Students' Gains in English as a Second or Foreign Language through Game-Based Learning: A Meta-Analysis

Jason S. Barrun[1,*], Sherill A. Gilbas[2]

[1]*Language Education Department, Sorsogon State University, Magsaysay St., Sorsogon City, 4700, Philippines*
[2]*Language Education Department, Sorsogon State University, Magsaysay St., Sorsogon City, 4700, Philippines*

ARTICLE INFO

Keywords:
game-based learning;
language domains;
English language learning;
meta-analysis;
students' achievement

Article History:
Received : 06/08/2024
Revised : 20/11/2024
Accepted : 24/12/2024
Available Online: 31/12/2024

ABSTRACT

Various digitally aided applications have prompted ESL/EFL teachers to adopt them in their classrooms for games and experiments; hence, the concept of play in pedagogical settings has grown and become more modern, modifying the structures of Game-based Learning (GBL). This study, therefore, aims to analyze the effectiveness of GBL through meta-analysis. Following the set of inclusion and exclusion criteria, 16 studies in East Asia, Middle East Asia, and South America conducted from 2018 to 2023 have qualified for this study with 14 studies on tertiary level (n=822); 1 on secondary (n=56); and 1 on primary (n=46). Studies were obtained from Google Scholar, OpenAlex, Scopus, and Crossref. Further, the researchers used Harzing's Publish or Perish software to exhaust the search process. Sample size, mean, and standard deviation were analyzed using the Jamovi software version 2.4 to determine the effect sizes (Hedge's g) and the results of moderator analysis, forest plot, funnel plot, and Classic Fail-Safe N test. Findings have shown that GBL, as an approach to teaching English, had a significant and positive effect (ES=1.19) on students' achievement in different language domains. Educational levels and the type of game-based learning employed as moderators were also proven to be factors that may affect learning outcomes. Hence, more studies testing the subgroups mentioned above should be conducted to facilitate more comprehensive comparative educational research in the future.

How to cite (in APA style): Barrun, J. S., & Gilbas, S. A. (2024). Students' Gains in English as a Second or Foreign Language through Game-Based Learning: A Meta-Analysis. *OKARA: Jurnal Bahasa dan Sastra, 18*(2), 244–265. https://doi.org/10.19105/ojbs.v18i2.15127

1. INTRODUCTION

With the evolution of the terms for the generation of learners over time and with its fast transition to newer terms such as Gen Z to Gen Alpha being the latest, the trends in the teaching-learning process have seemed to be congruently on the rise as the recent digital natives of today are craving for a more relevant, interactive and innovative way of teaching. In turn, teachers have also pushed themselves to incorporate games in the classroom, be it in the form of cards, milling around, board games, reward play, and others which over the years have been upgraded into digital games like Kahoot, Pictionary, Prodigy Math, Minecraft for education, and Genially to name a few. This transition from a traditional game-based approach to digital game-based learning is a response that teachers and educators,

***Corresponding Author**: Jason A. Barrun jason.barrun@deped.gov.ph

in general, have poured efforts into to be up-to-date and flexible with the technologies that can improve learning outcomes. With the rise in popularity and ubiquitous presence of digital games, Reinhardt and Thorne (2016) argue that "it has become easier to imagine digital games as authentic, consequential, and widely applicable L2 learning resources" (p. 416).

Education has become a more modern evolution of the older and widely accepted concept of play in pedagogical settings. Game-based learning has a long and storied history. Authors of a research paper on game-based learning note that strategy-heavy board games like xiangqi, mancala, and chess, to name a few examples which have been played for thousands of years, to the ancient Greeks and Romans who understood games and play-based learning as effective didactic tools. Psychologists have long lauded the benefits of play and games in both cognitive development and learning. Roughly six decades ago, even perhaps the most well-known developmental psychologist, Jean Piaget, famously described "play" as being intertwined with cognitive development in children.

As defined by Plass, Homer, & Kinzer (2015), Game-based learning (GBL) can be defined as gameplay that incorporates educational objectives. In game-based learning, students are kept engaged, and their attention is sustained so they can participate interactively in the discussion and activities, hence bringing out active learning from them. This experience of learning the lesson meaningfully and firsthand is one of the student-centered approaches being highlighted over the years to yield significant improvements in learning outcomes, which is also present in the wide array of literature reviews on teaching pedagogies. As (Erhel & Jamet, 2013; Stefanie Vasquez et al., 2017; and Serrano, 2019) asserted, GBL has a characteristic to motivate students in terms of concrete learning experiences in the classroom.

So and Seo (2018) posited that despite the increasing interest in games and gamification in recent years, games in general still suffer from the prevalent public perception that gameplay is merely an entertainment medium. In particular, they are often criticized for their negative effects, such as game addiction, violent behaviors, and isolation in social life.

However, technology has undoubtedly opened a new era of significant contribution to the education sector, influencing classrooms for an enriched learning experience and improved outcomes, leading to the so-called "serious games," which entertain users and have additional purposes such as training, social awareness, and education. Existing meta-reviews of the effects of serious games suggest that games can be a viable learning approach in schools when there is a tight coupling between technology and pedagogy (e.g., Connolly, Boyle, MacArthur, Hainey, & Boyle, 2012; Girard, Ecalle, & Magnan, 2013)

The use of GBL in the classroom has been widely significant; hence, the researchers felt the need to summarize the existing studies through a meta-analysis to identify its effectiveness in enhancing students' achievement in the various domains of the English language. Meta-analysis is utilized to synthesize available evidence for a given question to serve as a roadmap of relevant research, a basis for designing new studies, and a guide for making institutional policies and educational curricula (Borenstein et al., (2010).

The purpose of this study is to provide an analysis of empirical research and valuable information on the students' achievement of learning the English language using game-based learning (GBL) along with the investigation of the different types of game-based learning and the educational levels GBL has been employed to. Specifically, the researchers aim to answer the following questions:

1. How effective is the use of a game-based learning approach in enhancing the students' achievement in learning the English language?
2. Is there a significant difference among the effect sizes of game-based learning in enhancing students' achievement in terms of educational level, language domain, and type of game-based learning employed?
3. What game-based learning approaches or strategies have been investigated?

2. METHOD

2.1 Research Design

This study employed a meta-analysis research design to assess the effectiveness of Game-based Learning (GBL) in enhancing the students' achievement of second/foreign language learning, mainly English, by examining the results of selected studies on the defined variables. Borenstein et al. (2010) defined meta-analysis as a core movement to systematically synthesize the quantitative results from a collection of evidence-based knowledge depending on the study's purpose and available data.

2.2 Study Search Procedure

The researchers developed a set of criteria to guide the selection of studies for inclusion and exclusion. A software program, Publish or Perish (Harzing, 2007), was used to search for peer-reviewed articles published between 2018 and 2023, which were indexed in Google Scholar, Crossref, OpenAlex, and Scopus. The researchers used systematic procedures anchored on the study's objectives to review and synthesize the quantitative findings and results on game-based learning that were selected using the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA 2020) and a set of inclusion and exclusion criteria. Figure 1 shows the flow of the search process using the PRISMA search strategy diagram.

