Expert Review of Contextual Learning through Extensive Reading in a Digital Game-Based System

T. Govender and J. Arnedo-Moreno

Abstract—There are many factors involved in language acquisition, but immersive language experiences and consistent exposure are two crucial components. Creating these learning-conducive settings can be challenging in a traditional classroom. However, learning a new language through context to connect words and definitions can be fun and engaging with a game-based approach. Consequently, it's important to create opportunities for this approach in a curriculum to diversify studies and excite learners about the language itself. Adding games based on their entertainment value may not translate into effective learning. Therefore, in this paper, a proposal for a gameful approach using a visual novel was evaluated. Ten industry experts from various fields of education were asked to review a game-based system that focused on acquiring language through reading and comprehension. The educational experts graded the system based on perceived acceptance and usability, learner experience and, qualitatively, its impact on education. The experts concluded that the system is an innovative and exciting implementation for language learning.

Index Terms—Educational model, expert validation, gamifying language education, digital game-based learning.

I. INTRODUCTION

Games can provide incredibly immersive environments for their users, and this component has led many researchers to explore the use of gameful approaches in language education. Immersion in a language environment and consistent exposure are two critical components for language learning to succeed. Many questions remain regarding how to implement strategies that involve game-based learning. Digital game-based language learning (DGBLL) is still evolving. As more tools and methods for interactive technology become available, the guidelines for best practices need to be updated for effective implementations.

Considering most DGBLL approaches are aimed at children, this paper proposes using a DGBLL system to engage adult language learners and encourage contextual learning through extensive reading (ER) to facilitate language acquisition. The proposed application of DGBLL for adult learners includes using the game Phoenix Wright: Ace Attorney [1], a visual novel (VN), that introduces learners to challenging new vocabulary using a well-written story and a complementary workbook. One of the motivations behind using a digital game is the cognitive benefits found in adults [2], including older adults [3]. Since the system is still in development, it is necessary to get critical feedback.

For this study, the validation of the system was based on the following two questions:
1. What is the perceived acceptance, usefulness, and feasibility of the system?
2. Does the system encourage self-efficacy and self-driven participation through the user experience?

The contribution of this paper is a proposal for a DGBLL that is still under development. The design of this DGBLL system is rooted in research and aims to advance knowledge of practical applications in adult education. Furthermore, this study seeks to inspire the possibility of incorporating older language learning techniques into the digital age. More specifically, studying the perspectives of experts in implementing a reading-based, self-driven system that could potentially be added to a language curriculum.

This study is structured as follows: Section II provides background information on similar DGBLL research; Section III goes through the design of the system as well as more detail regarding the choice of the game; Section IV describes the research design, which includes the design of the survey and semi-structured interview. Section IV also contains the results of the study, which are then discussed in Section V; Finally, Section VI concludes this paper and discusses future work in the field and for this system.

II. BACKGROUND

Adult learning has not been as heavily studied as pedagogy, and there is still much to uncover in this field regarding language acquisition. It's easy to assume that pedagogical tools and techniques would be effective for adults, but abundant evidence suggests otherwise [4]. Adults have been found to learn more effectively when exposed to problems that can be solved with knowledge gained through the learning process. Adult learners want to learn relevant knowledge and often engage in self-directed learning [5].

By having adult education experts go through a course designed for adult learners, the system can be improved to align more closely with current andragogical trends. Therefore, the system utilises game-based learning, which uses a fully-fledged game to promote self-driven learning and increase engagement [6]. As mentioned previously, andragogy suggests tapping into an adult's self-directed nature. Research has found that game-based approaches encourage autonomous language learning and self-efficacy [7]. Self-efficacy is often used to predict the effectiveness of a self-directed learning system because higher self-efficacy often correlates to higher academic achievement [8]. This prediction can be made due to the nature of people feeling more confident as they master something [9].

The digital game chosen for this study is a VN. VNs are interactive, dynamic stories and can be defined as "a narrative-based digital medium that players can guide by making decisions, thus altering the outcomes" [10]. A few studies have used or proposed VNs as a means of delivering language education to adult learners. For example, Agusakim [11] suggested using a VN to help students stay motivated in a distance learning program where students are learning independently. Another study found that VNs can be beneficial in learning conversational English because of how dialogues are structured [12].

