawi, A.; Meccawy, Z. The Impact of Gamification on Self-Assessment for English Language Learners in Saudi Arabia. Int. J. Educ. Pedagog. Sci. 2019, 13, 117–122. [CrossRef]
- Almalki, L.; Meccawy, M. Investigating Students' Performance and Motivation in Computer Programming through a Gamified Recommender System. *Comput. Sch.* 2022, 39, 137–162. [CrossRef]
- 22. Alvarez, J.; Djaouti, D. An introduction to Serious game Definitions and concepts. *Serious Games Simul. Risks Manag.* 2011, 11, 11–15.
- 23. McClarty, K.L.; Orr, A.; Frey, P.M.; Dolan, R.P.; Vassileva, V.; McVay, A. A Literature Review of Gaming in Education. *Gaming Educ.* 2012, 1, 1–35.

- 24. Ault, M.; Craig-Hare, J.; Frey, B.B.; Ellis, J.D.; Bulgren, J.A. The effectiveness of Reason Racer, a game designed to engage middle school students in scientific argumentation. *J. Res. Technol. Educ.* **2015**, *47*, 21–40. [CrossRef]
- Taub, M.; Sawyer, R.G.; Smith, A.; Rowe, J.P.; Azevedo, R.; Lester, J.C. The agency effect: The impact of student agency on learning, emotions, and problem-solving behaviors in a game-based learning environment. *Comput. Educ.* 2020, 147, 103781. [CrossRef]
- Yien, J.; Hung, C.; Hwang, G.; Lin, Y. A game-based learning approach to improving students' learning achievements in a Nutrition course. *Turk. Online, J. Educ. Technol.* 2011, 10, 1–10.
- 27. Kusuma, G.P.; Suryapranata, L.K.P.; Wigati, E.K.; Utomo, Y. Enhancing historical learning using Role-Playing Game on mobile platform. *Procedia Comput. Sci.* 2021, 179, 886–893. [CrossRef]
- Hwang, W.; Shih, T.K.; Zhao, H.; Shadiev, R.; Chen, S.Y. Evaluating listening and speaking skills in a mobile game-based learning environment with situational contexts. *Comput. Assist. Lang. Learn.* 2015, 29, 639–657. [CrossRef]
- 29. Lam, Y.W.; Hew, K.F.; Chiu, K.F. Improving argumentative writing: Effects of a blended learning approach and gamification. *Lang. Learn. Technol.* **2018**, 22, 97–118.
- 30. Aunio, P.; Mononen, R. The effects of educational computer game on low-performing children's early numeracy skills—An intervention study in a preschool setting. *Eur. J. Spec. Needs Educ.* **2017**, *33*, 677–691. [CrossRef]
- Jagušt, T.; Botički, I.; So, H.J. Examining competitive, collaborative and adaptive gamification in young learners' math learning. Comput. Educ. 2018, 125, 444–457. [CrossRef]
- 32. Panskyi, T.; Rowińska, Z.; Biedroń, S. Out-of-school assistance in the teaching of visual creative programming in the game-based environment—Case study: Poland. *Think. Ski. Creat.* **2019**, *34*, 100593. [CrossRef]
- Alsaadi, B.; Alsaadi, B.; Alghamdi, A.; Alfhaid, M.; Almuallim, N.; Meccawy, M. Learning While Playing: Introducing Programming Concepts to Children in Minecraft. *Int. J. Online Biomed. Eng. (Ijoe)* 2022, 18, 4–24. [CrossRef]
- Yang, K.; Lu, B. Towards the successful game-based learning: Detection and feedback to misconceptions is the key. *Comput. Educ.* 2021, 160, 104033. [CrossRef]
- 35. Cheng, Y.; Lou, S.; Kuo, S.; Shih, R. Investigating elementary school students' technology acceptance by applying digital game-based learning to environmental education. *Australas. J. Educ. Technol.* **2013**, 29. [CrossRef]
- Domínguez, A.; Saenz-De-Navarrete, J.; de-Marcos, L.; Fernández-Sanz, L.; Pagés, C.; Martínez-Herráiz, J. Gamifying learning experiences: Practical implications and outcomes. *Comput. Educ.* 2013, 63, 380–392. [CrossRef]
- 37. Daubenfeld, T.; Zenker, D. A Game-Based approach to an entire physical chemistry course. *J. Chem. Educ.* **2014**, *92*, 269–277. [CrossRef]
- İnce, E.Y. Educational games in higher education. In *Simulation and Gaming*; Cvetković, D., Ed.; InTech eBooks: London, UK, 2018. [CrossRef]
- 39. Duncan, K.J. Examining the effects of immersive Game-Based learning on student engagement and the development of collaboration, communication, creativity and critical thinking. *TechTrends* **2020**, *64*, 514–524. [CrossRef]
- Hamari, J.; Shernoff, D.J.; Rowe, E.; Coller, B.D.; Asbell-Clarke, J.; Edwards, T. Challenging games help students learn: An empirical study on engagement, flow and immersion in game-based learning. *Comput. Hum. Behav.* 2016, 54, 170–179. [CrossRef]
- Papastergiou, M. Digital Game-Based Learning in high school Computer Science education: Impact on educational effectiveness and student motivation. *Comput. Educ.* 2009, 52, 1–12. [CrossRef]
- 42. Schaaf, R. Does digital game-based learning improve student time-on-task behavior and engagement in comparison to alternative instructional strategies? *Can. J. Action Res.* **2012**, *13*, 50–64. [CrossRef]
- 43. López-Faican, L.; Jaén, J. EmoFindAR: Evaluation of a mobile multiplayer augmented reality game for primary school children. *Comput. Educ.* **2020**, *149*, 103814. [CrossRef]
- 44. Hung, H.T.; Yang, J.C.; Hwang, G.J.; Chu, H.; Wang, C.C. A scoping review of research on digital game-based language learning. *Comput. Educ.* **2018**, *126*, 89–104. [CrossRef]
- Noroozi, O.; Dehghanzadeh, H.; Talaee, E. A systematic review on the impacts of game-based learning on argumentation skills. Entertain. Comput. 2020, 35, 100369. [CrossRef]
- 46. Zhonggen, Y. A Meta-Analysis of Use of Serious Games in Education over a Decade. *Int. J. Comput. Games Technol.* **2019**, 2019, 4797032. [CrossRef]
- 47. Tapingkae, P.; Panjaburee, P.; Hwang, G.; Srisawasdi, N. Effects of a formative assessment-based contextual gaming approach on students' digital citizenship behaviours, learning motivations, and perceptions. *Comput. Educ.* **2020**, *159*, 103998. [CrossRef]
- 48. Mavridis, A.; Tsiatsos, T. Game-based assessment: Investigating the impact on test anxiety and exam performance. *J. Comput. Assist. Learn.* **2016**, *33*, 137–150. [CrossRef]
- Cruaud, C. The playful frame: Gamification in a French-as-a-foreign-language class. *Innov. Lang. Learn. Teach.* 2016, 12, 330–343.
   [CrossRef]
- 50. Enayat, M.J.; Haghighatpasand, M. Exploiting adventure video games for second language vocabulary recall: A mixed-methods study. *Innov. Lang. Learn. Teach.* 2017, 13, 61–75. [CrossRef]
- 51. Ho, J. Gamifying the flipped classroom: How to motivate Chinese ESL learners? *Innov. Lang. Learn. Teach.* **2019**, *14*, 421–435. [CrossRef]
- 52. Bagunaid, W.; Meccawy, M.; Meccawy, Z.; Allinjawi, A.A. Tracking English language learners' progress through a gamified social visualisation tool to increase motivation and performance. *Int. J. Technol. Enhanc. Learn.* **2022**, *14*, 286–306. [CrossRef]

