Teaching Language and Culture with a Virtual Reality Game

Alan Cheng[1], Lei Yang[2], and Erik Andersen[1]

[1]Department of Computer Science, [2]Department of City and Regional Planning, Cornell University
{ayc48, ly292, ela63}@cornell.edu

ABSTRACT

Many people want to learn a language but find it difficult to stay engaged. Ideally, we would have language learning tools that can make language learning more enjoyable by simulating immersion in a foreign language environment. Therefore, we adapted *Crystallize*, a 3D video game for learning Japanese, so that it can be played in virtual reality with the Oculus Rift. Specifically, we explored whether we could leverage virtual reality technology to teach embodied cultural interaction, such as bowing in Japanese greetings. To evaluate the impact of our virtual reality game designs, we conducted a formative user study with 68 participants. We present results showing that the virtual reality design trained players how and when to bow, and that it increased participants' sense of involvement in Japanese culture. Our results suggest that virtual reality technology provides an opportunity to leverage culturally-relevant physical interaction, which can enhance the design of language learning technology and virtual reality games.

Author Keywords

language learning, video games, virtual reality

ACM Classification Keywords

H.5.0. Information Interfaces and Presentation: General

INTRODUCTION

Learning a second language is a goal shared by many, from children in bilingual environments to adult immigrants seeking employment to people wishing to travel abroad. However, many people find it difficult to stay engaged and also to learn vocabulary and grammar in context. Although many people learn languages through popular online tools like DuoLingo [36] and Rosetta Stone [26], much of the vibrancy of the real-world experiences that can make language learning fun and relevant do not translate to these systems.

On the other end of the spectrum, studying abroad provides ample opportunities to be exposed to a foreign language, assimilate into another culture, and apply language in context. Indeed, many studies have analyzed the positive effects of studying abroad on language acquisition [31, 15]. However, studying abroad is not an option for everyone due to financial cost, lack of opportunity, or insufficient time. Ideally, we

would have immersive language tools that simulate the experience of being in a foreign language environment as deeply as possible, so that an aspiring learner can learn both language and culture from observation, as well as harness their potential for motivation.

The emergence of high-quality, low-cost virtual reality headsets in recent years has sparked a surge in interest in the technology, including for educational purposes [17, 22]. Virtual reality has also been used for virtual tourism [16, 5] to situate people in other places without actually being present.

In this paper, we discuss our experiences in designing a virtual reality game for learning language. Building off of *Crystallize* [9, 10], an existing 3D video game for learning the Japanese language, we created a new version of this game that works with the Oculus Rift, a virtual reality headset system. In particular, we explored whether we could use virtual reality to design game mechanics around culturally-relevant embodied physical interaction, such as bowing in Japanese greetings.

To evaluate the effectiveness of our VR game design process, we conducted a formative user evaluation with 68 participants. Our results provide initial evidence that porting to VR and adding VR-specific game mechanics to *Crystallize* was useful for increasing involvement in Japanese culture and teaching players how to bow. However, players encountered some challenges, including feeling sick while using the virtual reality headset. The impact on learning itself was also inconclusive. Nevertheless, this formative evaluation indicates that we were able to leverage some benefits of VR, which will inform future development of the game. Furthermore, the integration of physical cultural artifacts like bowing has implications for the design of language learning technology and virtual reality games.

RELATED WORK

Virtual reality and education

Virtual reality has great potential for education across a wide spectrum of fields. One such field is surgical education. Currently, aspiring surgeons learn surgical skills in the operating room, and while this provides the most realistic environment, there is much left to be desired from a pedagogical perspective. The focus is on the patient rather than the learner, and steps in the surgery can not be adapted to the student's needs. However, a VR surgical environment would allow students to practice and learn without any danger to an actual patient. The VR environment can cater to students by allowing them to make mistakes, repeat steps, and pause as needed [17]. Furthermore, surgical competence can both be taught and measured automatically with VR software [22].

Construct3D is an augmented reality mathematics and geometry learning tool that allows students to draw and visualize virtual three-dimensional objects in real three-dimensional space. Student participants in a pilot study by Kaufmann et al. were able to learn how to use the VR tool quickly and thought that it created a good environment for experimentation, which suggests that tools like *Construct3D* can be effective supplements to a traditional classroom curriculum [19].

