Shan, L. (2023). AI-enhanced adaptive Chinese language learning using VR "Mondly" app and "Uptale" platform. *International Journal of Technology in Teaching and Learning, 19* (2), 107-127.

AI-Enhanced Adaptive Chinese Language Learning Using VR "Mondly" App and "Uptale" Platform

Limei Shan
Lehigh University

This paper explores the impact of AI applications in adaptive learning on learners with varying levels of Chinese proficiency and diverse learning needs. The study investigates the robustness of VR technology in enhancing Chinese oral proficiency and pragmatic skills among 1st year and 2nd year learners. It compares the established VR app Mondly with Uptale, an Immersive VR authoring platform, focusing on AI-based voice recognition and analysis features. Additionally, it compares the adaptability of the Uptale VR platform with Mondly by analyzing factors such as difficulty adjustments, interactive features, learner engagement, and support for learner autonomy. It was found that VR technology, in general, enhances the Chinese learning experience. Uptale's additional features extend its potential utility in student assessment. These features include custom learning modules that incorporate students' self-shot 360-degree videos and interactive elements, such as multiple-choice questions, audio prompts, speaking/recording tags, feedback mechanisms, and analysis reports.

Keywords: AI voice analysis, VR experience, Mondly, Uptale, OPI

INTRODUCTION

Amid and after the pandemic, there's a rising demand for foreign language instruction in a more task-based and experiential setting, prompted by the suspension of study abroad programs and the difficulties in practicing with language partners on campus due to social distancing. Virtual Reality (VR) technology serves as an excellent alternative for these experiences. By replicating physical experiences in a virtual realm through spatial tracking and 3D visuals, VR offers an immersive platform. The immersive nature of VR technology enables the creation of a contextual environment that closely mirrors real-life, facilitating the achievement of both communication and language acquisition objectives.

The rapid growth of Natural Language Processing (NLP) technology and its related Artificial Intelligence (AI) applications presents both challenges and opportunities for foreign language teaching and learning. Maximizing the potential of AI to enhance the language acquisition process poses a significant challenge for language instructors.

Limei Shan is a teaching associate professor of Chinese at Department of Modern Languages & Literatures at Lehigh University. She can be reached at lis208@lehigh.edu.
DOI: https://doi.org/10.37120/ijttl.2023.19.2.02

Given the limited utilization of VR in the context of Chinese language acquisition, the researcher selected 2nd year Chinese as Foreign Language (CFL) learners in the Fall of 2022 to participate in a trial study aimed at evaluating the effectiveness of VR as an innovative tool for task-based Chinese Teaching (Shan, 2023). In this prior study, second-year Chinese students experienced both Mondly and Uptale in a scenario focused on "Dining Out" and affirmed that VR technology effectively enhances Chinese language acquisition (Shan, 2023) in line with other research in the field of VR-assisted language learning. However, there remains a gap in research concerning different proficiency levels, and the effectiveness and experience may vary with prolonged exposure to the target language.

BACKGROUND

Motivated by this research gap and building upon the success of the previous year's study, this research extends its scope to 1st year Chinese language learners, focusing on the theme of "Making Requests." It further compares two VR technologies: the established VR app Mondly and Uptale, an Immersive VR authoring platform. This paper presents the following contributions: (1) it conducts the first comparative study between the well-established VR App Mondly and the tailored/customized VR platform Uptale, specifically focusing on AI-based voice recognition analysis features; (2) it examines the robustness of both VR technologies in terms of enhancing Chinese oral proficiency and pragmatics, proposing that the Uptale VR is more adaptive to different proficiency levels of CFL learners; (3) it explores the feasibility of using Uptale to conduct OPI assessments instead of traditional face-to-face interviews.

The paper specifically highlights the features of the Uptale platform, which can be useful for further research on VR-facilitated language learning. These features include custom learning modules that integrate students' self-recorded 360-degree videos and interactive elements, such as multiple-choice questions, audio prompts, speaking/recording tags, feedback mechanisms, and analysis reports.

LITERATURE REVIEW

Previous research in the field of VR-assisted language teaching, including recent studies and the Lehigh trial study, shared a common focus on several key aspects. These include the types of participants involved, which encompass primary, secondary, and university-level learners, the specific research objectives, with an emphasis on effectiveness and the learning experience, as well as considerations related to language and language proficiency. Furthermore, these studies explored the types of VR technology employed and assessed the benefits and drawbacks of VR integration (Hua & Wang, 2023).

The findings from these studies consistently reported positive outcomes. Notably, they emphasized the significant affective benefits of VR in language learning. VR was found to "promote positive attitudes and emotions, such as liking, interest, enthusiasm, enjoyment, and satisfaction, which were translated into positive actions, enhancing task engagement". Additionally, VR was shown to contribute to "positive psychology, fostering motivation, confidence, and a willingness to communicate." It also had a "mitigating effect on negative psychological factors, like foreign language anxiety", and it can help improve "teacher-student relationships" (Hua & Wang, 2023, p. 5).