The storytelling element also has an important place in adult learning. Story-driven content has been shown to produce similar academic results to traditional content in self-regulated learning, even in highly technical subject matter, but with a higher level of enjoyment [13]. Stories help to contextualise content and aids in connecting the learning material and its application – a crucial andragogical element.

Chen [14] provides further evidence for the contextual learning that can take place through story-driven systems. More importantly, this study focuses on language learning in adults and produced promising results for implementing such a system.

Reading is an integral part of learning a language. The process of reading stories naturally exposes a learner to new vocabulary that learners can understand within the context. Extensive reading (ER) is an approach that encourages learners to read through longer texts such as newspaper articles or novels for understanding. Unlike listening, reading is a process that can easily change pace according to the reader's situation, for example, slowing down or re-reading sections to process complex concepts. ER has been found to be effective in adult students as well [15]. The system in this study uses a more gameful approach to ER to encourage self-driven, autotelic experiences.

III. DGBLL System Design

The system proposed for language learning in this study comprises two parts: the digital game and the complementary material. When looking at andragogy, the fact that students will be learning vocabulary that will immediately become usable in the game should help to improve their recollection and solidify the new knowledge.

A. Digital Game

The system focuses on ER as a means of acquiring language through context., This study uses a VN to deliver the ER content in a way that adds interactivity and encourages voluntary participation. As the name suggests, a VN is a video game driven by a story and often contains branching scenarios with multiple endings. A commercially available VN called Pheonix Wright: Ace Attorney was chosen based on the following criteria:

1. The game controls are straightforward (the entire game can be played with only the mouse), and a tutorial is built into the first episode to onboard new players (See Fig. 1.). Less complicated gameplay allows users to focus more on the game content.

Fig. 1. Example of onboarding in the first episode of the game, which helps to introduce law jargon and controls to new players.

2. Ace Attorney provides fundamental game design elements commonly found in DGBLL applications, including instant feedback (see Fig. 2.), narrative and theme [16]. Given that these elements also align with andragogical findings, this VN is a good candidate for testing the system.

Fig. 2. Example of instant feedback when making choices in the game.

3. Voice acting is absent from the game, which can isolate the extensive reading component.
4. Although there are many challenging new words, the sentence structure remains relatively simple and conversational throughout the game.
5. Dialogue is paced in a way that helps to break up information. The pacing also aids in contextual learning through the changes in visuals based on the characters' emotions.
6. The game is conveniently divided into episodes, with further divisions within each episode (i.e., Episode 2, Day 2) that help to keep consistent gameplay time between checkpoints.
7. There is a heavy reliance on the user's ability to understand the context to progress in the story.
8. Finally, Ace Attorney has been received very well among the gaming community, with 94.8% positive feedback.[1]

B. Complementary Material

The system uses an interactive complementary workbook built in Articulate Storyline 360 [17]. The workbook serves several functions: the first function is to guide learners through the controls and mechanics present in the game; second, the workbook contains vocabulary that is used to prime students for each episode; third, the workbook offers activities with instant feedback; finally, it includes optional activities that can be used to measure the engagement of the system.

Although video games are designed to be as intuitive as

---

[1] https://steamdb.info/app/787480/

possible, people indeed have varying levels of experience with digital game environments. People with higher levels of experience will find a new game easier to navigate. Those with little to no experience may feel uncertainty and anxiety, especially with non-intuitive user interfaces [18]. For this reason, the complementary workbook contains the game controls and a short section that has tips and guidelines to assist with the onboarding process.

Ace Attorney contains an extensive collection of law jargon and challenging words. A short vocabulary section was placed at the start of each episode's workbook section to mitigate some of the stress. This priming is especially necessary at the beginning of the game since the law jargon may be entirely novel for language students. Therefore, the initial vocabulary section contained the most words (nine words in total) and the number of words decreased to a maximum of six words in successive episodes.

