11th International Conference of the Immersive Learning Research Network
Selected Academic papers for the iLRN Proceedings

DOI: https://doi.org/10.56198/tw4hct56

Virtu-ELLE Vocabulary: Evaluating Three Second Language Acquisition VR Games

Yingzi Kong[1] and Emily K. Johnson[1]

[1] University of Central Florida, Orlando, FL, USA
yingzi.kong@ucf.edu
ekj@ucf.edu

Abstract. Virtual reality (VR) continues to captivate educators across fields with its ability to create a sense of immersion. Its ties to entertainment gaming also position VR to increase learner motivation. These factors and others have prompted scholars to explore VR for applications such as second language acquisition (SLA). This is an early study in a broader series of studies investigating how interactive games can assist in SLA. This paper reports the results of a 72-participant study to evaluate the ability of three VR games to teach Portuguese vocabulary terms. Participants were beginner or non-Portuguese speakers who were able to travel to our lab on campus and play a VR game for 30 minutes. The games are part of a suite of games we have developed specifically for SLA, aiming to enhance the student experience in university language courses through immersive and interactive environments. Our findings indicate that novice learners can gain SLA vocabulary knowledge by playing a VR game, even in an isolated laboratory context. They also suggest that different designs of VR games impact their efficacy.

Keywords: Second Language Acquisition, VR Games, Vocabulary Learning, Video Games.

1 Introduction

Video games hold great potential as educational tools. Research suggests that educational video games, broadly, can increase learning outcomes, as well as learner engagement and motivation [1]. Educators from across fields have turned to video games in an effort to augment instruction, including language learning [2, 3]. The affordances of virtual reality (VR) are especially conducive to learning, with headsets allowing the player to feel immersed in a new environment while also blocking out distractions. Scholars continue to investigate the use of VR for language learning; Dhimolea, Kaplan-Rakowski, and Lin's 2022 metareview found that VR could be effective for vocabulary retention but that more work is needed in all aspects of SLA with VR [4].

Part of our overarching research goals include investigating whether or not language learners find VR engaging and helpful in their SLA journey or if they prefer computer-based or mobile-based games. VR is helpful for creating a sense of immersion and for limiting outside distractions for the learner, such as other students or mobile phone notifications, but these affordances may not be worth the additional resources required to create, maintain, and play games with VR headsets at this state of the technology, especially as companies aggressively sunset equipment after only a few years of its release. Broadly, our work seeks to discern the best uses and contexts for immersive SLA games.

To better understand how language instruction can best be supported and enhanced through the use of immersive and interactive games, we conducted a series of studies around our specialized materials, a suite of educational L2 video games we call "ELLE the EndLess LEarner," to assess how effective they are as educational tools in different contexts. The ELLE suite encompasses several games we have created to teach L2 vocabulary terms, including both VR and PC games; however, this article focuses on three VR games played on an Oculus (now Meta) Quest 1 headset. The study reported here aims to determine the efficacy of these VR games as vocabulary teaching aids. The games were designed to augment instruction in language courses; however, we have removed that context for this study to determine if the games were effective as standalone interventions to teach second language (L2) vocabulary terms when isolated from the context of a language course.

2 Background

Second language acquisition has benefited from technological advances, and video games have been used in a variety of ways to teach and improve language knowledge [3, 5]. In the absence of widely available L2 games where instructors can easily customize learning content for focused practice, much research around the use of video games for second language acquisition explore the interaction of players in commercial entertainment multiplayer games [6] and how commercial video game use for player purposes of L2 learning supports key sociocultural theory concepts including collaboration, assistance, and co-construction of meaning, and communities of practice [7]. Scholars have seen success in modifying commercial multiplayer games to enhance L2 vocabulary acquisition [8]. Even a commercial music video game (*Parappa the Rapper 2*) resulted in L2 vocabulary retention [9]. Most video games studies on L2 vocabulary learning remain limited to incidental L2 vocabulary acquisition though the playing of commercial games [10]. In a systematic review of L2 video games, Li [11] identified a need for more research on the integration of video games into language teaching. Studies have also suggested that the integration of VR technology to enhance vocabulary acquisition in the learning process has the potential to surpass traditional classroom methods, possibly due to increased motivation [12].