- 53. Alharbi, A.S.; Meccawy, Z. Introducing Socrative as a tool for formative assessment in Saudi EFL classrooms. *Arab World Engl. J.* **2020**, *11*, 372–384. [CrossRef]
- 54. Khayati, M.E.Z.; Hadi, M.S. Explicit Vocabulary Instruction with English Language Learners: A Charade Game. *Engl. Lang. Focus* (*ELIF*) **2021**, *3*, 39–48. [CrossRef]
- 55. Bandura, A. Organisational applications of social Cognitive Theory. Aust. J. Manag. 1988, 13, 275–302. [CrossRef]
- 56. Bonney, C.R.; Cortina, K.S.; Smith-Darden, J.P.; Fiori, K.L. Understanding strategies in foreign language learning: Are integrative and intrinsic motives distinct predictors? *Learn. Individ. Differ.* **2008**, *18*, 1–10. [CrossRef]
- 57. Zheng, D.; Young, M.; Brewer, R.A.; Wagner, M. Attitude and Self-Efficacy Change: English language learning in virtual worlds. *Calico J.* 2009, 27, 205–231. [CrossRef]
- 58. Yang, J.C.; Quadir, B.; Chen, N.S. Effects of the badge mechanism on Self-Efficacy and learning performance in a Game-Based English Learning Environment. *J. Educ. Comput. Res.* 2015, *54*, 371–394. [CrossRef]
- 59. Lu, Y.; Lien, C. Are they learning or playing? Students' perception traits and their learning Self-Efficacy in a Game-Based Learning environment. *J. Educ. Comput. Res.* 2019, *57*, 1879–1909. [CrossRef]
- 60. Frick, R.W. Accepting the null hypothesis. Mem. Cogn. 1995, 23, 132–138. [CrossRef]
- 61. Gribbons, B.; Herman, J.L. True and Quasi-Experimental designs. Pract. Assess. Res. Eval. 1996, 5, 14. [CrossRef]
- 62. Joët, G.; Usher, E.L.; Bressoux, P. Sources of self-efficacy: An investigation of elementary school students in France. *J. Educ. Psychol.* **2011**, *103*, 649–663. [CrossRef]
- 63. RPG Maker MZ | RPG Maker | Make Your Own Video Games! (n.d.). Available online: https://www.rpgmakerweb.com/ products/rpg-maker-mz (accessed on 30 November 2023).
- 64. Etikan, İ.; Musa, S.A.; Alkassim, R.S. Comparison of convenience sampling and purposive sampling. *Am. J. Theor. Appl. Stat.* **2016**, *5*, 1. [CrossRef]
- 65. Nachar, N. The Mann-Whitney U: A Test for Assessing Whether Two Independent Samples Come from the Same Distribution. *Tutor. Quant. Methods Psychol.* **2008**, *4*, 13–20. [CrossRef]
- 66. Myers, L.; Sirois, M.J. Spearman Correlation Coefficients, Differences between; John Wiley & Sons (Wiley): London, UK, 2014. [CrossRef]

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.