The sense of *presence* is the defining experience of virtual reality [28]. The main factors thought to influence presence are: high-resolution information, environmental consistency, possibility for navigation and interaction with other avatars, and similarity of a virtual avatar to one's own body [34]. Immersion and involvement are both necessary for experiencing presence, and presence leads to increased learning and performance [38]. Therefore, we focused our design of the virtual reality version of *Crystallize* around enabling presence.

Language learning tools

There are many computer-based tools for learning languages. MicroMandarin utilizes the user's real world location to suggest vocabulary words [14]. Tip Tap Tones trains users to recognize tones in Chinese [13]. Dearman et al. used a desktop wallpaper to teach vocabulary [12]. We build on these ideas and make them more engaging by creating a holistic experience in a 3D virtual reality video game.

Rosetta Stone [26] is a highly successful tool that mainly teaches language through a series of pictures. A typical task features a set of four or more pictures that each show a certain situation, such as a boy eating or a girl running, and asks the user to identify the picture that most closely matches a phrase in the target language. Rosetta Stone offers many advantages over traditional curricula: the learner receives immediate feedback, information is presented in a visual context, and meaning is often learned through inference. Our work builds on these ideas but also features an interactive 3D environment that provides a deep, situated, virtual reality context and prioritizes experimentation and completion of tasks and goals in that environment.

DuoLingo [36] is another successful tool that teaches language by asking the learner to translate sentences. DuoLingo also has a highly structured learning progression and has achievements and point systems to motivate users. Although 34 hours of DuoLingo has been shown to be equivalent to a first-semester college course in Spanish, many students stopped using it within 2 hours [35]. Our work also uses game-like elements to incentivize users, but it teaches linguistic concepts in a situated physical context rather than through pure translation.

There are some existing video games that focus on teaching languages. A good summary can be found in [29]. For example, *Sanjigenjiten* uses a 3D environment to teach vocabulary words [18]. This game makes use of visual context to teach vocabulary meanings, but it does not provide a deep learning progression. Zengo Sayu [25] is a virtual environment for learning Japanese in which the player explores a location and hears audio cues of words. However, it is not really a game.

Crystallize is a 3D language learning game that teaches language material in *its physical context*, motivated by the theories of situated cognition [8] and encoding specificity [32]. The game provides opportunities for players to learn words from context and gamifies language learning through a quest system that requires players to solve challenges that involve learning new words and using them to construct target sentences. The conceptual components of this game involve understanding specific vocabulary words and understanding grammatical constructions. Two evaluations of *Crystallize* have been conducted: a preliminary lab study with 42 participants [9] and an "in the wild" study with 186 participants from Reddit [10]. Both studies showed statistically significant learning gains.

Interaction design for virtual reality

Interaction in virtual reality is a major challenge. Early work [33] found that presence is better when users walk around in the real world (corresponding to movement in the virtual world), compared to walking in place or pressing a button to "fly" around. Our work attempts to create a similar sense of presence but with the act of bowing instead of walking.

Other work has focused on enabling high-fidelity interaction with virtual environments. Much of this work has focused on providing haptic feedback [4, 6] for hand motions. In this work, we explore the potential of a different kind of virtual reality input, which is to use head motions of the headset itself to enable culturally relevant physical input for a language learning scenario.

Recently, Tan et al. [30] studied virtual reality games and found that VR can heighten game experiences. They also found that although virtual reality can make players sick, they still had a positive reaction to the experience. In this work, our goal was to leverage some of the benefits of VR while abating negative impacts as much as possible.

CRYSTALLIZE IN VIRTUAL REALITY

Since a major goal of *Crystallize* is to gamify language learning by simulating immersion in a target language environment, ideally it should try to maximize the sense of immersion. Therefore, we extended *Crystallize* by adding virtual reality support using the Oculus Rift (Figure 2). Our porting of *Crystallize* to VR was primarily motivated by the following two-part design question: (1) How can we best leverage VR to improve a language learning game, and (2) Can VR help us "gamify" cultural interaction by introducing physical cultural artifacts (like bowing) into the user interface?

To explore the impact of VR on *Crystallize*, we created a new VR demo that is separate from the existing game. The port itself is fairly simple, with only minor changes. The camera angle was changed from third-person to first-person. The original UI has been mostly preserved, with some adjustments to make viewing it in VR more comfortable.