Sykes and Reinhardt [13] assert that video games that effectively support language learning also follow best practices in L2 teaching. First, effective L2 games provide the learner with dynamic, learner-driven tasks and goals. Next, they include opportunities for a variety of types of interactions. Players are able to interact with the game itself, for learning to take place. Additional interaction types include "wraparound tasks" that take place outside of the game but help the learner mentally engage with the game's content. Third, effective L2 learning games provide immediate, individualized feedback to players, providing just-in-time guidance. Fourth, the game world builds an educationally relevant space with narrative and context to help immerse the player in the content. Fifth, effective language learning games are motivating—continually engaging players with difficult or complicated challenges where they can learn.

VR games can certainly incorporate all the features mentioned by Sykes and Reinhardt. Studies suggest that even a mobile version of VR can create an environment that encourages interaction, creates a strong sense of presence, and facilitates motivation and learning [14]. Alvi studied 200 students in an English course who played the VR commercial L2 learning game *Mondly* in 15-minute sessions throughout the six weeks of their academic term in addition to their coursework [15]. Alvi found that the VR game provided an immersive and interactive experience that enhanced learners' sense of flow, with realistic scenarios that enhanced learning outside of the classroom. Additionally, Thrasher's work suggests that VR helped reduce L2 learners' speaking anxiety, especially over time [16]. Clearly, VR has potential to benefit SLA learners.

This study focuses on the research question, "How do educational VR games impact the vocabulary knowledge of novice language learners?" We assess how effectively each of our VR games teaches Portuguese vocabulary terms in an isolated lab context.

ELLE VR Games . We have written about the design of the ELLE backend structure in the past, describing one of our computer games [17] as well as the early versions of two of the VR games, Spin 'N' SpELLE and Highrise hELLEp [18]. All of these games were co-designed with a series of computer science undergraduates at our large, public, high-research activity (R1) Hispanic Serving Institution (HSI) in the southeastern United States. Each game was co-developed with a group of three to six undergraduates as part of their two-semester capstone computer science course. The second author served as the faculty sponsor for the multidisciplinary project and met regularly with each group throughout the two semesters providing design guidance and feedback.

Because she has sponsored so many groups over the past several years who have all worked within the same backend structure, the ELLE suite of games all connect with a robust mySQL database housing both term data and gameplay data to assist with SLA research. The games and database all connect to a single website, where language faculty can enter new modules in any language by entering written terms, uploading images, and uploading audio files or by recording directly on the website with their computer's microphone. All of the ELLE games can pull terms from any given module, allowing teachers to customize the term lists while providing a wide range of practice options for students. Students can monitor their progress and view term lists using the website, as well.[1]

This paper reports the results of a participant study evaluating three VR games in the ELLE suite: hELLE's Kitchen, Spin 'N' SpELLE, and Highrise hELLEp. They are all developed in Unity, connected to a user-friendly website and database, and all are played on Oculus (now Meta) Quest 1 headsets. Though these games were

[1] The URL for the ELLE website is: https://chdr.cs.ucf.edu/elle/games

designed to augment postsecondary language coursework, this study evaluated them outside of classroom context to assess their efficacy as standalone learning tools.

hELLE's Kitchen VR. hELLE's Kitchen is a game that emphasizes practicing vocabulary in categories, helping learners create mental associations between related words. This is designed according to the assumption that novice language learners may also benefit from learning terms in related categories together. Grouping words into three-term categorically related "recipes" like "clothing for bad weather" or "items for feet" is thought to be beneficial for L2 learners. This is supported by second language acquisition research, such as Kang's [24] and Mulder et al.'s [25] emphasis on context, as well as Sykes and Reinhardt's [13] fourth recommendation for effective L2 learning games.

In hELLE's Kitchen VR, players enter a virtual kitchen and receive a "recipe" to follow. Recipes can be actual food-based recipes, asking for ingredients to make various dishes like spaghetti, but they can also involve broader uses of the term "recipes." These might include thematic vocabulary sets, such as "school supplies," or activities that require preparing for a specific scenario, like gathering items needed for a rainy day. Players move around the kitchen to locate "ingredients," images of vocabulary items on the shelves and in the refrigerator, placing each item in a large pot on the stove. If a player attempts to add an incorrect ingredient to the pot, it will bounce out, and the player will hear a thumping noise to emphasize that the selection was wrong. Correct ingredients will disappear when added to the pot, and the player will hear a bubbling sound to indicate that the ingredient is cooking. Once all correct ingredients are added, the final "dish" will appear in the player's virtual hands, and they "serve" it to the head chefs at the kitchen window. Then they will get the next recipe or exit the game.

