Improving Language Learning by an Interact-to-Learn Desktop VR Application: A Case Study with Peinture

Xiao Liu* Shaanxi Normal University, China
Shuwei Zhang† Shaanxi Normal University, China
Tao Xu‡ Northwestern Polytechnical University, China
Yun Zhou§ Shaanxi Normal University, China

2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW) | 978-1-6654-8402-2/22/$31.00 ©2022 IEEE | DOI: 10.1109/VRW55335.2022.00061

Abstract

Mastering vocabulary is a daunting task for second language learners. Desktop Virtual Reality (VR) can inspire and engage learners with interact-to-learn activities to facilitate memorizing vocabulary. This paper proposes a desktop VR application for color vocabulary learning called Peinture. Peinture consists of a lecturing video and an interactive coloring virtual space. The talking teacher in the lecturing video is generated automatically through text-to-speech and lip synthesis approaches. Learners can paint and interact in a theme when or after lecture learning. A preliminary user study based on performance, self-report and eye tracking data analysis shows that interact-to-learn activities provided by Peinture were beneficial for second language learning. However, there is still considerable room to upgrade such educational VR.

Keywords: Educational VR, Second Language Learning, Interact-to-Learn, Eye Tracking.

Index Terms: Human-centered computing—Virtual Reality

1 Introduction

Mastering vocabulary is a daunting task for second language learners. They spend considerable time understanding and memorizing English words. Evidence from embodied cognition research [6] on language processing suggests that the learner's bodily sensations and actions can enhance comprehension. With rich operations and context, desktop Virtual Reality (VR) may offer learners an embodied learning experience to facilitate memorizing vocabulary. Desktop VR is a type of VR that provides less immersion [9, 12] but can be used on a large scale for education. In desktop VR, users interact with 3D virtual scenes and objects via mouse or keyboard as common input [7]. A recent meta-analysis examined the effect of desktop VR applications and reported a positive learning effect. Some empirical studies also proved that well-designed non-immersive desktop VRs could be highly conducive to learning and training [4, 7].

This paper proposes a desktop VR application for color vocabulary learning called Peinture. The purpose of Peinture is to inspire and engage learners with interact-to-learn activities to build their vocabulary of color. Peinture consists of a lecturing video and an interactive coloring virtual space. The talking teacher in the lecturing video is generated automatically through text-to-speech and lip synthesis approaches. Learners can paint and interact in a theme when or after lecture learning. A preliminary user study based on performance, self-reports, and eye tracking data analysis showed that learners succeed in their English Language task performance and gain a good learning experience via interact-to-learn activities with Peinture.

*e-mail: 200462@snnu.edu.cn
†e-mail: shuwei@snnu.edu.cn
‡e-mail: xutao@nwpu.edu.cn
§e-mail: zhouyun@snnu.edu.cn

2 Peinture and Interact-to-Learn Activities

We propose a desktop VR application for vocabulary learning called Peinture. The purpose of Peinture is to inspire and engage second language learners with interact-to-learn activities to build their vocabulary of colors. We create Peinture with Unity. Most of the 3D models in Peinture were found in the Unity Asset store. Peinture consists of a lecturing video on the left and an interactive 3D space for painting on the right (see Fig. 1). The size of the lecturing area is 768*1080, and the interactive area is 1152*1080.

The lecturing video consists of visual materials on the left and a talking teacher on the right. This talking instructor explained each English word and illustrated its origins, usage, and examples. Visual materials offer the key points and illustrations of the color. The video of the instructor is not prerecorded but generated automatically by an engine proposed in [11], based on text-to-speech and lip synthesis approaches. This engine generates videos by importing merely an image or a clip of a human and a text. The learner can control the play of the video through the UI.

In the interactive area, learners can paint the objects when or after lecture watching. Pressing the buttons of the color activate color selecting and word pronouncing. When learners press the keys of W, S, A, D or ↑, ↓, ←, → on the keyboard, they can move to four directions: front, back, left, and right in the scene, from a first-person perspective. Additionally, they can use the mouse to paint and right drag-n-drop to rotate the object.