Creating a sense of presence

Some elements of presence were created from the virtual reality environment itself: the culturally-relevant background and the first-person perspective. Three of the factors described by

Figure 1. Overhead view of the *Crystallize* virtual reality scenario that we set up to test the ability of virtual reality to communicate cultural and language information. Players navigated through a Japanese teahouse and solved language learning questions by interacting with non-player-controlled characters.

Figure 2. Using Unity, we made *Crystallize* work with the Oculus Rift.

Figure 3. An example of an eavesdropping dialogue. The player can listen in on NPC dialogues and collect new words to add to the inventory.

Usoh [34] were present: ability to navigate, ability to interact with other virtual avatars, environmental consistency, and high-resolution graphics. However, there was no virtual avatar to represent the player (if the player looked down, he or she would not see a virtual representation of himself or herself).

We also increased the sense of presence through embodying physical actions where possible. Although overall navigation was done with the keyboard (to avoid needing to play the game in a large room where one can walk around freely; besides, the Oculus does not support full-motion VR), we did include some basic gestural information. In Japanese culture, one often bows to other people when meeting them in formal settings. This is called *o-jigi* (お辞儀). Therefore, the VR version of *Crystallize* was designed so that players were expected to bow

when meeting new people. The head-tracking of the Oculus Rift was set up to detect if the player actually bowed.

The demo itself

The demo takes place in a traditional Japanese teahouse, which presents Japanese featured elements, including sliding shoji doors made from wood and paper, wooden floor passages, a bamboo courtyard, Japanese paintings, and a tatami (bamboo floor mat) room, with NPCs scattered about (Figure 1).

The VR demo is a self-contained experience separate from the main *Crystallize* game, and it uses a subset of the mechanics present in *Crystallize* along with the VR integration. In terms of gameplay, it comprises eight dialogues with NPCs scattered throughout the teahouse. There are two types of dialogues, all conducted solely in Japanese: (1) eavesdropping dialogue, in which two NPCs speak to each other and the player must learn new words from their conversation by dragging those words

Figure 4. An example of a multiple-choice conversation dialogue in VR. The player must select an appropriate response based on what the NPC said.

Figure 5. An example of a sentence-construction conversation dialogue in VR. The player must arrange vocabulary words in the correct order in order to respond to the NPC.

into the inventory (Figure 3), and (2) conversation dialogue, in which the player engages a single NPC in conversation. In conversations, the player is prompted to select the appropriate response from multiple choices (Figure 4) and, in some cases, asked to construct a sentence given several vocabulary words (Figure 5). Also, the player is prompted to bow in response to an NPC's bow when engaged in conversation.

Gameplay progresses in a linear fashion. Whenever a conversation with an NPC is available, the conversation must be successfully completed before the player is allowed to move on to the next conversation. Occasionally the player will be lacking a word needed to complete a dialogue; in this case, the player needs to seek out the appropriate eavesdropping NPC group to obtain the vocabulary word. The series of dialogues, along with their categorization and required player actions to complete, is listed in Figure 6.

The vocabulary words selected were chosen to include set greeting phrases like *sayounara* (Goodbye.), as well as some words and grammatical particles that can be arranged to form sentences, such as *namae* (name), used in the sentences *o namae wa?* (Your name?) and *watashi no namae wa [name] desu* (My name is [name]). This mix of vocabulary types allowed us to provide sentence construction challenges to the player in addition to multiple choice prompts, as mentioned previously.

TEACHING PLAYERS TO BOW

In Japanese culture, greetings such as *konnichiwa* (Hello.) and *hajimemashite* (Nice to meet you.) are often accompanied with a small bow, and the other party responds with a similar greeting and bow. Knowing how to bow and, just as important, when to bow is a key part of communication in Japan, so we wanted to teach players by integrating this cultural behavior into *Crystallize*.

Bowing is one example of physical cultural behaviors, and mastering such cultural artifacts is an important aspect of becoming proficient in a foreign language. However, cultural competence has often not received as much attention as linguistic material [11, 37]. Therefore, we explored whether we could use virtual reality to integrate cultural artifacts into the experience. In particular, we tried to leverage the sense of presence that virtual reality can provide to enable a natural mechanism for nonverbal cultural interaction. Physicality plays a significant role in communication [7, 21] and there have been attempts in character animation to emulate it [20].

The simplest way to implement a bowing mechanic would involve the player clicking a button or pressing a key to perform a bow. However, with this control scheme, the player's action (clicking or pressing) does not capture the physicality of the bowing motion. Instead, we realized that the act of bowing could be naturally adapted to virtual reality—because players already know how to bow in real life, we can have players apply that knowledge in the cultural context of a Japanese conversation through our VR simulation.