Fig. 1. hELLE's Kitchen VR showing sample Spanish terms in a clothing module with the category, "for work."

Spin 'N' SpELLE VR. Spin 'N' SpELLE is another VR game in the "ELLE" suite. This game emphasizes spelling practice, especially the use of accent marks, a particular challenge for native English speakers to grasp. Set in a whimsical virtual children's room, players hold magic wands in both hands to ensure accessibility for both left- and right-handed players. The wands allow players to easily summon letter blocks from a nearby treasure chest, and players use the VR controller to rotate the block to select a specific letter. Then, they place the block on a shelf next to an image and English text for the term they are required to spell. Players can also add accent mark stickers to the blocks when needed. When the word is correctly spelled with the proper accent marks, the shelf the word sits on will disappear in a puff of purple smoke, and a new shelf and word will appear.

Fig. 2. Spin 'N' SpELLE VR showing sample images and English terms to be spelled in the L2.

Highrise hELLEp VR. The third VR game is Highrise hELLEp. This game is flexible in that it can utilize categories and also handle one-to-one translation matching, basing its selection on the module's configuration in the database. Category names are displayed on the roadside sign, and terms that are labeled in the database as a certain part of speech or certain category appear on the burning balconies. Randomly selected other terms from the module are displayed on balconies that also appear to be on fire, and the player needs to point the water hose at the terms that match the category, extinguishing those flames. If the term does not match the category, the fire will not be put out by the water. Categories were not used in this study; participants matched individual terms with their translation.

Fig. 3. Highrise hELLEp VR showing Portuguese terms.

3 Methods

3.1 Participants

We recruited students through several faculty listservs at our university, encouraging faculty to share the information about our study with their students. The IRB-approved recruitment email was sent out a few times over the course of the Spring 2024 semester. To participate in the study, participants had to be at least 18 years of age, *not* fluent in Portuguese (beginner level or lower), able to travel to our campus location, and able to play a video game with visual and auditory components (with or without the use of glasses and/or hearing aids).

Participants did not have to be enrolled in courses at our university in order to participate. At the end of the study, each participant received a $20 gift card for Amazon.

Participants self-reported their demographics on the post-test. Of the 24 participants who played hELLE's Kitchen VR, 12 identified as female, 9 as male, 1 as nonbinary, and 2 chose not to answer. For the 24 Highrise hELLEp VR, 11 reported identifying as female, 8 male, 2 non-binary, 1 demi-femme, 1 transfemme, and 1 chose not to answer. The 24 participants playing Spin 'N' SpELLE VR identified themselves as follows: 7 female, 15 male, and 1 nonbinary, with 1 choosing not to respond.

For race/ethnicity, participants in hELLE's Kitchen VR responded: 3 Asian or Pacific Islander, 1 Black or African American, 4 Hispanic/Latino/Latina/Latinx/Latin, 15 White or Caucasian or European, and 1 chose not to answer. For Highrise hELLEp VR, participants reported: 2 Asian or Pacific Islander, 4 Black or African American, 6 Hispanic/Latino/Latina/Latinx/Latin, 11 White or Caucasian or European, and 1 wrote in multiple races/ethnicities. Spin 'N' SpELLE VR participants were: 3 Asian or Pacific Islander, 0 Black or African American, 9 Hispanic/Latino/Latina/Latinx/Latin, 11 White or Caucasian or European, and 1 wrote in multiple races/ethnicities.

Participants were also asked to type their age into our Qualtrics form. hELLE's Kitchen VR participant ages ranged from 18 to 52, with an average age of 22.35, a median age of 21, a mode age of 21, and 1 choosing not to respond. Highrise hELLEp VR participants spanned from 18 to 31, with an average age of 21.00, a median age of 20, a mode age of 19. Finally, participants playing Spin 'N' SpELLE VR reported ages ranging from 18 to 45, with an average age of 21.40, a median age of 20, a mode age of 19.

We also asked participants if they had experience playing video games, and on what kinds of platforms. Of the 24 hELLE's Kitchen VR participants, 21 reported that they had been playing video games for several years, 2 indicated that they did not currently play video games but used to, and one responded that they rarely or never played video games. Three of these participants also listed a VR headset as a platform on which they played video games. Of the 24 in the Highrise hELLEp VR condition, 21 reported playing video games for several years, and 3 selected the not currently but used to play video games option. Three of these participants also selected at least one VR platform as one they use to play video games. In the Spin 'N' SpELLE VR group of 24 participants, 21 indicated that they had been playing video games for several years, 3 that they did not currently but used to, and 4 of these listed a VR platform as one on which they play video games.