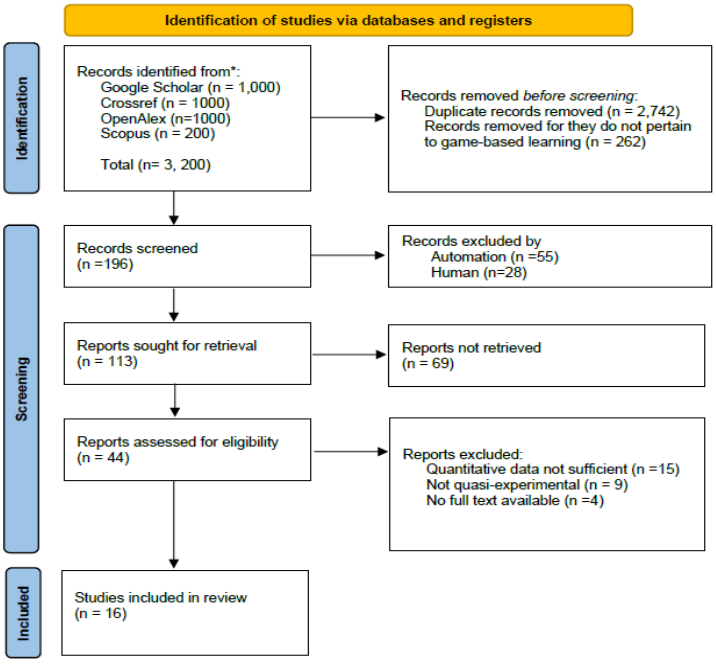

Fig. 1. PRISMA (2020) Flowchart and Search Strategy Results on the Effectiveness of Game-Based Learning on Students' Achievement of English Language Learning

Varied keywords such as "game-based learning," "language learning," and "learning the English language" were used for the search process, including similar keywords but with the phrase English language. Furthermore, a Boolean operator "in" was used for the search string to connect the term game-based learning to the phrase English language learning. These words were encoded in the Publish or Perish "keywords" bar to identify the studies easily. Then, in the "year section," the researcher delimited the scope of the search from 2018 up to 2023.

2.3 Inclusion and Exclusion Criteria

Research articles relevant to the context of this study utilizing quantitative research design from 2018 up to 2023 were investigated. Specifically, inclusion criteria protocols have been set in selecting journal articles, to wit: (a) must be a research article from a peer-reviewed journal published from 2018 up to 2023; (b) must include an explicit reference to game-based learning in its title or abstract; (c) must use language learning as the dependent variable; (d) must utilize quasi-experimental design; (e) must focus on a certain language domain such as but not limited to Oral Language Fluency, Writing & Composition, Grammar Awareness and Structure, Vocabulary Development, Reading Comprehension, Study Strategies and Attitude towards language, literacy, and literature; and (f) must provide sufficient quantitative data such as sample size, mean, and standard deviation to allow effect size computations.

In addition, studies that measured the effects of game-based learning on other variables such as motivation, self-efficacy, interest, and similar learning conditions were included in the set of studies to be examined, as long as appropriate quantitative data were provided for the student's achievement in the language domains and other variables separately. Moreover, the researcher used the PRISMA search strategy flowchart in Figure 1 to filter the collected journal articles based on the inclusion criteria specified.

2.4 Coding Procedures

The research studies included in this meta-analysis investigated the effectiveness of GBL on students' achievement of learning the English language as culled from eligible peer-reviewed journals and were coded as follows: (a) study labels (author's last name and year of publication); (b) databases; (c) educational level to which the study was implemented; (d) focused language domain; (e) type of GBL employed; and (f) outcome measure characteristics (sample size, mean and standard deviation).

2.5 Effect Size Calculation

The researchers utilized *Hedge's g* to determine the effect size of the gathered data. The Hedge's g statistic measures the effect size of the difference between means. In addition, Hedge's g statistic is generally preferred to Cohen's d statistic because it has better small sample properties and better properties when the sample sizes are significantly different (National Institute of Standards and Technology (2017). Hedge's g is more accurate than Cohen's d, most especially when sample sizes are very small (< 20) (Glen, 2016; Hedges & Olkin, 1985). Data obtained were interpreted using the values .8, .5, and .2 with descriptions of large, medium, and small effect sizes, respectively (Cohen, 1988).

In analyzing the data statistically, the researchers utilized the Jamovi software version 2.4, developed by the Jamovi Project (2023). Using this software program, the gathered

data were grouped and compared, and other necessary statistics like effect sizes (fixed and random), heterogeneity, and forest plots were identified. Further, it is recommended that the test for funnel plot asymmetry be used for at least 10 studies (Harbord et al., 2009). This recommendation led the researchers to use the aforementioned software to create a funnel plot to illustrate the publication bias in the pooled studies. Egger's test was also used, being the most widely used approach to test funnel plot asymmetry.

The final part of the Publication Bias Analysis calculates a Failsafe-N, first described by Rosenthal (1979), a test of combined significance. The failsafe number is the number of missing studies averaging a z-value of zero that should be added to make the combined effect size statistically insignificant. Hence, a big fail-safe N with a small p-value discredits the null hypothesis and asserts a general relationship between the response and predictors, making the studies free of publication bias. The Begg-Mazumdar test was also part of the calculation in the Jamovi software, which calculated the p-value. If the p-value is less than .05, then there is publication bias.

3. RESULTS

A sample size of 924 students (primary, secondary, and tertiary level) from the 16 qualified empirical research studies was identified for inclusion in this meta-analysis. Table 1 presents the frequency of studies, the student's educational levels, and the various language domains being investigated.

Table 1
Number of students' educational levels and language domains studied

Educational Level	**Frequency**	**Percentage**
Primary Level	1	6.25%
Secondary Level	1	6.25%
Tertiary Level	14	87.5%
Language Domain	**Frequency**	**Percentage**
Vocabulary Development	5	31.25%
Grammar Awareness	1	6.25%
Grammar Awareness & Vocabulary Development	1	6.25%
Reading Comprehension	1	6.25%
Self-Efficacy in Reading	1	6.25%
Self-Efficacy in Writing	1	6.25%
Self-Efficacy in Listening	1	6.25%
Self-Efficacy in Speaking	1	6.25%
Attitude towards learning English	1	6.25%
Motivation to learn English skills in general	1	6.25%
Motivation to learn vocabulary	1	6.25%
Motivation to learn speaking	1	6.25%
TOTAL	**16**	**100%**

As shown in Table 1, 14 studies obtained from 2018 up to 2023 utilized tertiary-level students (n = 822), one study was conducted with secondary students (n = 56), and one study used primary-level students (n = 46). Of the 16 studies, 6 are from East Asia (Taiwan = 6), 9 are from Middle East Asia (Turkey = 5, Iran = 3, Saudi Arabia = 1) and 1 is from South America (Ecuador = 1). It can be noticed that most of the studies are from tertiary

(educational level) countries, and the countries where GBL was employed were dominated by Asia. The major reason for the minimal number of qualified articles was that some of the gathered articles did not meet the set inclusion criteria, and some lacked the necessary statistical information presented in the PRISMA 2020 diagram.

3.1 Effectiveness of GBL in Enhancing Students 'Achievement in Learning the English Language

Table 2 presents the overall effect size, heterogeneity analysis, and confidence intervals based on the analysis effect model generated using the meta-analysis in Jamovi software (2023).

Table 2
Overall Effect Size for English Language Learning

	k	Estimate	SE	Z	P	CI Lower Bound	CI Upper Bound
Random	16	1.19	0.263	4.55	<.001	0.679	1.708
Fixed	16	0.984	0.0730	13.5	<.001	0.841	1.127

Note: Tau^2 Estimator: Hedge's g

A total of k =16 studies were included in this analysis. The estimated average standardized mean difference based on the random-effects model was g = 1.19 (95% CI: 0.679 to 1.708), interpreted as a large positive effect size per the National Institute of Standards and Technology (2017). The average outcome results differed significantly from zero (z = 4.55, p = <.001), implying that the use of game-based learning in student learning achievement of language domains has a significantly large and positive effect as determined by the overall weighted random effect size of 1.19.

Furthermore, the heterogeneity analysis in Table 3 was found to be significant (p-value < 0.001), and the Q-stat with 15 degrees of freedom is 164.083, indicating that the studies included in the meta-analysis do not share common effect sizes and are, therefore, significantly heterogeneous. In addition, according to the Q-test, the true outcomes appear to be heterogeneous (Q (15) = 164.0832, P <. 0001, tau^2 =1.0098, I^2 = 92.1734%), A 95% prediction interval for the true outcomes is given by -0.8424 to 3.2288. Hence, although the average outcome is estimated to be positive, in some studies, the true outcome may, in fact, be negative. Further, the results denote that the appropriate method to synthesize the studies in this meta-analysis is the random-effect method (Borenstein et al., 2011).