Using the angle of the head mount, we attempted to detect when the player bowed. If the head was lowered by at least 22.5° and then raised above that threshold again, the motion registered as a successful bow. We experimented with various angles until we found one that minimized both discomfort and the chance of unintentionally triggering a bow. Ideally, we would have emulated the standard Japanese greeting bow (15° from the hip while standing), but since the Oculus can only track the head, a 22.5° bow of the head seems to be a reasonable approximation [27].

In the game demo, we teach the cultural practice of bowing by integrating it into NPC conversations at the appropriate times. Just like in real Japanese conversations, NPCs in the game may bow while saying greeting phrases. When this occurs, the player must bow in return in order to successfully continue the conversation.

Tutorial design

Our tutorial design for teaching bowing was motivated by two principles. First, there is empirical support for the common belief that text tutorials are not necessarily effective, particularly in situations where the player can learn through experimentation instead of reading [2]. Second, work in math education has found positive benefits of gradually fading out tutorials to enable a smooth transition from solely following instructions to performing the task without help [24, 3]. Therefore, the

Number	Type	Content
1	Eavesdropping	Learn *konnichiwa* (Hello.)
2	Eavesdropping	Learn *sayonara* (Goodbye.)
3	Conversation	Reply with "Hello." and learn to bow
4	Eavesdropping	Learn *hajimemashite* (Nice to meet you.)
5	Eavesdropping	Learn *o namae wa?* (Your name?), *[name] desu* (I'm [name].)
6	Conversation	Reply with "Nice to meet you.", bow, reply with "I'm [name].", learn *yoroshiku onegaishimasu* (Please take care of me.), bow
7	Eavesdropping	Learn *watashi no namae wa [name] desu.* (My name is [name])
8	Conversation	Bow*, reply with "Hello.", bow, reply with "My name is [name])
9*	Conversation	Reply with "Nice to meet you.", bow, reply with "Your name?", reply with

Figure 6. The list of in-game dialogues. Items marked with an asterisk (*) were added in a later version of the demo.

Figure 7. Using the Oculus Rift head tracking, we trained players to bow when greeting other characters in Japanese. A few seconds after the NPC bows, the player is given a prompt to bow. The third panel shows the player's field of view mid-bow.

bowing tutorials are minimal in text and gradually fade out over time.

When an NPC bows to the player in the game, a prompt with the words "Please bow." appear on the screen afterward. During the first instance, the prompt appears right away to instruct the player when to bow. From then on, the game lengthens the time between the NPC's bow and the message prompting the player to bow. The second time, the game waits one second before prompting. The third time, two seconds; the fourth and each subsequent time after that, four seconds. If the player completes a bow before the prompt appears, that suggests the player both acknowledged the NPC's bow and knew to execute the correct response by bowing in return, therefore signaling that the player has most likely learned the desired behavior. The sequence of events involved in a bowing exchange is portrayed in Figure 7.

FORMATIVE USER STUDY
We conducted a formative user study in order to gain initial insights on the design of our VR port of *Crystallize*. This study focused on evaluating the following questions: (1) Does the VR version of the game improve language acquisition and stimulate interest in the language's culture more effectively than the non-VR version, and (2) Could players can be taught cultural behaviors, like bowing, through the help of VR?

To test these hypotheses, we conducted a study with a total of 68 participants recruited from the university community through events. The only requirement to participate in the study was a self-reported lack of familiarity with Japanese language and culture. This study took place over two weeks.

To evaluate the impact of the addition of virtual reality and the bowing mechanic on the effectiveness of the game design, we created two versions of the demo. One was played without virtual reality on a regular computer screen. The other was played with the Oculus Rift head-mounted display. The VR and non-VR versions were the same except that the VR version included bowing.

Experimental protocol
Participants were informed that they would be playing a video game for learning Japanese. Participants then completed a pre-game survey in which to provide some demographic information, such as gender and ethnicity, state their level of interest in Japanese culture and language, and take a pretest for Japanese vocabulary proficiency. The pretest involved matching Japanese words (that appear in the demo) to their English definitions.