3 User Study

This user study evaluated the role of a desktop VR application called Peinture in improving learning outcomes and experience. We adopted the cognitive load theory [10], embodied cognition theory [6] and split-attention effect [1] to inform the hypotheses about interact-to-learn with a desktop VR. Overall, we assume that VR interaction causes a higher cognitive load but leads to more learning.

- **H1** Participants in the Peinture condition will gain a higher knowledge comprehension compared to the non-VR condition.
- **H2** Participants in the Peinture condition will report a higher cognitive load compared to the non-VR condition.
- **H3** Participants in the Peinture condition will split attention in learning under the Peinture condition.

3.1 Participants and Experimental Conditions

We recruited 6 Chinese graduate students (6 females) in this user study, and English is their second language. Their ages ranged from 22 to 23 years old (average = 22.50, SD = 0.55). Since tasks and tests involved colors, all participants completed a color vision test before the experiment. All participants saw colors in a normal way.

This study employed a within-group design. All participants took part in two conditions: Peinture and non-VR video conditions. In the Peinture condition, each participant was asked to learn five English words about color by watching a video lecture and painting objects in a scene in line with the theme of the words in desktop VR. In the non-VR video condition, each participant was asked to learn another five

978-1-6654-8402-2/22/$31.00 ©2022 IEEE
DOI 10.1109/VRW55335.2022.00061

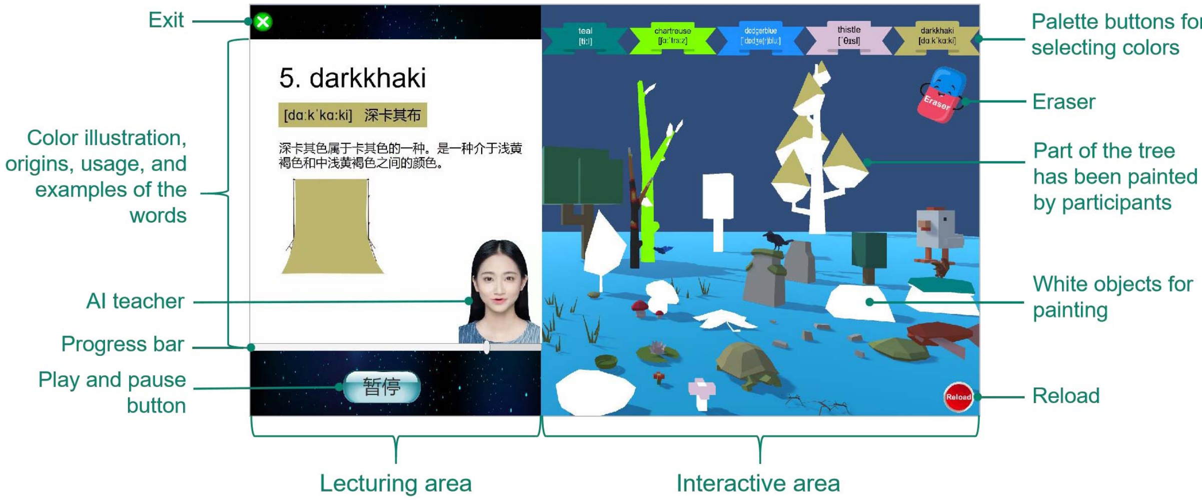

Figure 1: The interface of Peinture, with a lecturing area for obtaining knowledge (left) and a painting area for interacting and practicing (right).

color words different from the Peinture condition by only watching a video lecture. An AI-generated talking instructor taught the words in each video, explained the meaning, and illustrated some examples. The order of the two conditions was counterbalanced with a 2*2 Latin square.

3.2 Measures

To compare two conditions, we used the measures of learning outcome, cognitive load, and attention. We also collected the eye tracking data of Peinture to study how participants assign their visual attention. Finally, we evaluated Peinture through self-report rating using seven-level Likert scales.