The compulsory workbook activities also feature an instant feedback element to keep the experience consistent throughout the system. The decision to use similar features was based on harmonizing the gameful approaches in the workbook activities with Ace Attorney.

The questions asked in the workbook comprise general comprehension questions, such as idiomatic phrases and slang (see Fig. 3), and recollection questions that ask the learner to remember events from the game. The questions are similar to the type of questions that are found in reading comprehension and literature tests. Optional questions were added to measure learner satisfaction and engagement with the system. If these questions are answered during the experiment, we can assume that system encouraged students to have an autotelic experience where the motivation to complete the work was self-driven.

Fig. 3. Examples of comprehension, recollection and optional questions are present in the complementary workbook.

Overall, the DGBLL system aims to find the effectiveness of ER in a context-based learning environment that contains words not commonly encountered in second or foreign language classes.

C. The Proposed Framework

The aim of the combination of the DGBLL and complementary material is to foster an active learning environment. Each episode of the game has a corresponding workbook section that starts with vocabulary for priming the student. The student then plays the game that activates the vocabulary and links new words and phrases to the primed vocabulary through context. Workbook activities serve to fortify the new knowledge, and then the student moves on to the next section. The cycle is illustrated in Fig. 4.
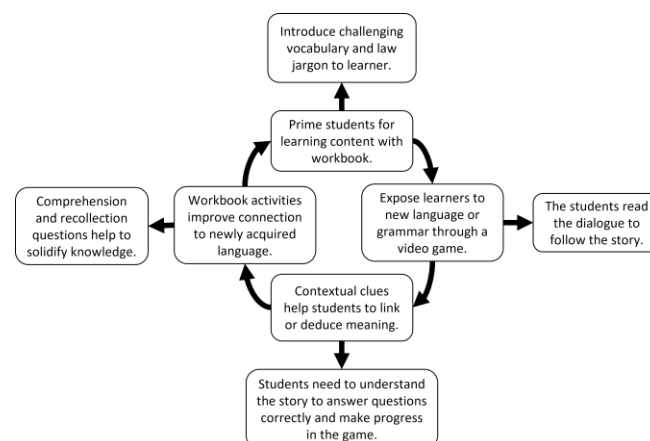

Fig. 4. The proposed learning cycle for the DGBLL system.

In theory, the students will go through the new vocabulary before the start of a new episode. They will then read the dialogue and gain an understanding of the story and the events that have occurred. They follow the story's main character, who uses evidence collected during the episode to find holes in the witnesses' testimonies. This involves the learners actively searching for potential clues and evidence in the scene (physical objects) and the dialogue (character remarks) to progress in the game. Once they have completed an episode, they will return to the workbook and test their knowledge with various types of questions.

IV. EVALUATION

The instruments for this study were chosen based on the project's stage of development, the research questions, and the fact that it involves game-based learning. We recruited experts to validate the system through a combination of quantitative data through a questionnaire and qualitative data that were gathered through interviews.

Ten experts were selected based on their areas of expertise. The group of experts comprised licensed teachers, instructional designers, and curriculum developers. Each expert was given the complementary material beforehand to read through. Since only one copy of the game was available, each expert was scheduled and attended an interview where the game could be played or streamed alongside the complementary material.

The sessions followed the same procedure. First, a brief discussion took place about the background of the study and the proposed framework. Then the expert played the game and interacted with the workbook. Not interfering with the expert was vital in determining the ease of use and the success of the onboarding (from both the workbook and VN). Each session was designed to cover the first episode of the VN (approximately 30 minutes of gameplay). The changes in the game's mechanics, such as exploration and collecting and

using evidence to open new dialogue options, were demonstrated afterwards. After completing the material, the experts filled in the questionnaire and conducted an optional interview to explain their observations.

For perceived acceptance and usefulness, Park's Technology Acceptance Model (TAM) [19] and for overall experience, O'Brien's refined user engagement scale (UES) [20] were adapted to provide valuable quantitative data (see Table II in the Appendix for more detail). The items on the questionnaire consisted of statements that could be rated on a five-point scale. The question order was randomized for each participant. Each expert also had an opportunity to provide qualitative data in a semi-structured interview after going through the system (by completing the game and complementary material).