To evaluate the impact of adding virtual reality (and bowing), we were primarily interested in players' first experience with the game. Therefore, players were randomly assigned to play either the VR or non-VR version. Participants were given basic instructions on how to play the game, such as how to interact with in-game characters by clicking on them. Each play session took approximately 15-20 minutes. Participants then completed a post-game survey in which they completed a vocabulary posttest that was identical to the pretest except for

item order. Participants then compared their level of interest in Japanese culture and language to before.

Next, in order to collect additional data on bowing and reactions to the two versions, participants were asked to play the version they did not play first. Then they completed a second post-game survey in which they provided additional feedback on the experience. In addition to the surveys, we recorded bowing data, including how long it took for the player to bow after the NPC bowed, whether the bow was performed before the prompt, and the angle of the bow.

Towards the end of the data collection, we added a ninth dialogue and an extra bow in the eighth dialogue (in VR), as shown in Figure 6. 15 of the 68 participants played these updated VR and non-VR versions. We analyzed all 68 participants in one group because we did not think that this change significantly affected players' reactions to the VR and non-VR versions.

RESULTS
We analyzed (1) whether players learned cultural information through the VR bowing mechanic, (2) how players reacted to the VR and non-VR versions, (3) negative effects of VR, (4) players' overall reactions to the experience, and (5) the impact of the VR demo on learning.

The VR demo trained players how to bow
As mentioned previously, participants were expected to bow in response to an in-game NPC bowing, with a textual prompt appearing on-screen if the participant does not bow within a few seconds. We define a "unprompted bow" as one that is performed before the prompt appears. If the participant performs an unprompted bow, it is likely that he/she recognized, without being explicitly reminded, that the appropriate cultural response to the other party bowing is to bow in return.

Aside from the first time bowing is introduced, where the prompt appears immediately, there are three bowing opportunities in the VR version of the game. We found that 34 out of the 68 participants (50.0%) performed at least one unprompted bow, which suggests that players were able to learn when to bow.

Overall, the average bow angle that players reached was $45.8°$. This angle is deeper than the culturally-appropriate standing bow angle of $15°$ and the game-specific detection angle of $22.5°$. This indicates that although the game appears to be training players when to bow, it may be necessary to provide additional angular feedback and refine the bow detection angle so that players learn the most appropriate angle. Furthermore, future work must verify experimentally that the bowing knowledge that players acquire transfers to real conversations in a way that native speakers would find appropriate.

The VR demo led to increased involvement in culture
Although all participants played both the VR and non-VR versions, the surveys given after playing each demo consisted of different questions. Therefore, we analyzed the impact of the demo version in a between-subjects manner because these were not repeated measures. Since the survey questions asked

for Likert-style responses, we used the non-parametric Mann-Whitney U-test to analyze the differences between the VR and non-VR versions. These results can been seen in Figure 8.

We found a statistically significant effect for the impact of the VR version on perceived involvement in Japanese culture. We found that players of the VR demo felt more involved in Japanese culture ($M = 3.18, SD = 1.03$) than the non-VR players ($M = 2.47, SD = 0.86$), $Z = 2.70, p = 0.007$.

VR made people dizzy
The virtual reality experience suffered from some technological and interface defects. According to the user experience feedback, 23.5% (16/68) of the comments expressed negative feelings such as "dizziness" and "sickness", and 27.9% (19/68) of the comments revealed user interface problems such as difficulty reading words in the inventory toward the bottom of the UI plane. These issues influenced the participants' self-evaluations.

In other words, despite our efforts to adapt the original UI of *Crystallize* to work similarly in VR, users still often struggled with reading and interacting with the UI in a VR environment. One participant wrote that "the combination of controlling the UI with the mouse felt disorientating", and another commented, "Having to click and drag a menu at the bottom of the screen wasn't particularly intuitive in the VR version", suggesting that trying to retain the original drag-and-drop UI in the VR version of the game may not have been the ideal approach for a natural VR experience. Our main takeaway here is that VR necessitates more significant changes, primarily to enable movement without sickness and user input without relying on the keyboard and mouse, which are blocked by the headset.

In addition, since the experiment was designed to complete within 30 minutes, no more than two questions were devoted to a particular research question. More questions need to be added and further debriefing is needed in the future.