To measure learning outcomes, we developed three tests for two conditions. Participants were required to determine the color bar for each color name (the word) in test one. In test two, learners needed to select which color appears in the given image. In test three, participants should choose the appropriate object that the given color can paint from four choices. For each test, 1 point would be added for the correct answer; otherwise, 0 point would be given. Thus, the total scores of the learn performance tests were 15. Concerning cognitive load, we collected intrinsic and extraneous cognitive load [2]. We also asked participants to report the attention levels of their learning.

To analyze how participants allocated visual attention to interactive and video lecture areas, we used a Tobii Nano eye tracker with a sampling rate of 60 Hz. Before starting learning with Peinture, each participant performed a gaze calibration task by Tobii. We collected real-time gaze data (i.e., fixations and saccades) during the learning experience of using Peinture. The measures include total fixation duration, total fixation count, saccade duration, and saccade count. These measures were collected for two AOIs (Area of Interests): the interactive and lecturing areas. The definitions of measures are as follows [5]:

- **Total fixation duration** refers to the total time spent on fixations.
- **Total fixation count** refers to the total number of fixations counted in an AOI.
- **Saccade duration** refers to the sum of saccadic time spent within an AOI.
- **Saccade count** refers to the total number of saccades counted within an AOI.

To evaluate the system and interaction of Peinture, we asked participants to rate the perceived usefulness, perceived ease of use, satisfaction, user experience, and learning experience of the Peinture.

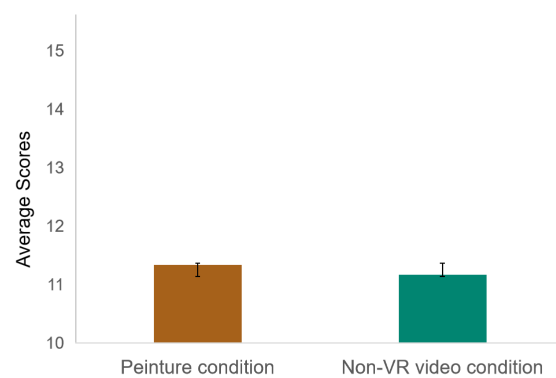

Figure 2: Learning performance.

3.3 Procedure

To ensure that participants did not know the words of color vocabulary in two conditions, we asked them to identify whether they mastered the words to be learned before the experiment. Once they had no prior knowledge of words, they could participate in the experiment. Participants first signed the informed consent form and filled out the questionnaire of basic information. Then they learned with Peinture or non-VR video in a counterbalanced order. The eye movements were collected in real-time. Participants answered the tests and filled out the tests and scales under each condition, respectively.

4 Results

Since the sample size in this preliminary work is not big, we present the results using descriptive statistics. As shown in Fig. 2, the overall performance of tests under Peinture condition (M = 11.33, SD = 1.82) was slightly higher than that of the non-VR condition (M = 11.16, SD = 3.37).

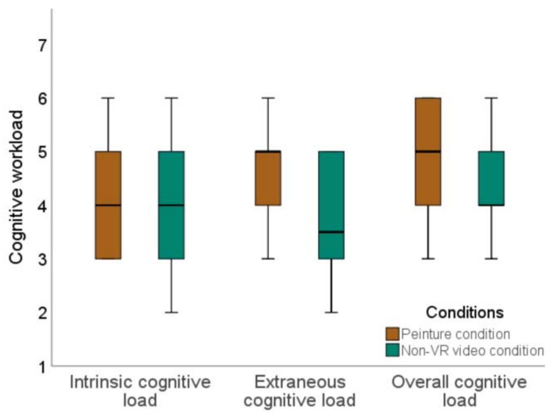

Figure 3: Cognitive workload.

Participants reported a higher intrinsic, extraneous, and overall cognitive load under the desktop VR condition compared with the non-VR condition (see Fig. 3). The medians of cognitive load levels under the desktop VR condition were all above 4.

As shown in Fig. 4, compared with the VR condition (Median = 4), participants were more able to concentrate when studying under the non-VR condition (Median = 4.5).