We did not ask participants to disclose any disabilities or diagnoses other than giving them the option to choose to play a PC game in lieu of a VR game. All participants in this study indicated that they felt able to play a VR video game with one of our Oculus Quest 1 headsets and its accompanying standard controllers.

The study sessions were held between January and April of 2024, coinciding with the spring semester at our university. While the full study also assessed PC games, this paper only focuses on the 72 participants who received the VR condition and one of three VR games designed by our team.

3.2 Materials

The study materials included a pretest, one game conditions, a posttest, and a structured post-interview. This paper reports the pretest and posttest results from the knowledge items on the pretest and posttest, administered via Qualtrics on computers in the lab.

The pretest contained a total of 34 items, with 20 translation items. Half of these asked the participant to type the Portuguese translation for an English term, and the other half asked the participant to type the English translation for a Portuguese term. Terms that would be difficult for non-Portuguese speakers to guess but could be represented with images were selected by a Portuguese language instructor whose native language is Portuguese, who is a Co-PI on the grant. The pretest also asked participants to describe their prior experience with any L2 as well as their gaming experience.

The posttest contained a total of 58 items, including the same 20 translation items repeated exactly from the pretest. The posttest also included additional items on user experience, but these results will be reported elsewhere.

3.3 Procedure

After participants responded to our recruitment email and arrived at our lab location, the researchers verbally reviewed the IRB-approved explanation of research and asked each participant to verbally confirm that they still wished to participate. Participants were advised that they could stop the study at any time. Once they consented, the game they played was determined randomly through a card draw.

Participants assigned to a VR game were given a brief verbal overview of the game that included the game's goals and how to work the handheld controllers by a research assistant, and then they played their assigned VR game for 30 minutes. During this time, participants were also asked to follow a think-aloud protocol, though this

data is still under analysis. The research assistant stayed nearby to answer technical questions and to ensure participant safety. Two participants chose to end their session early, after about 20 minutes instead of the planned 30 minutes. One became frustrated with game glitches, and the other found the headset too uncomfortable to continue. Since, both participated for two thirds of the allotted time, we have included their data in our results.

After playing their assigned VR game for 30 minutes, participants were asked to complete the Qualtrics posttest on another computer in the lab, participate in a post-interview, and then they received their gift card.

4 Results

4.1 Knowledge Items

Overall, participant knowledge of these Portuguese terms increased across all of the conditions—an average increase of 4 terms per game. The condition where participants learned the most terms was Highrise hELLEp, with a mean knowledge gain of 7 from pretest to posttest. Participants in the hELLE's Kitchen condition learned about 3.5 new terms on average, and participants in the Spin 'N' SpELLE condition learned the least—only one or two new terms each (Table 1).

Table 1. Knowledge item scores across games.

Game Condition	Pretest Mean	Posttest Mean	Mean Gain
hELLE's Kitchen	0.75	4.25	+3.500
Highrise hELLEp	1.208333	8.208333	+7.000
Spin 'N' SpELLE	1.625	3	+1.375
All VR	1.194444	5.152778	+3.958

The discrepancies in participant knowledge increases can be better understood from the perspective of term exposure. While these games are intended to supplement a language course, we study them here in insolation to assess each one outside of instructional contexts. Therefore, it makes sense to conclude that because our participants have little to no experience with Portuguese, the more exposure they had to terms in a given game, the more likely they were to learn that term. Spin 'N' SpELLE requires players to produce the Portuguese term and properly spell it, including with proper use of accent marks—which results in a lot of guessing for these novice players. Highrise hELLEp, on the other hand, displays many Portuguese terms all at the same time, and basically asks the player to match the term with its English translation. This allows those players to see a lot more terms in comparison to the Spin 'N' SpELLE condition. hELLE's Kitchen also displays several terms at once (usually 4 or so), and it asks players to match the text term with the image. This game structure, like that of Highrise hELLEp, affords complete novices a way to learn the terms through trial and error.

Table 2. ANOVA Results.