In terms of the established VR app Mondly and its impact on language learning, much of the research has been centered on English as a Foreign Language (EFL) learning. These studies concluded that "VR Mondly significantly enhanced EFL learners' listening comprehension and retention." VR users demonstrated improved recall across various measures (Tai, 2022, p. 16). Furthermore, research has explored the impact of VR, particularly Mondly, on EFL learners' vocabulary acquisition. The findings indicated that

"Mondly positively facilitated EFL learners' vocabulary learning and retention in terms of word meaning and contextual usage." VR users outperformed video watchers, owing to the scaffolding and feedback provided by avatars, which offered "opportunities for repeated practice" and "enhanced vocabulary learning and retention" (Tai, Chen, & Todd, 2022, p. 912).

Shan's study (2023) delved into language proficiency and pragmatics gains among 2nd year Chinese learners, employing a scaffolded Task-Based Language Teaching (TBLT) design that involves higher-level cognitive processes.

An emerging trend in VR-assisted language learning, highlighted by Hua and Wang (2023) and Parmaxi (2020), is the increasing immersion, interactivity, and customization of VR technologies. However, insufficient research has been conducted on customized VR platforms, particularly the leading one, Uptale. No studies have compared these two types of VR technologies: established VR apps and customized VR platforms. Additionally, there's a growing need for a comparative study among learners of different proficiency levels who have used both Mondly and Uptale. This paper addresses this research gap by proposing pedagogical suggestions on how to integrate VR technology into language learning and how to choose the right type of VR technology to better serve students' needs and teaching outcomes.

VR TECHNOLOGIES INVESTIGATED IN THIS STUDY

Virtual Reality simulates physical experiences in a virtual world by employing spatial tracking and 3D visuals. It is generally categorized into two types of VR: low-immersion VR (LiVR, such as computer screen-based) and high-immersion VR (HiVR with enclosed boxes or head-mounted displays [HMD]) (Kaplan-Rakowski & Gruber, 2019). Lehigh University's Data Visualization & XR Learning Lab is equipped to provide users with multiple VR platforms capable of offering a variety of curated immersive virtual reality experiences. Due to the high level of immersion afforded by an HMD, it effectively provides the user with a sensation of physical presence in a computer-generated virtual world. Thus, in this study, we utilized interactive 360-degree immersive experiences through an HMD—Oculus provided by CITL Visualization Lab. For the few students who opted to use a desktop computer with a two-dimensional monitor, which falls under LiVR, the research results excluded that part of the data due to the small sample size.

THE MONDLY APP AND UPTALE PLATFORM.

In recent years, most of the research in VR facilitating language acquisition has emphasized two types of VR technology: established commercial platforms and more tailored platforms (Hua & Wang, 2023). Prior to this study, the researcher tested several off-the-shelf language learning apps and VR social platforms, including VR Chat, Mondly, Allspace VR, Noun Town, and Hanyu VR. Some of these platforms focused on specific language acquisition areas, such as vocabulary (Noun Town), or were only available in a LiVR format (Hanyu VR is only accessible on phones and desktops). Others, such as VR Chat, were suitable for informal language practice, but there were numerous variables that couldn't be controlled in the virtual chat room. Based on it being the highest-rated VR app available for language learning and the positive reviews provided by Lehigh's 2nd year learners, Mondly was ultimately selected as the primary established VR platform for this comparative study.

Uptale, a French company established in 2017, is a pioneering platform that empowers educators and trainers to craft and share interactive 360-degree immersive experiences. It facilitates the development of custom immersive lessons using 360 video and imagery captured by students themselves, with no prior coding knowledge required. Through Uptale's platform, real-world scenarios are simulated as faithfully as possible, ensuring an

authentic learning experience (Uptale, 2023). The Uptale customized VR experience is accessible through either a VR headset or in 2D via a web browser. While Uptale is widely used for training in organizations and enterprises, there has been limited research in language acquisition, especially in Chinese language teaching and learning. As a new, innovative VR authoring platform, all the customized features that Uptale provides perfectly align with the requirements of Chinese language teaching and learning.

THE EMBEDDED AI VOICE ANALYSIS TECHNOLOGY IN MONDLY AND UPTALE.

Natural language processing (NLP) is a subfield of computer science, specifically within artificial intelligence (AI), focused on enabling computers to comprehend text and spoken words in a similar manner to human beings. Artificial intelligence-driven language understanding forms the foundation for various applications, such as "language modeling, chatbots, personal assistants, question answering, text summarization, speech-to-text conversion, sentiment analysis, machine translation, and more" (Rothman, 2021). Many of these widely used AI-powered voice analysis technologies are employed in both Mondly and Uptale.