Table 3
Heterogeneity Analysis

Tau	Tau^2	I^2	H^2	R^2	df	Q	p
1.005	1.0098(SE=0.4031)	92.17%	12.777		15.000	164.083	<.001

Further, I^2 obtained a high score of 92.17%, suggesting that moderator or subgroup analysis is worthwhile (Borenstein et al., 2011). This may back up and strengthen the claim that game-based learning is a successful teaching and learning method across educational levels and settings. Table 4 capsulizes the results in the accepted studies to show the distribution of effect sizes. Figure 2 presents the forest plot and a detailed analysis of each meta-analyzed study, providing context for the analysis.

Table 4
Effect Size Distribution in Included Studies

	Hedge's g	SE	SE^2	LL	UL	p-value
Ahmed et al. (2022)	1.96	0.54	0.30	1.30	2.67	<.001*
Ahmed et al. (2022)	1.41	0.45	0.20	0.80	2.06	<.001*
Ahmed et al. (2022)	3.01	0.74	0.55	2.24	3.83	<.001*
Alhebshi & Gamlo (2022)	0.79	0.33	0.11	0.25	1.35	0.004*
Solano (2022)	0.86	0.37	0.14	0.28	1.41	0.003*
Kao (2020)	1.61	0.65	0.42	0.80	2.49	<.001*
Chen & Yeh (2019)	3.06	0.61	0.37	2.43	3.76	<.001*
Chen & Yeh (2019)	3.05	0.61	0.37	2.41	3.75	<.001*
Altiok & Baser (2018)	0.04	0.26	0.07	-0.45	0.54	0.865
Altiok & Baser (2018)	-0.07	0.26	0.07	-0.57	0.42	0.776
Altiok & Baser (2018)	0.50	0.28	0.08	-0.01	1.01	0.055
Altiok & Baser (2018)	0.75	0.30	0.09	0.24	1.27	0.004*
Altiok & Baser (2018)	0.17	0.26	0.07	-0.33	0.67	0.502
Chen et al. (2018)	0.61	0.34	0.12	0.02	1.21	0.043*
Wu (2018)	0.78	0.31	0.09	0.27	1.30	0.003*
Hung (2018)	0.90	0.37	0.13	0.32	1.51	0.003*

*p-value < alpha =0.05

The observed standardized mean differences (Hedge's g), as shown in Table 3, range from -0.07 to 3.06, with the majority of the estimates being positive (94%), which implies that the majority of the distribution of the effect sizes in studies favored the experimental group (with GBL) group over the control (non-GBL) group using traditional approaches. The estimate of the standardized mean difference based on the random-effects model is g= 1.19 (with 95% CI: 0.679 to 1.708). Therefore, the average outcome differed significantly from zero (z= 4.55, p = <.001). The overall distribution of the effect sizes is also evident in the forest plot in Figure 2 below.

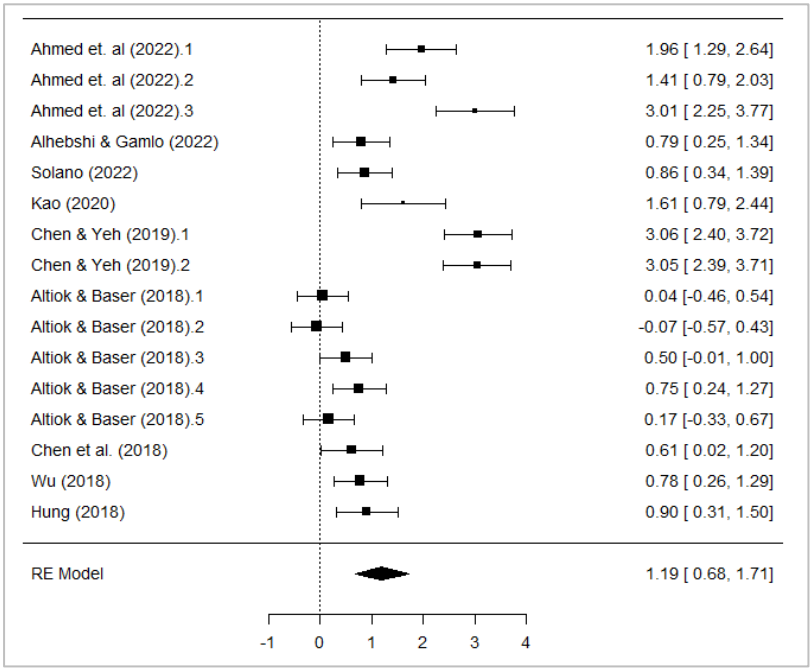

Fig. 2. Forest Plot of the Included Studies

The forest plot in Figure 2 illustrates the individual effect size for each study included in this analysis. This forest plot provides visualizations of each study's unweighted d values and their 95% confidence intervals (CIs). This plot was generated using the software suite Jamovi 2023. In addition to the upper and lower CI values reported for each study, CIs are visualized by the width of the center lines, with wider lines indicating wider CIs. Further, the squares in the middle of each line indicate the size of the standard error in each study. Larger squares indicate less standard error, which typically coincides with larger samples and tighter CIs (Blair & Blair, 2015). The rhombus at the bottom of the figures and the lines stemming from it indicate the unweighted aggregate mean effect size (Cohen's d) and the aggregate 95% CIs. The CIs do not cross zero, which indicates statistically significant positive effects from the treatment. These findings suggest that, using Plonsky and Oswald's (2014) L2 research benchmarks, learners can generally expect small to medium positive effects on learning outcomes in DGBL contexts.

The researchers conducted the Classic Fail-Safe N Analysis to validate further the obtained effect of game-based learning (GBL), of which the studies strongly favor the experimental group. Table 5 shows the results of this analysis.

Table 5
Publication Bias Statement

Test Name	Value	p
Fail-Safe N	1285.000	<.001
Begg and Mazumdar Rank Correlation	0.767	<.001
Egger's Regression	4.439	<.001
Trim and Fill Number of Studies	0.000	

Note: Fail-safe N Calculation using the Rosenthal Approach

From the results of the Classic Fail-Safe N analysis in Table 5, the meta-analysis of 16 empirical studies is valid (p <.001), hence making the effect sizes of this study sufficient, valid, and resistant to publication bias. This outcome is supported by the Begg-Mazumdar Test, as shown above in Table 5, and by the funnel plot in Figure 3.

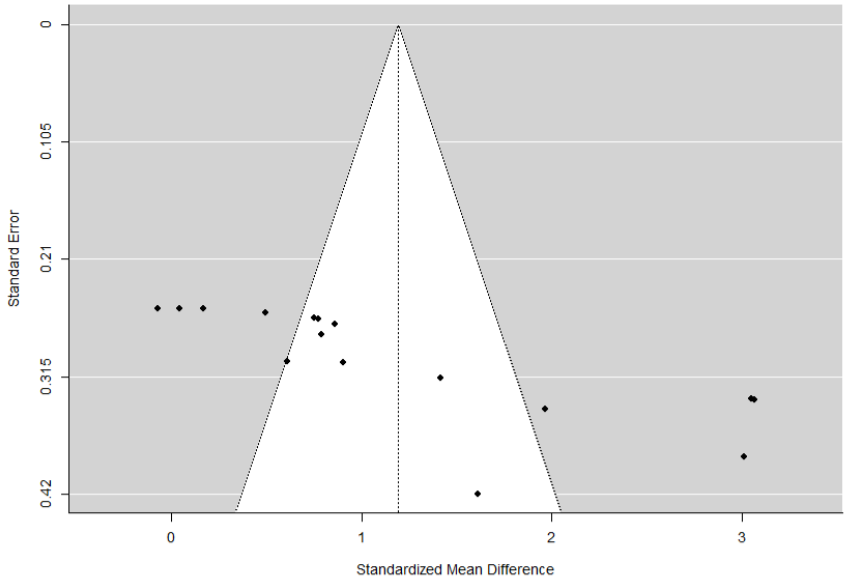

Fig. 3. Funnel plot showing publication bias status of included studies

As shown in Figure 3, the funnel plot illustrated eight outliers among the studies included in this meta-analysis, which account for its asymmetry. However, the Begg-Mazumdar test result in Table 5 yields a p-value of 0.767 (p > .05), indicating no publication bias among studies. Funnel plot asymmetry is not a guaranteed test for publication bias, especially for small studies (Harbord et al., 2009). Meanwhile, based on Janhavi and Anwaya (2017), the Begg-Mazumdar test is fairly powerful for meta-analysis, with more than or equal to 75 studies, but has low power with less than 25 studies. In this regard, the researchers considered the results of Classic Fail-Safe N. They moved directly to the moderator analysis to determine the significant difference in effect sizes between groups (educational levels, language domains, and types of GBL employed).