Source of Variation	SS	df	MS	F	P-value	F crit
Rows	406.4931	23	17.67361	1.928246	0.012533	1.623598
Columns	928.1181	5	185.6236	20.25212	1.89E-14	2.293205
Error	1054.049	115	9.16564			
Total	2388.66	143				

We conducted a repeated measures ANOVA to examine the learning efficacy of the three games. The analysis indicates a significance difference between pretest and posttest scores, $F(23, 115) = 1.93$, $p = 0.013$, indicating that participants' vocabulary knowledge significantly improved after playing each of the games. Additionally, there was a significant difference in learning gains between the three games, $F(5, 115) = 20.25$, $p < 0.001$, suggesting that the games differed significantly in their efficacy.

Post-hoc comparisons using pairwise t-tests with Bonferroni correction (comparing learning gains of each game against the others) indicated that the vocabulary learning gains were significantly greater for participants playing Highrise hELLEp compared to hELLE's Kitchen ($p = 0.0074$) and Spin 'N' SpELLE ($p = 0.0003$). There was no significant difference, however, in learning gains between hELLE's Kitchen and Spin 'N' SpELLE ($p = = 0.1824$). Comparisons displayed below in Table 3.

Table 3. Pairwise Comparisons of Learning Gains.

Comparisons	Mean Difference	t Stat	p-value (two-tail)	Significance ($p < 0.0167$)
hELLE's Kitchen vs Highrise hELLEp	-3.96	-2.803	0.0074	Yes
hELLE's Kitchen vs Spin 'N' SpELLE	1.25	1.354	0.1824	No
Highrise hELLEp vs Spin 'N' SpELLE	5.21	3.945	0.0003	Yes

The game with the lowest learning gains was Spin 'N' SpELLE, with an average gain of just 3 terms, and 7 of the 24 making no learning gains at all. Participants playing hELLE's Kitchen had larger learning gains than participants playing Spin 'N' SpELLE; mean posttest scores for hELLE's Kitchen were of 4.25, and for Spin 'N' SpELLE, it was just 3.00. Although this difference suggests that hELLE's Kitchen may be more effective in enhancing vocabulary knowledge, the difference is not statistically significant ($p = 0.1824$), indicating that the observed variation could be due to random chance rather than an effect of the game.

The participant with the largest vocabulary gain on the knowledge items learned 16 words, translating 4 correctly on the pretest and all 20 correctly on the posttest after 30 minutes of playing Highrise hELLEp. In contrast, the poorest performance was seen among participants who played Spin 'N' SpELLE. Specifically, 7 out of 24 Spin 'N' SpELLE participants made no gains, 8 out of 24 gained only one word.

4.2 Qualitative Insights

To complement and better understand the knowledge findings, we analyzed the post-interview recordings of the lowest and highest performer in each of the three VR game conditions. Because our isolated lab study targeted novice Portuguese learners, we selected participants who each scored a 0 on the pretest, indicating no prior knowledge of Portuguese (at least not the twenty words used in our study). From that group chose the lowest and highest knowledge gains from 0 for each condition. For hELLE's Kitchen, the low performer did not show any gains, correctly translating 0 terms on the posttest, and the high performer translated 10 terms correctly on the posttest for a gain of 10. For Spin 'N' SpELLE, the low performer again gained 0, while the high performer gained 2. While the highest posttest score on Spin 'N' SpELLE was 10, that participant scored 5 on the posttest and was excluded from this secondary analysis. For Highrise hELLEp, the low performer again had a gain of 0 and the high performer gained 10.

In their post-interviews, all six participants stated that their interest in learning the Portuguese language improved as a result of playing their assigned game. Both high performers from hELLE's Kitchen and Spin 'N' SpELLE reported enjoying the VR environment of their respective games. The high performer from Highrise hELLEp described the game as "fun" and "good," expressing confidence in the game's educational value by noting, "I learned at least four words."

A frequently mentioned negative emotion across all groups was frustration and confusion, mainly due to technical glitches, unclear gameplay instructions, insufficient language instruction, and a lack of feedback mechanisms. Meanwhile, the low performers from hELLE's Kitchen and Spin 'N' SpELLE felt the games did not aid their learning at all.