The mean, median and standard deviation were calculated for each question and area of interest for the total population ($N = 10$). The agreement among raters was also evaluated. Table I shows the results of the computations done on the quantitative data.

TABLE I: RESULTS FROM QUESTIONNAIRE ANALYSIS

Category	Question	Mean (μ)	Std dev (σ)	Kappa (κ)	Median
PE	E1	4.8	0.4	0.56	5
	E2	4.8	0.4	0.56	5
	E3	4.7	0.46	0.42	5
	Total	4.77	0.42	0.51	5
PU	U1	5	0	1	5
	U2	4.8	0.4	0.56	5
	U3	5	0	1	5
	Total	4.93	0.25	0.85	5
AT	A1	4.8	0.4	0.56	5
	A2	4.9	0.3	0.75	5
	A3	4.9	0.3	0.75	5
	Total	4.87	0.34	0.69	5
SE	S1	4.6	0.49	0.33	5
	S2	4.7	0.46	0.42	5
	Total	4.65	0.48	0.37	5
SN	N1	4.9	0.3	0.75	5
	N2	4.7	0.46	0.42	5
	Total	4.8	0.4	0.58	5
FA	F1	4.4	0.8	0.22	5
	F2	4.3	0.78	0.14	4.5
	F3	4.6	0.49	0.33	5
	Total	4.43	0.72	0.23	5
AE	T1	4.8	0.4	0.56	5
	T2	4.9	0.3	0.75	5
	T3	4.8	0.4	0.56	5
	Total	4.83	0.37	0.62	5
RW	R1	4.7	0.46	0.42	5
	R2	4.9	0.3	0.75	5
	R3	4.9	0.3	0.75	5
	Total	4.83	0.37	0.64	5
Overall Total		4.77	0.46	0.57	5

V. DISCUSSION

From the total mean (μ) value of 4.77 (out of 5), a consistently high median value and low standard deviation (σ) of 0.46, we can describe the feedback as positive for each category with high reliability. Additionally, we assessed the reliability of agreement between experts with a combination of Fleiss' fixed-marginal multi-rater kappa [21] and Randolph's free-marginal multi-rater kappa [22]. This

resulted in a Free-marginal kappa (κ) = 0.57 with a 95% Confidence Interval (CI) of [0.47, 0.66] and a Fixed-marginal Kappa of -0.00 with a 95% CI of [-0.20, 0.20]. The overall agreement percentage was equal to 65.56%, which is above average.

A. What is the Perceived Acceptance, Usefulness, and Feasibility of the System?

The TAM questions scored a μ of 4.82, σ of 0.39, median of 5 and κ of 0.67. Questions U1 and U2 had a very high agreement rate, at 1 – with the PU category having the highest μ value. PE, AT and SN also showed positive results with an above-average agreement. Therefore, the experts agree that the system is practical and will be accepted if implemented. This information was also reiterated in the qualitative interviews, with one expert remarking that the system is "A great way to assist students with vocabulary and comprehension". Two concerns were raised regarding implementation. Given the cost of the game and the need for equipment, the feasibility in educational settings with large student populations will need to be further evaluated.

B. Does the System encourage Immersion, Self-Efficacy, and Self-Driven Participation through the User Experience?

The UES questions scored a μ of 4.82, σ of 0.39, median of 5 and κ of 0.51. Although SE was placed in the TAM section, the category is also present in UES and will be included in the user experience analysis.

Overall, the experts were excited about the learning experience. However, two categories showed poor agreement amongst raters. Both the SE and FA categories scored a κ below 0.4. Although the standard deviation for SE indicates that the raters scored similarly, FA was an outlier with a low κ, the lowest mean value and the highest standard deviation. SE was also expected to do well since this system is self-driven. Although the κ value shows a poor rating, the μ is 4.6, and the median value is 5, which means at least half of the experts agreed that the system does promote self-efficacy. However, issues with κ have been documented [23].