3.2 Significant Differences among the Moderators of the Included Studies

Table 6 shows the results of the moderator analysis to substantiate the results of the significant differences among the three moderators in the included studies namely educational level, language domain, and type of GBL employed.

Table 6
Moderator Analysis

Moderator	Subgroups	k	Hedge's g	95% CI		p-value	Test of Heterogeneity		
				LL	UL		Q	Df (Q)	p
Educational Level	Primary Level	1	0.61	0.02	1.21	0.043*	47.871	2	<.001*
	Secondary Level	1	0.79	0.25	1.35	0.004*			
	Tertiary	14	2.24	2.07	2.42	<.001*			
Language Domain	Vocabulary Development	5	1.19	0.94	1.44	<.001*	116.604	11	<.001*
	Grammar Awareness	1	1.61	0.80	2.49	<.001*			
	Grammar Awareness & Vocabulary Development	1	0.83	0.28	1.41	0.003*			
	Reading Comprehension	1	3.05	2.41	3.75	<.001*			
	Self-Efficacy in Reading	1	0.04	-0.45	0.54	0.865			
	Self-Efficacy in Writing	1	-0.07	-0.57	0.42	0.776			
	Self-Efficacy in Listening	1	0.49	-0.01	1.01	0.055			
	Self-Efficacy in Speaking	1	0.75	0.24	1.27	0.004*			
	Attitude towards learning English	1	0.17	-0.33	0.67	0.502			
	Motivation to learn English skills in general	1	2.99	2.24	3.83	<.001*			
	Motivation to learn vocabulary	1	0.61	0.02	1.21	0.043*			

(continue on the next page)

Table 6 (*Continue*)

	Motivation to learn speaking	1	0.90	0.32	1.51	0.003*			
Type of GBL Employed	Digital Game-Based Learning	12	1.12	0.95	1.29	<.001*	81.43	2	<.001*
	Mixed (Traditional & Digital) GBL	3	3.03	2.64	3.44	<.001*			
	Digital GBL (integrated w/ augmented reality)	1	0.61	0.02	1.21	0.043*			

Random-effects Model, *p < alpha = 0.05

As presented in Table 6, the moderator analysis was performed to identify the significant difference in effect sizes among the students' achievement according to their educational level, the English language domain they were investigated in, and the type of GBL employed for the study.

First, on the educational level where game-based learning was used, the secondary and tertiary levels obtained large and positive effect sizes of 0.79 and 2.24, respectively, on improving the students' achievement. Meanwhile, the primary level has a medium effect size of 0.61. However, it can be noticed that the frequency distribution as to the level was uneven; the tertiary level had the highest (k = 14) while the elementary and secondary levels had (k = 1). Though there is uneven distribution due to the inclusion criteria specified to accept studies for this meta-analysis, it can still be implied that GBL can have a medium to large positive effect size across educational levels if used in an English classroom. This implies that if in the future there are more studies that qualify in the inclusion criteria, the same to even higher range of effect sizes will likely be deduced because of the positive effect sizes game-based learning has proven throughout the years. In this study, a set of exclusion and inclusion criteria were set, and therefore, the results that the studies were heading toward in relation to the research questions were yielded. This is why there is a separate analysis for each of the moderators mentioned.

In addition, the heterogeneity results (Q = 47.871 > df; and p < 0.05) showed significant differences, demonstrating that the primary, secondary, and tertiary levels share no common effect sizes. This finding indicates that the effect of GBL on students' achievement, when compared with the traditional method, can vary according to the educational level. Secondly, analyzing the results of the effect sizes as to the language domains that were investigated, the domain of Reading Comprehension garnered the highest with 3.05, followed by motivation to learn English skills in general with 2.99; and the third is Vocabulary Development with 1.19. These three gained large and positive effect sizes along with other language domains, such as Grammar Awareness with 1.61, Motivation to learn speaking with 0.90, Grammar Awareness and Vocabulary Development with 0.83, and Self-Efficacy in Speaking with 0.75. Meanwhile, Motivation to learn vocabulary with 0.61 and Self-Efficacy in Listening had medium effect sizes. The domains with small effect sizes are Attitude toward learning English with 0.17, Self-Efficacy in Reading with 0.04, and Self-Efficacy in Writing with -0.07.

The heterogeneity results among the language domains showed significant differences with (Q =116.604 > df; and p < .05), indicating that the domains do not share common effect sizes. This implies that the different language domains where GBL was investigated can be included as factors affecting students' achievement in learning the

English language. Furthermore, the effect sizes may vary depending on the domain. They may be explained later in this study as being affected by the specific set of GBL strategies employed.

Lastly, there is no mention of traditional game-based learning or the type of game-based learning employed in the included studies. Digital game-based learning, which is the other type, was utilized, and for the purpose of this meta-analysis, the researchers operationally defined two other types of GBL, namely Mixed Game-based Learning (a combination of traditional and digital GBL) and Digital GBL (with integration of Augmented Reality). The specifications of these types will be given in Table 7.

Among the 16 included studies, the dominantly used type is Digital GBL (k=12) with a large and positive effect size of 1.12, followed by the Mixed GBL (k=3) with 3.03, the largest effect size among the three types and the last is Digital GBL (with integration of Augmented Reality) with a medium effect size of 0.61. The heterogeneity test again showed a significant difference with (Q > df, and p <.05). The three types do not share common effect sizes, indicating that the effect of GBL, when used against the traditional method, would vary according to the type of GBL employed. These results could be further associated with the differences among the various approaches/strategies in each type of GBL, as shown in Table 7.

3.3 Game-based Learning Approaches and Strategies that Were Investigated

Table 7 presents the various combinations of strategies and techniques for implementing Game-Based Learning. Of the 16 included studies, 12, or 75%, utilized digital GBL, 3, or 18.75%, used a mix of digital and traditional GBL, and 1, or 6.25%, used digital GBL with the integration of augmented reality.

Table 7
Game-based Learning (GBL) Types, Approaches and Strategies Used

Author (s)	Year of Publication	GBL Types, Approaches, and Strategies Used
Ahmed et al.	2022	**Digital:** Use of Kahoot as a platform/game for 80-item vocabulary instruction where; - Students entered the game PIN; - One new vocabulary was displayed on the screen. Its meaning was provided for the students in a multiple-choice format; - Four different meanings were provided, and the students were required to guess and select the right option. *This procedure continued to teach ten words in each session, and students responded to ten multiple-choice questions about vocabulary by clicking the options. -When all students answered and selected their options, the quiz results were shown to see who had the highest score and was the winner of the game.
Ahmed et al.	2022	**Digital:** Spaceteam ESL (English as a Second Language) is a crazy and fun English learning game that the students play with their classmates using phones or tablets for vocabulary learning where

(continue on the next page)

Table 7 (*continued*)

Author (s)	Year of Publication	GBL Types, Approaches, and Strategies Used
		- The game had a place where the students could practice slowly saying and listening to all the words that were in the game; - Then, the meanings, synonyms, explanations, and definitions of the word were provided for the students; - Usage of words in a sentence to help the students learn the words in a related context.
Alhebshi & Gamlo	2022	**Digital:** Use of Quizizz application in a Mobile-Assisted Language Learning (MALL) - Students answer the vocabulary activities in an interactive atmosphere; - Learning of 7 words to be derived against the target vocabulary from the course content.
Solano	2022	**Digital:** Use of Genially Tool where: - Students underwent 5 grammar and 5 vocabulary games; - Students shared their perceptions using Canvas forums.
Kao	2022	**Digital:** English Extras in Business with A, An, and The - A digital game developed to acquire an English article system where there is; - Simulation of a virtual business context where students communicate and have conversational turn-taking; - Giving metalinguistic feedback through an online database for explanations of the proper use of articles; - Providing summative feedback at the end of the game
Chen & Yeh	2019	**Mixed GBL:** Student-generated questioning (SGQ) as an active learning strategy to complement game-based learning (Kahoot) in flipped classrooms, particularly for vocabulary learning and reading comprehension in language learning through: - Engagement of students in outside class activities before attending an onsite class - Watching authentic videos selected from VoiceTube, - Answering SCGQ worksheets - Creation of at least 5 questions and answers for the video
Altiok & Baser	2018	**Digital;** Game-based learning activities with Kinect technology on students' self-efficacy beliefs and attitude toward English. (Kinect is a motion sensing input device with four major components: an RGB camera, 3D depth sensors, a multi-array microphone, and built-in processing cores. Gestures or voice commands can be employed to operate any system developed with the help of a software development kit without a handheld remote or pedal controllers (Kinect for Windows, 2017). -Games were designed to interact with Kinect technology to be applied in the classroom environment.