Interestingly, the high performer in hELLE's Kitchen expressed dissatisfaction with how items were grouped within each conceptual "recipe," stating that the groupings "didn't make sense." They also noted that some visual elements did not effectively represent the corresponding L2 terms. However, this clearly didn't impede their ability to learn 10 words in just 30 minutes of play. The low performer from Highrise hELLEp described the game as "not really interactive" and "not really challenging," while still praising its layout as "nice" and "fun." The high performer reported enjoying the, "feeling of success of getting it right," adding, "I liked, you know, finding that I had matched the words and was learning and getting them correct repeatedly."

The low performer of the hELLE's Kitchen and the high performer of the Highrise hELLEp specifically mentioned that they think the games relied more on short-term memory/temporary memory. Both the high and low performers of Spin 'N' SpELLE found the VR setting or physical interaction enjoyable but found the game frustrating due to a lack of feedback, limited learning and the repetitive music. They both complained about the game mechanics to be confusing and didn't feel it helped with learning. The low performer also hoped the game could add in pronunciation practice. Summaries of these participants' interview answers are reported in Table 4.

Table 4. Participant interview insights from hELLE's Kitchen (HK), Spin 'N' SpELLE (SS), and Highrise hELLEp (HH).

Interview Question	HK High	HK Low	SS High	SS Low	HH High	HH Low
Emotions during gameplay	▪Confusion ▪Curiosity ▪Satisfaction	▪Frustration ▪Aggravation	▪Frustration ▪Annoyance	▪None	▪Frustration ▪Fun ▪Confusion ▪Sense of achievement	▪Frustration ▪Boredom
Elements they liked most	▪VR setting	▪Language learning	▪VR setting	▪Physical interaction	▪Feelings of success	▪"Nice" game ▪"fun" layout
Elements they liked least	▪Incoherent word categories	▪Glitches	▪Not enough interaction, such as correction or explanation	▪Not learning anything	▪No instruction or feedback of the words	▪Simple game mechanics
Increased interest in language	▪Yes	▪Yes	▪Yes	▪Yes	▪Yes	▪Yes
L2 aspects the game improved	▪Vocabulary	▪Vocabulary	▪No SLA ▪Patience	▪None	▪Word recognition	▪None
Most useful game features	▪Visual elements (pictures representing words)	▪Nothing (pictures help a little bit)	▪Timer	▪Visual elements	▪Visual elements	▪Mechanics control
Suggested changes to the games	▪Add pronunciation	▪More hints	▪Improved mechanics (feedback system) ▪Various music	▪More audio ▪Instruction on game mechanics ▪Instruction on words to be learned	▪Add reinforcement mechanics ▪Reduced gameplay time ▪Tailored word lists	▪Add a feedback system ▪More simulation ▪More complex mechanics ▪Add translation

5 Discussion and Implications

Participants who had the highest learning gains were those in the Highrise hELLEp condition. This could be due to the repeated exposure to the new vocabulary terms and the opportunities to figure out the term translations through low-stakes trial and error. In the interview feedback, some participants identified repetition as a tedious element, but also noted that the repetition likely improved their vocabulary retention. While categorizing words under relevant contexts may enhance vocabulary acquisition, selecting appropriate categories remains subjective.

Conversely, Spin 'N' SpELLE inherently requires players to already know the translation of the term and to correctly spell it. This means that a participant who did not know the term had to guess, letter by letter, with feedback only involving the correctness of the whole word—inclusive of accents and other diacritical marks. The fact that participants learned the least from this game confirms our assumptions as well as SLA research that novice L2 language learners need repeated exposure to new terms. Since Spin 'N' SpELLE focused on spelling without exposing the novice to the target vocabulary terms at all, it did not effectively teach novice learners new words. The interview results aligned with this assumption; interactivity alone is insufficient without feedback mechanisms. Repeated exposure to new terms as well as feedback is critical for reinforcing learning.

The post-interviews revealed the importance of interactivity, feedback, and repetition in game-based vocabulary learning. While all participants reported increased interest in learning the Portuguese language with this novel technology, their emotional experiences varied based on the game's design. Participants grew frustrated when feedback and gameplay instructions were insufficient, emphasizing the need for a more cohesive design that balances VR engagement with language instruction, clarifies and refines

gameplay mechanics, and improves audio elements. Notably, Highrise hELLEp's combination of interactivity, sound effects, and repetition seems to have contributed to the highest vocabulary gains, suggesting that thoughtfully designed repetition can enhance retention. These findings underline the need for educational games to incorporate adaptive feedback, cohesive mechanics, and motivational elements to support diverse learners effectively [13, 19]. While some participants expressed concerns that these games educate primarily on short-term or temporary memory, it is logical to assume that as we implement these games into the contexts of semester-long language courses for which they were ultimately designed, these short-term gains, practiced over a longer period of time, may ultimately translate into more lasting L2 knowledge.