Participants enjoyed the VR experience
While many participants felt motion sickness and dizziness while playing, a majority of participants reacted positively to the experience: 58.8% (40/68) of participants wrote positive comments towards using VR to interact with cultural context. Not only did 17.6% (12/68) of participants report feeling more immersed when using VR compared to non-VR, but some participants also noted that they felt more involved in the game and its world. One wrote that he/she felt "more involved within the game world, more connected to the people I was speaking to", while another noted, "I felt like at least I was 'looking' at a person at eye-level, and it was fun to 'talk' with them!" This feedback reinforces the idea that, on top of situating the user in a cultural context, VR can also help users connect with that culture and the people within it.

Furthermore, participants found bowing in VR to be a unique and enjoyable experience, with 17.6% (12/68) praising the bowing feature in particular as part of their feedback. However, one participant did express some confusion as to when bowing was supposed to occur ("[...] it seemed to happen at odd

Question	Condition	Size	Mean	SD	Results
Compared to before, how interested are you in the Japanese language now?	VR	34	3.65	0.69	$Z = 1.28$
	Non-VR	34	3.35	0.85	$p = 0.20$
Compared to before, how much are you interested in Japanese culture now?	VR	34	3.41	0.67	$Z = 1.55$
	Non-VR	34	3.09	0.67	$p = 0.12$
How interesting was the content presented in the game?	VR	34	3.29	0.87	$Z = 1.87$
	Non-VR	34	2.88	0.81	$p = 0.06$
How much attention did you pay to the content presented in the game?	VR	34	3.50	0.96	$Z = 1.19$
	Non-VR	34	3.18	0.97	$p = 0.23$
How much did you feel involved in Japanese culture while playing the game?	VR	34	3.18	1.03	$Z = 2.70$
	Non-VR	34	2.47	0.86	$p = 0.007$

Figure 8. Analysis of survey questions. Response options used a Likert scale, spanning from 1 (None/not at all) to 5 (Very much/a great deal).

moments between snippets of conversation."), which suggests that there is still room in the game to teach the proper contexts in which to bow in Japanese culture. An interesting way to fill in this gap would be to allow players to initiate bows to NPCs on their own volition and provide feedback through the NPCs as to whether the bows were performed at an appropriate time.

Impact of the VR demo on learning was inconclusive

To measure language acquisition, we computed the difference in score between the pre-test and the post-test as the number of words learned for each participant. Both tests contained the same eight vocabulary words and had the player match each Japanese word to its equivalent, though the ordering of words was different for the two tests.

We used an unpaired t-test to analyze the number of words learned between participants in the VR and non-VR conditions. Though participants had self-reported a lack of familiarity with the Japanese language, eight participants scored a perfect 8 out of 8 on the pre-test. For this analysis, we omit those participants, as well as one participant that scored 3 points lower on the post-test compared to the pre-test.

Data	Condition	Size	Mean	Results
Words learned	VR	31	4.77	$t(57) = 0.889$
	Non-VR	28	5.39	$p = 0.377$

The number of words learned by participants in the VR condition ($M = 4.77, SD = 2.77$) turned out to be lower than that of those in Non-VR ($M = 5.39, SD = 2.56$), though the difference was not statistically significant. We suspect that the user interaction issues and the cognitive overhead in adapting to the unfamiliar VR interface and mastering the UI may have contributed to less effective learning [23].

Design implications

Our results provide some evidence that players can learn about culture and cultural behaviors, such as bowing, through immersive VR environments that are designed to enable them. They show that virtual reality can provide the basis for game designs that stimulate virtual participation and feelings of involvement in a foreign culture.

Most importantly, they suggest that the sense of immersive presence provided by VR yields a valuable opportunity to open up additional input channels for meaningful physical

interaction. This is particularly useful for games and simulations for language learning due to the importance of physical interaction. After viewing an instruction like "please bow," players could figure out what to do fairly naturally because it corresponded to an embodied action. Hopefully, virtual reality technology (such as the HTC Vive hand-held controllers) will become useful for training other kinds of culturally-grounded physical interaction, such as gestures and posture.

Limitations

We acknowledge the presence of multiple confounding factors, such as that bowing existed only in the virtual reality version. Therefore, we analyzed the results of this experiment in order to evaluate the design of the virtual reality game, and not to draw general conclusions about the impact of virtual reality. Furthermore, the study may have conflated players' natural enthusiasm for VR with their responses to the game experience. However, we believe this weakness can also be viewed as a strength: since motivation is an issue in language learning, harnessing natural enthusiasm for VR seems beneficial.