(continue on the next page)

Table 7 (*continued*)

Author (s)	Year of Publication	GBL Types, Approaches, and Strategies Used
		- Kinect Camera technology is a manually controlled device that commands the user's gestures and speech by providing a natural user interface before the user perceives his body position, movements, voice, and so on.
Chen et al.	(2018)	Digital GBL (with integration of augmented reality); - Use of augmented reality in game-based learning with "What's the Word" a mobile game-based English vocabulary practice system based on Keller's Attention-Relevance-Confidence-Satisfaction (ARCS) model adapted from Keller's unpublished IMMS (Instructional Materials Motivational Scale).
Wu	2018	**Digital;** Use of a mobile game-based English vocabulary practice system based on the Attention-Relevance-Confidence-Satisfaction (ARCS) model - Students learned English in the classroom through traditional lecture-style instructions; - During the English vocabulary review in the last half of class, they practiced using the mobile game-based English vocabulary practice system; - Instructor and technicians helped learners troubleshoot any problems with the learning activity and system use.
Hung	2018	**Digital;** Leveraging the use of technology-enhanced board games (TEBGs) to nurture student engagement in gamified flipped classrooms with: - Integration of Quick Response (QR) codes to deliver digital content via mobile technology (QR codes were mainly used to provide the students with immediate access to the digital content presented on web pages, from where they could view the learning questions, review the source video and look up vocabulary items, using their smartphones or any other mobile devices.

It can be noted that the 16 included studies did not use a sole traditional game-based learning type, as the years when these studies were published catered to Gen Z learners who are now more adept at integrating technology into their learning, hence matching their interests and needs. Digital GBL has become popular because the sound effects, animation, and interaction in the game captivated learners. As stated in the studies of Waugh & Norman, 1965; and Chen & Chunga, 2008, such a responsive learning method stimulated what was taught, prolonged the memory, and motivated the brain nerves to connect old and new knowledge and to develop stable and complete knowledge which increased learners' learning interest and motivation and strengthened their problem-solving ability (Chou, 2012; Hwang, Wu, & Chen, 2012).

Among these various strategies and techniques is the use of Kahoot as a platform. The studies of Ahmed et al. (2022) and Chen and Yeh (2019) utilized Kahoot interactively. As affirmed in the studies of Wang, 2015; and Woo, 2014, game-based learning, like Kahoot, is an effective learning tool because it uses attractive graphic and audio user interfaces to increase students' attention and participation. Other popular digital GBL

platforms were used, like Genially, as in the study of Solano (2022), and the Quizziz app in Alhebshi & Gamlo (2022). Meanwhile, Keller's Attention-Relevance-Confidence-Satisfaction (ARCS) model was also prevalent in use, especially in enhancing learners' motivation to learn specific language domains in English, as studied by Chen et al. (2018) and Wu (2018). It was also the study of Chen et al. that experimented with integrating augmented reality into a digital game used by the students to enhance motivation to learn English vocabulary.

The techniques and specific flow of activities presented in the included studies ensured that students experiencing game-based learning firsthand are kept engaged, motivated, and active creators of their learning. GBL can keep students engaged and motivated (Annetta, Minogue, Holmes, & Cheng, 2009) and lead to deeper learning through an immersive environment in which students explore concepts, reflect on personal experiences, and solve problems (Wiburg, Chamberlin, Trujillo, Parra, & Stanford, 2018).

5. DISCUSSION

Game-based learning is an active learning technique capsulizing a constructivist, inquiry-based, and collaborative approach that uses games to improve student learning. In this case, the learning comes from playing the game, which promotes critical thinking and problem-solving skills, allowing students to experience the learning firsthand. As evidenced by this meta-analysis, a significant number of studies have already been undertaken to prove the effectiveness of GBL in improving the competence of learners of the English language concerning specific language domains. Hence, this meta-analysis tried to evaluate the effectiveness of GBL as an approach used in English teaching and in enhancing students' learning achievements.

The study results showed that GBL is widely used at the tertiary level, with a 14 out of 16 sample size, and in different English language domains. The meta-analysis of the 16 empirical studies obtained from 2018 to 2023 involving 924 students yielded an overall effect size of 1.19, which is a large and positive effect. This result implies that the use of GBL is generally effective in increasing students' (primary, secondary, and tertiary) level of achievement in English language learning, thereby supporting the studies of Erhel & Jamet (2013), Serrano (2019) and Hartt et al., (2020), that it is evident that the use of game-based learning helped to yield success and impact on student's performance, achievement, and/ or motivation.

Analyzing each study involved, the studies of Chen & Yeh (2019) with two involved language domains, Ahmed et al. (2022), another study by Ahmed et al. (2022) and Kao (2020) yielded the five highest significant and positive effect sizes of 3.06, 3.05, 3.01, 1.96 and 1.61 respectively. Meanwhile, the studies of Ahmed et al. (2022), Alhebshi & Gamlo (2022), Solano (2022), Altiok & Baser (2018 on Self-Efficacy in Speaking), Wu (2018), and Hung (2018) also yielded large and positive effect sizes of 1.41, 0.79, 0.86, 0.75, 0.78 and 0.90 respectively.

On the other hand, the study of Altiok and Baser (2018; on Self-Efficacy in Listening) garnered a medium effect size of 0.50. Another study with medium effect sizes is Chen et al. (2018), with 0.61. The studies with small effect sizes are within a single study but on different language domains. These are Altiok & Baser (2018 Attitude towards learning English) with 0.17, Altiok & Baser (2018 Self-Efficacy in Reading) with 0.04, and Altiok & Baser (2018 Self-Efficacy in Writing) with -0.07. The overall effect size (ES= 1.91) is large

and positive, indicating that GBL is an effective approach to teaching English, specifically in learning the different language domains. Further, the result of the meta-analysis favors the experimental side (p <.001), signifying that game-based learning is more effective than the traditional method of teaching English.

The presence of variability in effect sizes in the language domains has led the researchers to analyze the publication bias. There was asymmetry on the funnel plot due to the outliers, which signaled that the researchers should proceed to the results of Classic Fail-Safe N. The results led to a p-value of <.001, making the study resistant to publication bias. Further, even the Begg-Mazumdar test yields a p-value of 0.767 (p > .05), strengthening the claim that no publication bias exists among studies. This result may support, in general, that the utilization of GBL in enhancing English language learning is more effective than the conventional method.

The I2 value is large (I2 = 92.17%), signifying the need to perform a subgroup or moderator analysis (Borenstein et al., 2011). Using moderator analysis, the effectiveness of GBL among the educational levels, the language domains investigated, and the type of GBL employed were examined. GBL, as applied to the three levels, the secondary and tertiary levels, yielded large and positive effect sizes of 0.79 and 2.24, respectively. Meanwhile, the effect size on the primary level was medium, with 0.61. GBL is more effective with adult learners because adult learners have the following characteristics: self-concept, prior experience, readiness to learn, orientation to learning, and motivation to learn (Knowles et al., 1998).