Though we gathered participant engagement data from the posttest as well as the think-aloud-protocol, it has not been fully analyzed yet. This data will provide additional insight as to specific mechanics, dynamics, and aesthetics of the games that participants enjoyed the most and least, further informing future work. This is an important next step, as learners who do not enjoy the game will not play it for very long or very often; the balance between learning and fun is something we continue to strive for in our work.

6 Conclusions and Future Work

The findings of this study suggest that even novice learners can increase their knowledge of L2 vocabulary from a VR game, as all three VR games showed learning gains. Highrise hELLEp was the most effective game for vocabulary acquisition, demonstrating the highest learning gains. A design element that may have factored into this result is the way that the game includes repeated exposure to terms. This aligns with established SLA principles, suggesting that repeated exposure is crucial for effective vocabulary acquisition.

The results also indicate that hELLE's Kitchen provided moderate learning gains, though not statistically significant, which aligns with the SLA research supporting the teaching of terms that have a shared context, such as Kang's [20] and Mulder et al.'s work on context [21], as well as Sykes and Reinhardt's [13] fourth recommendation for effective L2 learning games. hELLE's Kitchen places items in context but only displays a few related terms at one time. Recipes in the game include no more than five terms, and these are the only terms displayed until the player completes the recipe. Only then do they receive a new recipe with different terms. The prolonged exposure to a few terms at a time is worth additional study as well—it may be more effective for long term retention than repeated exposure to a longer list of new terms.

Spin 'N' SpELLE was the least effective, likely due to its focus on spelling without sufficient contextual support and lack of immediate feedback. Further investigation is needed to determine what elements of the game were beneficial and to participants with what kind of prior knowledge, if it might help learners with pronunciation, or those more advanced in their L2 studies who need to improve their spelling accuracy.

Analysis is still ongoing for the data collected during this study. Future work includes analyzing the engagement items on the posttest as well as the think-aloud-protocol recordings and post-interview recordings. This study is also limited, as it is isolated from other SLA contexts and limited to one session. The ways that these games can support long-term vocabulary retention is the subject of planned future studies in real classroom settings. Another related limitation lies in the restricted linguistic scope of the current research, as it primarily focuses on vocabulary gaining in SLA. To provide a more comprehensive evaluation of second language acquisition, future games and studies should incorporate additional skill areas—such as speaking, listening, and grammar—to capture a more holistic perspective of learners' linguistic development.

Additional work at the term level is also needed. On the pretest, 44 participants were able to correctly translate *a pasta de dentes* (toothpaste) as well as *a toalha* (towel). On the posttest, this increased to 102 participants for toothpaste but stayed at 44 for towel: two participants translated it incorrectly on the pretest but correctly on the posttest, and two other participants got it correct on the pretest but incorrect on the posttest. One of these participants played Spin 'N' SpELLE, and the other played Highrise hELLEp; both participants typed a different term that was included in their game.

This next phase of this study involves integrating these games in the context of a language course, where learners have been exposed to the vocabulary terms and the games perform the role of reinforcement will be additionally informative. We predict that in the context of a course with learners who are practicing vocabulary retention rather than initial learning of terms will perform better with all games.

Overall, participants increased vocabulary knowledge no matter which VR game they were assigned in the study. Participants who were able to translate the most words on the posttest played the game with the highest amount of exposure—seeing the text of the Portuguese terms the most often allowed participants to perform the best on the written posttest. Though some of the ELLE games include spoken audio that pronounces the terms in

the target language, neither speaking nor listening comprehension items were included on the posttest, which will also need to be investigated in future work.