CONCLUSIONS

We adapted *Crystallize*, a 3D language-learning game, to virtual reality by adding integration with the Oculus Rift. In particular, we explored whether we could use virtual reality to design game mechanics around culturally-relevant embodied physical interaction, such as bowing in Japanese greetings.

We conducted a formative user study to evaluate the design of the ported game. Through adding VR and bowing, we observed a statistically significant increase in the participants' sense of cultural involvement. However, there is no obvious evidence that the language learning outcomes improved thus far. As for the usability of the VR game, participants were able to play through the entire demo and enjoy the experience, despite the problems introduced by using VR with the Oculus (motion sickness and inability to see the keyboard). This suggests that our fairly direct port of *Crystallize* was usable.

In future work, we plan to take steps to increase the sense of immersion further, such as exploring speech recognition as a primary input mechanism. Speech input would also have the side effect of reducing dependence on UI, which we expect will decrease player confusion with interacting with the UI in VR. We also hope to release both the VR and non-VR versions of the game online to gather longitudinal data and investigate whether participants would return to the game.

As 3D reconstruction technology [1] improves, this provides additional opportunities to feature *real* locations for language learning exercises. We plan to investigate whether this will increase engagement, as well as develop techniques to "gamify" these locations by adding non-player-controlled characters and interactive situations.

REFERENCES

1. Sameer Agarwal, Noah Snavely, Ian Simon, Steven M Seitz, and Richard Szeliski. 2009. Building rome in a day. In *2009 IEEE 12th international conference on computer vision*. IEEE, 72–79.

2. Erik Andersen, Eleanor O'Rourke, Yun-En Liu, Rich Snider, Jeff Lowdermilk, David Truong, Seth Cooper, and Zoran Popovic. 2012. The impact of tutorials on games of varying complexity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 59–68.

3. Robert K Atkinson, Alexander Renkl, and Mary Margaret Merrill. 2003. Transitioning From Studying Examples to Solving Problems: Effects of Self-Explanation Prompts and Fading Worked-Out Steps. *Journal of Educational Psychology* 95, 4 (2003), 774.

4. Mahdi Azmandian, Mark Hancock, Hrvoje Benko, Eyal Ofek, and Andrew D Wilson. 2016. Haptic Retargeting: Dynamic Repurposing of Passive Haptics for Enhanced Virtual Reality Experiences. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*. ACM, 1968–1979.

5. BBC. 2015. British Museum offers virtual reality tour of Bronze Age. `http://www.bbc.com/news/technology-33772694`, *BBC News* (2015).

6. Hrvoje Benko, Christian Holz, Mike Sinclair, and Eyal Ofek. 2016. NormalTouch and TextureTouch: High-fidelity 3D Haptic Shape Rendering on Handheld Virtual Reality Controllers. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*. ACM, 717–728.

7. Ray L Birdwhistell. 1952. *Introduction to kinesics*. University of Louisville.

8. John Seely Brown, Allan Collins, and Paul Duguid. 1989. Situated cognition and the culture of learning. *Educational researcher* 18, 1 (1989), 32–42.

9. Gabriel Culbertson, Erik Andersen, Walker White, Daniel Zhang, and Malte Jung. Crystallize: An Immersive, Collaborative Game for Second Language Learning. In *CSCW 2016*.

10. Gabriel Culbertson, Shiyu Wang, Malte Jung, and Erik Andersen. Crystallize: An Immersive, Collaborative Game for Second Language Learning. In *CHI 2016*.

11. Louise Damen. 1987. *Culture learning: The fifth dimension in the language classroom*. Vol. 11478. Addison Wesley Publishing Company.

12. David Dearman and Khai Truong. 2012. Evaluating the implicit acquisition of second language vocabulary using a live wallpaper. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 1391–1400.

13. Darren Edge, Kai-Yin Cheng, Michael Whitney, Yao Qian, Zhijie Yan, and Frank Soong. 2012. Tip tap tones: mobile microtraining of mandarin sounds. In *Proceedings of the 14th international conference on Human-computer interaction with mobile devices and services*. ACM, 427–430.

14. Darren Edge, Elly Searle, Kevin Chiu, Jing Zhao, and James A Landay. 2011. MicroMandarin: mobile language learning in context. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 3169–3178.

15. Barbara F Freed. 1995. *Second language acquisition in a study abroad context*. Vol. 9. John Benjamins Publishing.

16. Daniel A Guttentag. 2010. Virtual reality: Applications and implications for tourism. *Tourism Management* 31, 5 (2010), 637–651.