Moreover, adult learners can identify what imperative knowledge of interest they lack to individually fill the gap (Linderman, 1926). Although this meta-analysis had proven that GBL is effective for adult learners, still this approach is effective in enhancing the students' achievement of primary and secondary learners. There were not enough studies for these mentioned levels, which qualified in the inclusion criteria of this meta-analysis; hence, there was a lack of comparison as to this aspect of subgroup under educational level. In addition, the heterogeneity results (Q= 47.871 > df; and p< 0.05) showed significant differences, which demonstrates that the primary, secondary, and tertiary levels share no common effect sizes; hence, the results may vary according to the educational level; GBL is employed to.

In this regard, the medium to large effect sizes in primary and secondary can also be attributed to the design of the game-based learning used. In this context, the use of the Quizzizz application in the study of Alhebshi and Gamlo (2022) and the touch of augmented reality in the study of Chen et al. (2018) suggested that interactive and modern designs of the games infused into the content of the subject and language domains to be taught can be a crucial part of the success of GBL. Gaydos (2015) concluded that "attending to design may help bring together the various perspectives that have already been applied to games. Explicitly defining design theories and improving how we share our design knowledge should enable the development of common artifacts and processes, a necessary first step for replicating findings, iterating on solutions, and moving research across disciplines" (p. 481).

GBL, as scrutinized under the language domains investigated in the included studies, also yielded varied effect sizes. The domain of Reading Comprehension garnered the highest with 3.05, followed by motivation to learn English skills in general with 2.99, and the third was Vocabulary Development with 1.19. These three gained large and positive effect sizes along with other language domains such as Grammar Awareness with 1.61, Motivation to learn speaking with 0.90, Grammar Awareness and Vocabulary Development

with 0.83, and Self-Efficacy in Speaking with 0.75. Meanwhile, Motivation to learn vocabulary with 0.61 and Self-Efficacy in Listening had medium effect sizes. The domains with small effect sizes are Attitude toward learning English with 0.17, Self-Efficacy in Reading with 0.04, and Self-Efficacy in Writing with -0.07. Further, the heterogeneity result is 116.604 ($p < .05$), indicating that the different language domains of English, as compared with the traditional method, can be factors to consider in terms of affecting students' achievement.

Peterson (2010) conducted a qualitative review of the literature and found that games supported second language acquisition in a variety of settings for a variety of language skills (e.g., speaking, listening, reading, writing, and vocabulary). Furthermore, he highlighted a variety of reasons why digital games can support second language acquisition. Among the reasons highlighted was substantial exposure to comprehensible input in the target language (Krashen, 1992), enhanced student motivation (Svensson, 2003), opportunities for authentic interaction in the target language with peers (Bryant, 2006; Peterson, 2006), lowering affective barriers that can negatively impact students' willingness to interact with their peers (Freiermuth, 2002). Consequently, these language domains should be carefully considered as a factor when undergoing game-based learning instruction, as the effectiveness would vary per domain.

The researchers also examined the varied strategies and techniques as encapsulated in the flow of instruction of the included studies. Each study had a unique way of implementing game-based learning, yet the commonly shared principle or approach in them is the fusion of constructivism, inquiry-based, and collaborative into one instruction. Given that all the studies had the characteristics and principles of a game design, the researchers proceeded to identify the unique features of each GBL in the empirical studies. These studies have their own individualized features ranging from interactive interfaces, real-time feedback, and collaborative-based games that might have enhanced student engagement and motivation, and the specific language domains targeted by the researchers. Echoing this perspective, Plass et al. (2015) suggested that a "more promising method to capture the uniqueness of game-based or playful learning can be found by focusing on how these learning environments are designed" (p. 262). In that respect, a GBL framework based on design would provide researchers with a foundation to examine how game characteristics work together and how individual game applications are unique.

In the study of Ahmed et al. (2022), students learned new vocabulary by scrutinizing and eliminating choices through a multiple-choice game in Kahoot. In another study by Ahmed et al. (2022), Spaceteam ESL was used, which involved teamwork in saying out loud the vocabulary, practice time for speaking and listening to the words, provision of synonyms, definitions, and explanations of the words, and finally, using them in a sentence.

The use of Quizzizz application in a Mobile-Assisted Language Learning (MALL) model was used by Alhebshi and Gamlo (2022) in learning words per session interactively based on the course content. The study of Solano (2022) utilized the Genially Tool in answering 5 grammar and vocabulary games and included a reflective session where students provided perceptions using Canvas forums. In the study of Kao (2022), a digital game dubbed "English Extras in Business with A, An, and The" contained a simulation of a virtual business context where students had conversational turn-taking and giving of metalinguistic feedback through the online database.

Meanwhile, Chen & Yeh (2019) incorporated student-generated questioning (SGQ) into a flipped classroom using Kahoot, which involved the engagement of students in

outside class activities before onsite, watching authentic videos selected from Voicetube, answering SGQ worksheets and creation of at least 5 questions and answers from the video viewed. The study by Altiok and Baser (2018) used games to interact with Kinect technology. This manually controlled device commands the user's gestures and speech by providing a natural user interface. In the study of Chen et al. (2018) there was the use of augmented reality in game-based learning with "What's the Word," a mobile game-based English vocabulary practice system based on Keller's Attention-Relevance-Confidence-Satisfaction (ARCS) model adapted from Keller's unpublished IMMS (Instructional Materials Motivational Scale).

The same ARCS model was used by Wu (2018) in a mobile game-based English vocabulary practice system where students learned English in the classroom through traditional lecture-style instruction. During the English vocabulary review in the last half of class, they practiced using the mobile game-based English vocabulary practice system. Lastly, the study by Hung (2018) leveraged the use of technology-enhanced board games (TEBGs) to nurture student engagement in a gamified flipped classroom with the integration of Quick Response (QR) codes to deliver digital content via mobile technology; QR codes were mainly used to provide the students with immediate access to the digital content presented on web pages, from where they could view the learning questions, review the source video and look up vocabulary items, using their smartphones or any other mobile devices.

Although the results of these studies are positive indicators of the efficacy of GBL, it is safe to recommend that the design of the game has a large influence on outcomes. The review by Boyle et al. (2016) discussed that game design can influence learning outcomes, in particular collaboration and competition, and recommended additional work to define game characteristics that must be systematically evaluated through multiple measures. Further, since GBL also has an aspect of problem-based learning as students solve games to uncover outcomes, it is important to develop the teachers' skills in the delivery of PBL approaches and strategies and its integration with the newly emerged and innovative teaching strategies (Funa & Prudente, 2021).

6. CONCLUSION

In conclusion, this meta-analysis involving 16 research studies with 924 students from various educational levels, different language domains being investigated, and different types of GBL employed demonstrated that Game-Based Learning (GBL) is effective in enhancing students' achievement in English language learning. The study's overall effect size of 1.19, as determined by the analysis, establishes GBL as an effective pedagogy in English language learning specifically for students with English as their second/foreign language (ESL/EFL). It is also concluded that there is a significant difference among the effect sizes of game-based learning in enhancing students' achievement in terms of educational level, language domain, and type of game-based learning employed. As affirmed by the heterogeneity results, teachers should consider these three subgroups as factors that would affect the learning of students.

This meta-analysis summarizes the strategies and methods utilized by studies that implement GBL at primary, secondary, and tertiary levels. The study found that different GBL strategies have varying effects on educational levels and the language domains investigated and would depend on the type of GBL used. The findings have several

pedagogical implications, including the importance of selecting GBL type, strategies, and instruction flow that are to be used based on specific student needs on the language domains they need to improve. First, the researchers recommend that with the limited number of studies on GBL in primary and secondary education, more researchers should explore the effect of GBL on these educational levels.

Secondly, since language domains can be considered factors on the effects of GBL due to the significant heterogeneity in the effect sizes, it is suggested that before undertaking a study with GBL in the classroom, future researchers should consider an in-depth literature review of the possible effect on the language domain to be investigated and that it should be the 'felt need' of the students. Lastly, incorporating the appropriate type of GBL is important. Since teaching for students nowadays entails interactive use of technology, teachers should deeply understand that students would want to learn and play in a way that is most comfortable with them and relevant to them. In turn, it is therefore recommended that teachers should be trained on different trends in digital game-based learning and/or mixed GBL (traditional with the integration of digital GBL) so they can keep abreast of the trends in teaching Gen Z and Gen Alpha learners. Further, careful preparation of the game design with its characteristics and principles should be taken into complete account as these may affect learning outcomes. Overall, the findings of this study support the use of GBL as an effective pedagogy for enhancing students' achievement in English language learning, especially in the context of ESL/EFL.