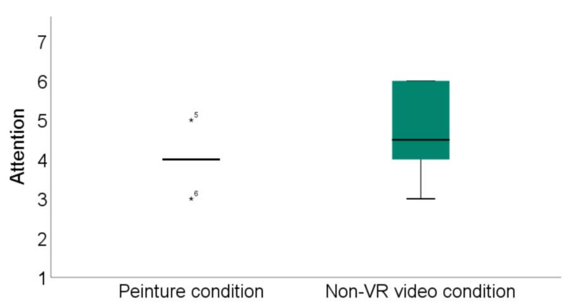

Figure 4: Attention.

Table 1 shows the results of eye tracking measures. Participants fixated longer and fixated more time in the interactive area compared with the video area. They also revisited the interactive area more. These results illustrate that participants spent more time interacting than watching the lecturing video to learn.

Participants were asked to rate perceived usability, ease of use, satisfaction, user experience, self-report of the learning experience of Peinture. As shown in Fig. 5, Peinture received all median above 4. Participants thought it interesting when interacting to learn.

5 Discussion

From the user study, we found that the performance, cognitive load and attention, and eye tracking results were in line with our hypothesis. First, participants gained slightly higher scores on the knowledge tests after interacting and learning with the desktop VR. Based on the current findings, we conclude that learning with the interaction in desktop VR might improve the performance of second language learning. This can be explained by embodied cognition theory that interaction enhances understanding while participants are involved in the interact-to-learn activities in Peinture. However, more data are required to collect to verify. Second, participants experienced a higher intrinsic, extraneous, and overall cognitive load under the desktop VR condition compared with the video condition. This is in line with the previous findings in VR for educational purposes [8], which indicated that such learning environments resulted in increased cognitive load. More interaction requires participants to use more cognitive resources to process information. Third, the self-report results of attention showed that participants split their attention under the Peinture condition. This finding was in line with the results of eye tracking data that indicated that participants put more attention into interactive learning than video watching. For the non-VR video condition, participants only needed to focus on the video, absorbing information through visual and auditory channels. However, when participants interacted to learn, they activated visual, auditory, and motor channels, leading to decreased control of attention.

We also found that learners fixated longer, fixated more time, and revisited more concerning the interactive area. When learners process one piece of information, they will look at it until they integrate knowledge into their cognitive structure or solve the problem. Thus, Longer fixation time indicates deeper processing, as well as higher fixation count and revisits indicate attention to the AOIs [5]. Thus, participants allocated their attention to the interaction while learning. Besides, participants rated Peinture with a high score from the perspective of perceived usefulness, perceived ease of use, satisfaction, user experience, and learning experience. The results showed that participants found Peinture interesting and liked this desktop application. Additionally, they thought that the interactive environment that Peinture offered could enhance understanding and memorizing English words.

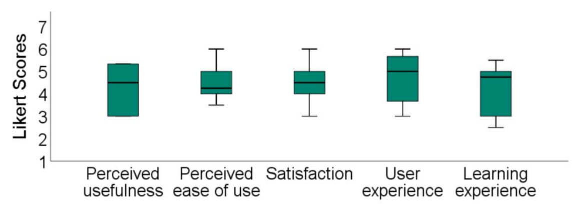

Figure 5: The boxplot of perceived usability, ease of use, satisfaction, user experience, self-report of learning experience of Peinture.

Table 1: Results of eye tracking measures.

Metrics	Interactive area		Video area	
	M	SD	M	SD
Total fixation duration	103.49	21.25	64.47	38.55
Total fixation count	293.17	104.54	233.50	138.93
Saccade duration	126.43	29.48	81.14	51.20
Saccade count	32.00	7.40	30.33	6.53

6 Conclusion, Limitations, and Future Work

In this study, we propose Peinture, a desktop VR providing interact-to-learn activities for second language learning. Despite the small sample size in this work-in-progress, we still benefited from obtaining empirical information. The user study's performance, usability, and user experience suggested that interact-to-learn activities provided by Peinture were beneficial for second language learning. However, cognitive load, attention, and eye tracking showed that providing interaction and instruction simultaneously may split attention and lead to overloading the learner's visual, auditory, and motor channels.