FA stuck out in the data set as the only category that contained 3's. This category comprises questions related to immersion. One of game-based learning's advantages is in the immersive environments that can be created to lure learners into positive learning cycles. So, this was a category that was expected to score well. The qualitative interview shed more light on why this happened. Both of the experts that scored 3's have young children. They recalled that during the time of the review, they would have to stop engaging with the system, for example, to help with a task or investigate a loud noise. Therefore, this score makes sense in the greater outlook of an adult-focused system where we can expect parents to be potential users. About the experience, one expert noted that "[They felt] that the system captures great gamified logic and concepts, to capture the attention of the students."

VI. CONCLUSION

This paper contains a proposal for a DGBLL system. Additionally, to strengthen the proposal, a validation step was included to refine the framework– an essential component of testing a new approach. There was a consensus from the

expert reviews that the system would be positively accepted in real-world implementations according to its perceived usefulness and useability. The experts also agreed that the learning experience was satisfactory. The gathered feedback will be used to create a new iteration of the material, which will then be tested on a group of language learners.

The system proposed in this study chose to isolate reading by using a visual novel that lacks voice acting. Therefore, an interesting study could entail observing the effect of voice acting versus a pure extensive reading system.

APPENDIX

TABLE II: QUESTIONNAIRE BREAKDOWN

Adapted from Paul's TAM	
Perceived ease of use (PE)	E1: I find the system easy to use.
	E2: Learning how to use the system is easy for me.
	E3: It is easy to become skilful at using the system.
Perceived usefulness (PU)	U1: The system would improve learning performance.
	U2: The system would increase academic productivity.
	U3: The system could make it easier to study course content.
Attitude (AT)	A1: Studying through the system is a good idea.
	A2: Studying through the system is a wise idea.
	A3: I feel positive toward the game-based system.
Self-efficacy (SE)	S1: I feel confident finding information in the system.
	S2: I have the necessary skills for using the system.
Subjective norm (SN)	N1: What game-based learning stands for is important for me as an educational expert.
	N2: I like using the system based on the similarity of my values and societal values underlying its use.
Adapted from O'Brien's refined UES	
Focused attention (FA)	F1: I lost myself in this experience.
	F2: The time I spent using the system just slipped away.
	F3: I was absorbed in this experience.
Aesthetic appeal (AE)	T1: This system was attractive.
	T2: This system was aesthetically appealing.
	T3: This system appealed to my senses.
Reward factor (RW)	R1: Using the system was worthwhile.
	R2: My experience was rewarding.
	R3: The content of the system incited my curiosity.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Conceptualization, T.G. and J.A.M.; methodology, T.G and J.A.M.; validation, J.A.M.; formal analysis, T.G.; investigation, T.G.; resources, J.A.M; data curation, T.G.; writing--original draft preparation, T.G.; writing--review and editing, T.G. and J.A.M.; visualization, T.G.; supervision, J.A.M.; project administration, J.A.M.; All authors have read and agreed to the published version of the manuscript.

REFERENCES

[1] Capcom, *Pheonix Wright: Ace Attorney*, Chūō-ku, Osaka, Japan: Capcom, 2001.

[2] D. Vlachopoulos and A. Makri, "The effect of games and simulations on higher education: A systematic literature review," *International Journal of Educational Technology in Higher Education*, vol. 14, no. 1, p. 22, Jul. 2017, doi: 10.1186/s41239-017-0062-1.

[3] D. Kaufman, L. Sauvé, L. Renaud, A. Sixsmith, and W. Mortenson, "Older adults digital gameplay: Patterns, benefits, and challenges," *Simulation & Gaming*, vol. 47, Apr. 2016, doi: 10.1177/1046878116645736.

[4] S. P. Forrest and T. O. Peterson, "It's called andragogy," *AMLE*, vol. 5, no. 1, pp. 113–122, Mar. 2006, doi: 10.5465/amle.2006.20388390.

[5] D. D. Pratt, "Andragogy after twenty-five years," *New Directions for Adult and Continuing Education*, vol. 1993, no. 57, pp. 15–23, 1993, doi: 10.1002/ace.36719935704.

[6] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, "From game design elements to gamefulness: Defining 'gamification'," in *Proc. the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 2011, p. 7.