The number of studies included in this meta-analysis may be small. However, the results can be considered valid because the researchers selected the studies based on the predetermined inclusion and exclusion criteria and analyzed them for selection bias using statistical methods. Moreover, this study is limited to the results provided by the pooled studies that examined the effectiveness of game-based learning (GBL) in enhancing students' achievement in terms of specific language domains investigated. These studies employed a quasi-experimental research design and explicitly reported pretest and posttest results. Using pretest and posttest results, a positive effect size has been anticipated because an intervention was implemented. As studies comparing control and experimental groups for GBL from 2018 to 2023 are limited, especially in terms of the varied educational levels and types of GBL employed, the researchers banked on the specified exclusion and inclusion criteria instead.

Acknowledgment
The authors would like to express their sincerest gratitude to Sorsogon State University for the opportunity to conduct this research. They would like to thank Dr. Ryan V. Dio for the guidance and support, Dr. Aaron A. Funa for the insightful discussion on Meta-Analysis, and Engr. Nardito Mediana. In addition, the authors sincerely thank their families: Jason- his wife Precious, and his daughter, Faith Yuna; and Sherill- her husband Reydennis, and her children, Shenise Rhae, Sophie, Shine Reyshelle and Rhaimund Louise, for their unconditional love and care.

Availability of Data and Materials
All the data generated and analyzed during the current study are not publicly accessible due to confidentiality concerns but are available from the corresponding author upon reasonable request.

Competing Interests
The authors declare that they have no competing interests.

Funding
The author(s) received no financial support for the research, authorship, and/or publication of this article.

Authors' Contribution

Jason S. Barrun conceptualized and led the research from planning to completion. Sherill A. Gilbas guided the implementation of data gathering and initial analysis of the results. Furthermore, Mr. Barrun analyzed the collected data, drafted the initial report, and revised/finalized the current form of the paper.

Authors' Information

JASON S. BARRUN is the Senior Education Program Specialist for Planning and Research of the Schools Division Office of Masbate City. He is a student in the Doctor of Philosophy in Language Education (PhD L.E.) program at Sorsogon State University, Sorsogon City, Philippines—his interests in research range from English teaching pedagogies to bilingualism/multilingualism and language.
Email: jason.barrun@deped.gov.ph; ORCID https://orcid.org/0009-0004-0609-2310
SHERIL A. GILBAS is the Head of the Admission Services Unit and the Program Chair of the Doctor of Philosophy Language Education Department at the Graduate School of Sorsogon State University, Sorsogon City, Philippines. She is the ninth-best researcher in the institution based on the Alper-Doger (AD) Scientific Index 2024. Her research interests include Humanities, Literature, and Language and Cultural Studies.
Email: sherillgilbas@sorsu.edu.ph; ORCID https://orcid.org/0000-0003-4764-1633

REFERENCES

Ahmed, A. A. A., Ampry, E. S., Komariah, A., Hassan, I., Thahir, I., Ali, M. H., . . . Zafarani, P. (2022). Investigating The Effect of Using Game-Based Learning on EFL Learners' Motivation And Anxiety. *Education Research International, 2022*, 1–9. https://doi.org/10.1155/2022/6503139

Ahmed, A. A. A., Sayed, B. T., Wekke, I. S., Widodo, M., Rostikawati, D., Ali, M. H., . . . Azizian, M. (2022). An Empirical Study on the Effects of Using Kahoot as a Game-Based Learning Tool on EFL Learners' Vocabulary Recall And Retention. *Education Research International, 2022*, 1–10. https://doi.org/10.1155/2022/9739147

Alhebshi, A. A., & Gamlo, N. (2022). The Effects of Mobile Game-Based Learning on Saudi EFL Foundation Year Students' Vocabulary Acquisition. *Arab World English Journal, 13*(1), 408–425. https://doi.org/10.24093/awej/vol13no1.27

Altiok, S., Yukselturk, E., & Baser, Z. (2018). Using Game-Based Learning with Kinect Technology in Foreign Language Education Course. *Educational Technology & Society, 21*(3), 159–173. https://doi.org/10.5072/zenodo.41162

Annetta, L. A., Minogue, J., Holmes, S. Y., & Cheng, M. (2009). Investigating the Impact of Video Games on High School Students' Engagement and Learning About Genetics. *Computers & Education, 53*(1), 74–85. https://doi.org/10.1016/j.compedu.2008.12.020

Blair, E., & Blair, J. (2015). *Applied Survey Sampling*. London: Sage.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2010). A Basic Introduction to Fixed-effect and Random-effects Models for Meta-Analysis. Research Synthesis Methods, 1(2), 97–111. https://doi.org/10.1002/jrsm.12

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2011b). Introduction to Meta-Analysis. John Wiley & Sons.

Boyle, E. A., Hainey, T., Connolly, T. M., Gray, G., Earp, J., Ott, M., . . . Pereira, J. (2015). An Update to the Systematic Literature Review of Empirical Evidence of the Impacts and Outcomes of Computer Games and Serious Games. *Computers & Education, 94*, 178–192. https://doi.org/10.1016/j.compedu.2015.11.003

Bryant, T. (2006). Using World of Warcraft and Other MMORPGs to Foster a Targeted, Social, and Cooperative Approach Toward Language Learning. Computers & Education. Retrieved from https://scholar.dickinson.edu/cgi/viewcontent.cgi?article=1781&context=faculty_publications

Cabrera-Solano, P. (2022). Game-Based Learning in Higher Education: The Pedagogical Effect of Genially Games in English as a Foreign Language Instruction. International Journal of Educational *Methodology, 8*(4), 719–729. https://doi.org/10.12973/ijem.8.4.719

Chen, C., & Yeh, H. (2019). Effects of Integrating a Questioning Strategy with Game-Based Learning on Students' Language Learning Performances in Flipped Classrooms. *Technology Pedagogy and Education, 28*(3), 347–361. https://doi.org/10.1080/1475939x.2019.1618901

Chen, C., & Chung, C. (2007). Personalized Mobile English Vocabulary Learning System Based on Item Response Theory and Learning Memory Cycle. *Computers & Education, 51*(2), 624–645. https://doi.org/10.1016/j.compedu.2007.06.011

Chen, S., Hung, C., Chang, Y., Lin, Y., & Lai, Y. (2018). A Study on Integrating Augmented Reality Technology and Game-Based Learning Model to Improve Motivation and Effectiveness of Learning English Vocabulary. *Technology Pedagogy and Education*, 24–27. https://doi.org/10.1109/ic3.2018.00015

Chou, M. (2012). Assessing English Vocabulary and Enhancing Young English as a Foreign Language (EFL) Learners' Motivation Through Games, Songs, and Stories. Education 3-13, 42(3), 284–297. https://doi.org/10.1080/03004279.2012.680899

Connolly, T. M., Boyle, E. A., MacArthur, E., Hainey, T., & Boyle, J. M. (2012). A Systematic Literature Review of Empirical Evidence on Computer Games and Serious Games. *Computers & Education, 59*(2), 661–686. https://doi.org/10.1016/j.compedu.2012.03.004

Schmidt, S. R., & Bohannon, J. N. (1988). In Defense of the Flashbulb-Memory Hypothesis: A Comment on McCloskey, Wible, and Cohen (1988). Journal of Experimental Psychology General, 117(3), 332–335. https://doi.org/10.1037/0096-3445.117.3.332

Cooper, H., Hedges, L. V., & Valentine, J. C. (2019). *The Handbook of Research Synthesis and Meta-Analysis*. https://doi.org/10.7758/9781610448864

Erhel, S., & Jamet, E. (2013). Digital Game-Based Learning: Impact of Instructions and Feedback on Motivation and Learning Effectiveness. *Computers & Education, 67*, 156–167. https://doi.org/10.1016/j.compedu.2013.02.019