However, this is a work-in-progress, and we still have a long way to go. We will continue to move along this line and conduct an extensive study with a significant sample size in our future work. First, previous studies [3] offered more evidence on the presentation of written text and graphics, or text and animation, in multimedia learning. However, investigation and evidence on presenting interaction and instruction in a VR learning environment are still lacking. We will focus on how to present interaction and instruction spatially and temporally in educational VR environment. Integrating interaction and instruction may be a solution. Second, based on a short interview of participants, we found that the learning would be impacted not only by the design of Peinture but also by learners' multitasking attitudes and abilities. This interact-to-learn application is beneficial for learning when the learner is good at multitasking; otherwise, it would confuse single-tasker and then detriment learning. Thus, we will investigate how such desktop VR support learning considering learning style.

ACKNOWLEDGMENTS

This work was supported by projects of the National Natural Science Foundation of China (62077036), and the national key research and development program of China under grant No.: 2018AAA0100501.

REFERENCES

[1] P. CHANDLER and J. SWELLER. The split-attention effect as a factor in the design of instruction. *British Journal of Educational Psychology*, 62(2), 1992. doi:10.1111/j.2044-8279.1992.tb01017.x.

[2] G. Cierniak, K. Scheiter, and P. Gerjets. Explaining the split-attention effect: Is the reduction of extraneous cognitive load accompanied by an increase in germane cognitive load? *Computers in Human Behavior*, 25(2):315–324, 2009. doi:10.1016/j.chb.2008.12.020.

[3] R. C. Clark, R. E. Mayer, and W. Thalheimer. E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning, 2016.

[4] I. Dubovi, S. T. Levy, and E. Dagan. Now I know how! The learning process of medication administration among nursing students with non-immersive desktop virtual reality simulation. *Computers & Education*, 113:16–27, Oct. 2017. doi:10.1016/j.compedu.2017.05.009.

[5] M.-L. Lai, M.-J. Tsai, F.-Y. Yang, C.-Y. Hsu, T.-C. Liu, S. W.-Y. Lee, M.-H. Lee, G.-L. Chiou, J.-C. Liang, and C.-C. Tsai. A review of using eye-tracking technology in exploring learning from 2000 to 2012. *Educational Research Review*, pp. 90–115, Dec. 2013. doi:10.1016/j.edurev.2013.10.001.

[6] G. Lakoff. Explaining embodied cognition results. *Topics in Cognitive Science*, 4(4):773–785, 2012. doi:10.1111/j.1756-8765.2012.01222.x.

[7] E. A.-L. Lee and K. W. Wong. Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education*, 79:49–58, Oct. 2014. doi:10.1016/j.compedu.2014.07.010.

[8] G. B. Petersen, A. Mottelson, and G. Makransky. Pedagogical agents in educational vr: An in the wild study. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, pp. 1–12, 2021. doi:10.1145/3411764.3445760.

[9] G. Robertson, M. Czerwinski, and M. van Dantzich. Immersion in desktop virtual reality. In *Proceedings of the 10th Annual ACM Symposium on User Interface Software and Technology*, UIST '97, p. 11–19. Association for Computing Machinery, New York, NY, USA, 1997. doi:10.1145/263407.263409.

[10] J. J. G. van Merriënboer and J. Sweller. Cognitive Load Theory and Complex Learning: Recent Developments and Future Directions. *Educational Psychology Review*, 17(2):147–177, June 2005. doi:10.1007/s10648-005-3951-0.

[11] T. Xu, X. Wang, J. Wang, and Y. Zhou. From textbook to teacher: an adaptive intelligent tutoring system based on bci. In *2021 43rd Annual International Conference of the IEEE Engineering in Medicine Biology Society (EMBC)*, pp. 7621–7624, 2021. doi:10.1109/EMBC46164.2021.9629483.

[12] Y. Zhou, S. Ji, T. Xu, and Z. Wang. Promoting knowledge construction: A model for using virtual reality interaction to enhance learning. *Procedia Computer Science*, 130:239–246, 2018. doi:10.1016/j.procs.2018.04.035.