[7] R. Zhang, G. Cheng, and X. Chen, "Game-based self-regulated language learning: Theoretical analysis and bibliometrics," *PLoS ONE*, vol. 15, no. 12, p. e0243827, Dec. 2020, doi: 10.1371/journal.pone.0243827.

[8] Z. An, C. Wang, S. Li, Z. Gan, and H. Li, "Technology-assisted self-regulated English language learning: Associations with English language self-efficacy, English enjoyment, and learning outcomes," *Front. Psychol.*, vol. 11, p. 558466, Jan. 2021, doi: 10.3389/fpsyg.2020.558466.

[9] T. Panc, A. Mihalcea, and I. Panc, "Self-efficacy survey: A new assessment tool," *Procedia - Social and Behavioral Sciences*, vol. 33, pp. 880–884, 2012, doi: 10.1016/j.sbspro.2012.01.248.

[10] T. Korhonen and R. Halonen, "On the development of serious games in the health sector - A case study of a serious game tool to improve life management skills in the Young," in *Proc.the 19th International Conference on Enterprise Information Systems*, Porto, Portugal, 2017, pp. 135–142, doi: 10.5220/0006331001350142.

[11] I. D. Agusalim, "Developing visual novel game of English conversation for DEP EEPIS," *Journal of Education and Practice*, vol. 6, pp. 113-124, 2015.

[12] M. A. Faizal, "The effects of conversation-gambits visual-novel game on students' English achievement and motivation," in *Proc. 2016 International Electronics Symposium (IES), Denpasar, Indonesia*, Sep. 2016, pp. 481–486. doi: 10.1109/ELECSYM.2016.7861054.

[13] J. Arnedo-Moreno and V. Garcia-Font, "A study on the design and application of fictional storytelling in online learning of computer security," *Applied Sciences*, vol. 11, no. 13, p. 6185, Jul. 2021, doi: 10.3390/app11136185.

[14] Z.-H. Chen, H. H.-J. Chen, and W.-J. Dai, "Using Narrative-based Contextual Games to Enhance Language Learning: A case study," *Journal of Educational Technology & Society*, 2018..

[15] B. Mason and S. Krashen, "Extensive reading in English as a foreign language," *System*, vol. 25, no. 1, pp. 91–102, Mar. 1997, doi: 10.1016/S0346-251X(96)00063-2.

[16] T. Govender and J. Arnedo-Moreno, "An analysis of game design elements used in digital game-based language learning," *Sustainability*, vol. 13, no. 12, Art. no. 12, Jan. 2021, doi: 10.3390/su13126679.

[17] Articulate, *Articulate Storyline 360*, New York, New York, United States: Articulate, 2016.

[18] C. Power, P. Cairns, A. Denisova, T. Papaioannou, and R. Gultom, "Lost at the edge of uncertainty: Measuring player uncertainty in digital games," *International Journal of Human–Computer Interaction*, vol. 35, no. 12, pp. 1033–1045, Jul. 2019, doi: 10.1080/10447318.2018.1507161.

[19] S. Y. Park, "An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning," *Journal of Educational Technology & Society*, vol. 12, no. 3, p. 13, 2009.

[20] H. L. O'Brien, P. Cairns, and M. Hall, "A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES short form," *International Journal of Human-Computer Studies*, vol. 112, pp. 28–39, Apr. 2018, doi: 10.1016/j.ijhcs.2018.01.004.

[21] J. L. Fleiss, "Measuring nominal scale agreement among many raters," *Psychological Bulletin*, vol. 76, no. 5, pp. 378–382, 1971, doi: 10.1037/h0031619.

[22] J. J. Randolph. (2005). Free-marginal Multirater Kappa (multirater K[free]): An alternative to fleiss' fixed-marginal Multirater Kappa. [Online]. Available: https://eric.ed.gov/?id=ED490661

[23] A. R. Feinstein and D. V. Cicchetti, "High agreement but low Kappa: I. the problems of two paradoxes," *Journal of Clinical Epidemiology*, vol. 43, no. 6, pp. 543–549, 1990, doi:10.1016/0895-4356(90)90158-l