Freiermuth, M. R. (2002). Connecting with Computer Science Students by Building Bridges. *Simulation & Gaming, 33*(3), 299–315. https://doi.org/10.1177/104687810203300305

Funa, A. A., & Prudente, M. S. (2021). Effectiveness of Problem-Based Learning on Secondary Students' Achievement in Science: A Meta-Analysis. *International Journal of Instruction, 14(*4), 69–84. https://doi.org/10.29333/iji.2021.1445a

Gaydos, M. (2015). Seriously Considering Design in Educational Games. *Educational Researcher, 44*(9), 478–483. https://doi.org/10.3102/0013189x15621307

Girard, C., Ecalle, J., & Magnan, A. (2012). Serious Games as New Educational Tools: How Effective Are They? A Meta-Analysis of Recent Studies. *Journal of Computer Assisted Learning, 29*(3), 207–219. https://doi.org/10.1111/j.1365-2729.2012.00489.x

Glen, S. (2016). Hedges' g: Definition, Formula. Statistics_How_To.com: Elementary Statistics for the rest of us. http://www.statisticshowto.com/hedges-g/

Harbord, R. M., Harris, R. J., & Sterne, J. A. C. (2009). Updated Tests for Small-Study Effects in Meta-Analyses. *The Stata Journal: Promoting Communications on Statistics and Stata, 9*(2), 197–210. https://doi.org/10.1177/1536867x0900900202

Hartt, M., Hosseini, H., & Mostafapour, M. (2020). Game On: Exploring the Effectiveness of *Game-Based Learning. Planning Practice and Research, 35*(5), 589–604. https://doi.org/10.1080/02697459.2020.1778859

Harzing, A. W. (2007). Publish or Perish Software. Available from https://harzing.com/resources/publish-or-perish

Freeman, P. R., Hedges, L. V., & Olkin, I. (1986). Statistical Methods for Meta-Analysis. *Biometrics, 42*(2), 454. https://doi.org/10.2307/2531069

Hsu, T. (2019). Effects of Gender and Different Augmented Reality Learning Systems on English Vocabulary Learning of Elementary School Students. *Universal Access in the Information Society, 18*(2), 315–325. https://doi.org/10.1007/s10209-017-0593-1

Hwang, G., Wu, P., & Chen, C. (2012). An Online Game Approach for Improving Students' Learning Performance in Web-Based Problem-Solving Activities. *Computers & Education, 59*(4), 1246–1256. https://doi.org/10.1016/j.compedu.2012.05.009

Hung, H. (2018). Gamifying the Flipped Classroom Using Game-Based Learning Materials. *ELT Journal, 72*(3), 296–308. https://doi.org/10.1093/elt/ccx055

Janhavi, K., & Anwaya, N. (2017). Comparison of Different Methods of Detecting Publication Bias. PhUSE 2017. https://www.lexjansen.com/phuse/2017/as/AS04.pdf

Kao, C. (2020). The Effect of a Digital Game-Based Learning Task on the Acquisition of the English Article System. *System, 95*, 102373. https://doi.org/10.1016/j.system.2020.102373

Kennedy, K. J., & Lee, J. C.-K. (Eds.). (2018). *Routledge International Handbook of Schools and Schooling in Asia (1st edition)*. Abingdon, Oxon; New York, NY: Routledge.

Knowles, M. S., Holton, E. F., & Swanson, R. A. (1998). *The Adult Learner: The Definitive Classic in Adult Education and Human Resource Development*. New York: Butterworth-Heinemann

Krashen, S. D. (1992). *The Input Hypothesis: Issues and Implications*. New Jersey: Laredo Publishing. Retrieved from https://ci.nii.ac.jp/ncid/BA35128831

Lindeman, E. C. (2013). *The Meaning of Adult Education*. Windham Press. Retrieved from http://ci.nii.ac.jp/ncid/BA12227685

Page, M. J., Moher, D., & McKenzie, J. E. (2021). Introduction to PRISMA 2020 and Implications for Research Synthesis Methodologists. Research Synthesis Methods, 13(2), 156–163. https://doi.org/10.1002/jrsm.1535

Peterson, M. (2006). Learner Interaction Management in an Avatar and Chat-Based Virtual World. *Computer Assisted Language Learning, 19*(1), 79–103. https://doi.org/10.1080/09588220600804087

Peterson, M. (2010). Massively Multiplayer Online Role-Playing Games as Arenas for Second Language Learning. *Computer Assisted Language Learning, 23*(5), 429–439. https://doi.org/10.1080/09588221.2010.520673

Peterson, M. (2012). Learner Interaction in a Massively Multiplayer Online Role-Playing Game (MMORPG): A Sociocultural Discourse Analysis. *ReCALL, 24*(3), 361–380. https://doi.org/10.1017/s0958344012000195

Plass, J. L., Homer, B. D., & Kinzer, C. K. (2015). Foundations of GBL. *Educational Psychologist, 50*(4), 258–283.

Plonsky, L., & Oswald, F. L. (2014). How Big is "Big"? Interpreting Effect Sizes in L2 Research. *Language Learning, 64*(4), 878–912. https://doi.org/10.1111/lang.12079

Poitras, M., Péléja, L., Lavertu, G., Langlois, A., Boulerice, K., Berthelot, P., & Vincent-Lamarre, P. (2020). A Replication of Waugh and Norman (1965) Primary Memory Study. *The Quantitative Methods for Psychology, 16*(2), r1–r7. https://doi.org/10.20982/tqmp.16.2.r001

Reinhardt, J. S., & Thorne, S. (2016). Metaphors for digital games and language learning. In F. Farr & L. Murray (Eds.), Routledge Handbook of Language Learning and Technology, 415–430. https://doi.org/10.4324/9781315657899.ch30

Rosenthal, R. (1979). The "File Drawer Problem" and Tolerance for Null Results, Psychological Bulletin, Vol. 86, No. 3, 838-641.

Routledge International Handbook of Schools and Schooling in Asia. (2018b). Routledge eBooks. https://doi.org/10.4324/9781315694382

Serrano, K. (2019). The Effect of Digital Game-Based Learning on Student Learning: Literature Review. *Graduate Research Papers, 943*. https://scholarworks.uni.edu/grp/943

Svensson, P. (2009). Virtual Worlds as Arenas for Language Learning. In U. Felix (Ed.), Language Learning Online: Towards Best Practice, 123–142. Retrieved from http://www.diva-portal.org/smash/record.jsf?pid=diva2:446390

The Jamovi Project. (2023). Jamovi (Version 2.4) [Computer software]. Retrieved from https://www.jamovi.org

Vásquez, S., Peñafiel, M., Cevallos, A., Zaldumbide, J., & Vásquez, D. (2017). Impact of Game-Based Learning on Students In Higher Education. *EDULEARN Proceedings, 1*, 4356–4363. https://doi.org/10.21125/edulearn.2017.1942

Wang, A. I., Øfsdahl, T., & Mørch-Storstein, O. K. (2007). Lecture Quiz: A Mobile Game Concept for Lectures. *International Conference on Software Engineering*, 305–310.

Wang, Z., Li, H., & Liu, F. (2014b). The Design of Mobile Learning English Word Memory Management Software Function Module. in *The Design of Mobile Learning English Word Memory Management Software Function Module*, 49–53. https://doi.org/10.1109/eitt.2014.16

Wiburg, K., Chamberlin, B., Trujillo, K. M., Parra, J. L., & Stanford, T. (2018). Transforming Mathematics Teaching Through Games and Inquiry. In *K-12 STEM Education: Breakthroughs in Research and Practice* (pp. 279–304). IGI Global Scientific Publishing. https://doi.org/10.4018/978-1-5225-3832-5.ch014

Woo, J. (2014). Digital Game-Based Learning Supports Student Motivation, Cognitive Success, and Performance Outcomes. *Educational Technology & Society, 17*(3), 291–307. Retrieved from http://ifets.info/journals/17_3/22.pdf

Wu, T. (2018). Improving The Effectiveness of English Vocabulary Review by Integrating ARCS with Mobile Game-Based Learning. *Journal of Computer Assisted Learning, 34*(3), 315–323. https://doi.org/10.1111/jcal.12244