Acknowledgements

The participant studies mentioned in this paper are funded by the U.S. Department of Education, PR/Award Number P017A230010. Opinions expressed do not represent the views of the U.S. Department of Education.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Zeng, J., Parks, S., Shang, J.: To learn scientifically, effectively, and enjoyably: A review of educational games. Human Behavior and Emerging Technologies. 2, 186–195 (2020). https://doi.org/10.1002/hbe2.188
2. Su, F., Zou, D.: Technology-enhanced collaborative language learning: theoretical foundations, technologies, and implications. Computer Assisted Language Learning. 35, 1754–1788 (2022). https://doi.org/10.1080/09588221.2020.1831545
3. Chapelle, C.A.: Computer Assisted Language Learning. The Handbook of Educational Linguistics. 585–595 (2008). https://doi.org/10.1002/9780470694138.ch41
4. Dhimolea, T.K., Kaplan-Rakowski, R., Lin, L.: A systematic review of research on high-immersion virtual reality for language learning. TechTrends. 66, 810–824 (2022).
5. Peterson, M.: Massively multiplayer online role-playing games as arenas for second language learning. Computer Assisted Language Learning. 23, 429–439 (2010). https://doi.org/10.1080/09588221.2010.520673
6. Peterson, M.: Learner interaction in a massively multiplayer online role playing game (MMORPG): A sociocultural discourse analysis*. ReCALL. 24, 361–380 (2012). https://doi.org/10.1017/S0958344012000195
7. Peterson, M.: The use of massively multiplayer online role-playing games in CALL: an analysis of research. Computer Assisted Language Learning. 29, 1181–1194 (2016). https://doi.org/10.1080/09588221.2016.1197949
8. Calvo-Ferrer, J.R., Belda-Medina, J.: The Effect of Multiplayer Video Games on Incidental and Intentional L2 Vocabulary Learning: The Case of Among Us. Multimodal Technologies and Interaction. 5, 80 (2021). https://doi.org/10.3390/mti5120080
9. deHaan, J., Reed, W.M., Kuwada, K.: The Effect of Interactivity with a Music Video Game on Second Language Vocabulary Recall. Language Learning & Technology. 14, 74–94 (2010).
10. Sullivan, P.M., Lantz, J.J., Sullivan, B.A. eds: Handbook of Research on Integrating Digital Technology with Literacy Pedagogies. IGI Global (2020). https://doi.org/10.4018/978-1-6684-3710-0.ch064
11. Li, J.: A Systematic Review of Video Games for Second Language Acquisition. In: Handbook of Research on Integrating Digital Technology With Literacy Pedagogies. pp. 472–499. IGI Global (2020). https://doi.org/10.4018/978-1-7998-0246-4.ch021
12. Alfadil, M.: Effectiveness of virtual reality game in foreign language vocabulary acquisition. Computers & Education. 153, 103893 (2020). https://doi.org/10.1016/j.compedu.2020.103893
13. Sykes, J.M., Reinhardt, J.: Language at Play: Digital Games in Second and Foreign Language Teaching and Learning. Pearson Education, Incorporated (2012).
14. Chen, Y.-L., Hsu, C.-C.: Self-regulated mobile game-based English learning in a virtual reality environment. Computers & Education. 154, 103910 (2020). https://doi.org/10.1016/j.compedu.2020.103910
15. Alvi, I.: Investigating students' adoption of virtual reality for L2-learning in India. Educ Inf Technol. 29, 8035–8056 (2024). https://doi.org/10.1007/s10639-023-12140-6
16. Thrasher, T.: The Impact of Virtual Reality on L2 French Learners' Language Anxiety and Oral Comprehensibility: An Exploratory Study. CALICO Journal. 39, 219–238 (2022). https://doi.org/10.1558/cj.42198
17. Kong, Y., Johnson, E.K.: Work-in-Progress—Cozy Games for Learning: Vocabulary Practice with Anim-ELLE Crossing. Immersive Learning Research - Academic. 253–257 (2024). https://doi.org/10.56198/U6C0WIZ5O
18. Johnson, E.K.: ELLE-ments of Learning: A Framework for Analyzing Multimodal Technical Communication Strategies in an Educational VR Game. In: 2021 7th International Conference of the Immersive Learning Research Network (iLRN). pp. 1–4 (2021). https://doi.org/10.23919/iLRN52045.2021.9459357
19. Nation, P.: Teaching and learning vocabulary. In: Handbook of Practical Second Language Teaching and Learning. Routledge (2022).
20. Kang, S.-H.: The effects of a context-embedded approach to second-language vocabulary learning. System. 23, 43–55 (1995). https://doi.org/10.1016/0346-251X(94)00051-7
21. Mulder, E., Ven, M.V.D., Segers, E., Verhoeven, L.: Context, word, and student predictors in second language vocabulary learning. Applied Psycholinguistics. 40, 137–166 (2019). https://doi.org/10.1017/S0142716418000504