17. Randy S Haluck and Thomas M Krummel. 2000. Computers and virtual reality for surgical education in the 21st century. *Archives of surgery* 135, 7 (2000), 786–792.

18. Robert Howland, Sachi Urano, and Junichi Hoshino. 2012. SanjigenJiten: computer assisted language learning system within a 3d game environment. In *Advances in Computer Entertainment*. Springer, 262–273.

19. Hannes Kaufmann, Dieter Schmalstieg, and Michael Wagner. 2000. Construct3D: a virtual reality application for mathematics and geometry education. *Education and information technologies* 5, 4 (2000), 263–276.

20. Sergey Levine, Philipp Krähenbühl, Sebastian Thrun, and Vladlen Koltun. 2010. Gesture controllers. In *ACM Transactions on Graphics (TOG)*, Vol. 29. ACM, 124.

21. David McNeill. 1992. *Hand and mind: What gestures reveal about thought*. University of Chicago press.

22. David Ota, Bowen Loftin, Tim Saito, Robert Lea, and James Keller. 1995. Virtual reality in surgical education. *Computers in Biology and Medicine* 25, 2 (1995), 127–137.

23. Joseph Psotka. 1995. Immersive training systems: Virtual reality and education and training. *Instructional science* 23, 5-6 (1995), 405–431.

24. Alexander Renkl, Robert K Atkinson, Uwe H Maier, and Richard Staley. 2002. From example study to problem solving: Smooth transitions help learning. *The Journal of Experimental Education* 70, 4 (2002), 293–315.

25. Howard Rose and Mark Billinghurst. 1995. Zengo Sayu: An immersive educational environment for learning Japanese. *University of Washington, Human Interface Technology Laboratory, Report No. r-95-4* (1995).

26. Rosetta Stone. Rosetta Stone.
`http://www.rosettastone.com/`. (????).

27. H Shibata, J Takahashi, and J Gyoba. 2015. [Subjective impressions of bowing actions and their appropriateness in specific social contexts]. *Shinrigaku kenkyu: The Japanese journal of psychology* 85, 6 (2015), 571–578.

28. Jonathan Steuer. 1992. Defining virtual reality: Dimensions determining telepresence. *Journal of communication* 42, 4 (1992), 73–93.

29. Julie E Sykes, Jonathon Reinhardt, Judith E Liskin-Gasparro, and Manel Lacorte. 2012. *Language at play: Digital games in second and foreign language teaching and learning*. Pearson Higher Ed.

30. Chek Tien Tan, Tuck Wah Leong, Songjia Shen, Christopher Dubravs, and Chen Si. 2015. Exploring Gameplay Experiences on the Oculus Rift. In *Proceedings of the 2015 Annual Symposium on Computer-Human Interaction in Play*. ACM, 253–263.

31. Koichi Tanaka and Rod Ellis. 2003. Study abroad, language proficiency, and learner beliefs about language learning. *JALT journal* 25, 1 (2003), 63–85.

32. Endel Tulving and Donald M Thomson. 1973. Encoding specificity and retrieval processes in episodic memory. *Psychological review* 80, 5 (1973), 352.

33. Martin Usoh, Kevin Arthur, Mary C Whitton, Rui Bastos, Anthony Steed, Mel Slater, and Frederick P Brooks Jr. 1999. Walking> walking-in-place> flying, in virtual environments. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*. ACM Press/Addison-Wesley Publishing Co., 359–364.

34. Martin Usoh, Ernest Catena, Sima Arman, and Mel Slater. 2000. Using presence questionnaires in reality. *Presence: Teleoperators and Virtual Environments* 9, 5 (2000), 497–503.

35. Roumen Vesselinov and John Grego. 2012. Duolingo effectiveness study. *City University of New York, USA* (2012).

36. Luis von Ahn. 2013. Duolingo: learn a language for free while helping to translate the web. In *Proceedings of the 2013 international conference on Intelligent user interfaces*. ACM, 1–2.

37. Paige D Ware and Claire Kramsch. 2005. Toward an intercultural stance: Teaching German and English through telecollaboration. *The Modern Language Journal* 89, 2 (2005), 190–205.

38. Bob G Witmer and Michael J Singer. 1998. Measuring presence in virtual environments: A presence questionnaire. *Presence: Teleoperators and virtual environments* 7, 3 (1998), 225–240.