Mondly is a foreign-language learning VR app developed by ATi Studios, designed to create an immersive virtual environment that replicates real-life scenarios, such as train rides, taxi trips, hotel lobbies, hotel rooms, and restaurants. The app utilizes state-of-the-art speech recognition technology from Nuance, a market leader in voice recognition, and is further enhanced by an object recognition engine capable of identifying objects within the text. It can understand millions of phrases across 33 languages.

Throughout the Mondly high-immersion experience, students wear an Oculus HMD and engage with a digital avatar in situational conversations, creating a sense of presence alongside virtual characters (Tai, 2022) As presented in Figures 1 and 2, the app can actively listen to the learner's responses, assessing both Chinese pronunciation and grammar structures. Positive feedback is provided in the form of a green checkmark that hovers over the transcript of the learner's speech. Following this encouragement, the avatar can seamlessly continue the conversation or offer verbal comments to motivate the learner to try again. For instance, the avatar may say phrases like "我不明白" (I don't understand), "请重复" (please repeat), or "请慢一点儿说" (please speak slower). During these interactions, learners can also perceive the avatar's feedback through its kinetic gestures, such as nodding, shaking its head, or using hand gestures.

Uptale, a specialized VR company, has established a partnership with Microsoft Teams and integrated Azure Speech services, which encompass the Text to Speech service featuring state-of-the-art neural voices and the Speech to Text service for voice recognition. In Uptale's VR environment, two significant features stand out which includes the real-time speech to text and batch transcription. The first feature transcribes audio in real-time as it is spoken through a microphone or from a file. What sets this technology apart is its ability to provide transcripts instantaneously as shown in Figure 3 and deliver feedback in the target language.

The second feature of Uptale is batch transcription which transcribes large volumes of audio stored in the system. Instructors can select audio files via a shared access signature (SAS) URI and asynchronously receive the transcription results as show in Figure 4 the Uptale User Report. This feature streamlines the process of digitizing, distributing, and monitoring student data generated during VR activities, offering valuable insights and analytics for language assessment and evaluations.

Figure 1. Mondly (positive feedback)

Figure 2. Mondly (negative feedback)

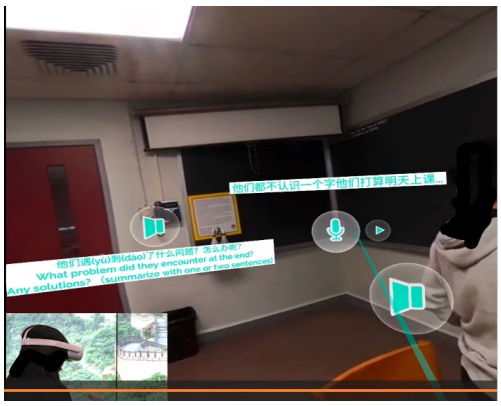

Figure 3. Uptale real-time speech to text

Figure 4. Uptale user report

METHODS

The research project received approval from the Institutional Review Board (IRB) of Lehigh University. Since it targets two pre-existing classes (1st and 2nd year) without the ability to create control and experimental groups, it didn't involve random samplings of all Chinese learners. Therefore, a quasi-experimental research type was employed in this empirical study.

RESEARCH QUESTIONS

The research questions for this study are described as follows:

1. How robust is VR technology (both Mondly and Uptale) in enhancing Chinese oral proficiency and pragmatic gain between 1st year and 2nd year learners?
2. To what extent does the Uptale VR platform demonstrate greater adaptability to the diverse proficiency levels of CFL learners compared to the VR App Mondly, considering factors such as difficulty adjustments, interactive features, learner engagement, and support for learner autonomy?

OVERVIEW

Taking the specific characteristics of Chinese educational linguistics into account, the aim of the research is to provide recommendations for Chinese language instructors on selecting and incorporating the appropriate VR technology to better align with their

teaching objectives. Notably, Uptale's additional features can expand the potential use of VR in student assessment.

During the 2022-2023 school year, this study involved two groups of participants, with ages ranging from 18 to 21. In the fall semester, ten 2nd year CFL learners participated (4 heritage learners and 6 non-heritage learners) using the Integrated Chinese Volume 2 textbook. The overall Chinese proficiency of these learners was at the ACTFL Novice High to Intermediate Low proficiency level. In the spring semester, nineteen 1st year learners took part (6 heritage learners and 13 non-heritage learners) using the Integrated Chinese Volume 1 textbook. Their Chinese proficiency was at ACTFL Novice Low to Novice High level.

DATA COLLECTION

Both groups utilized the established Mondly VR application and the new customized VR platform Uptale as part of their Chinese curriculum design. Both quantitative and qualitative data were collected for the descriptive analysis. For the quantitative data collecting, the Oral Proficiency Interview (OPI) was adopted as an instrument to measure CFL learners' oral proficiency and pragmatic gains and the rating scale was based upon the ACTFL Proficiency Guidelines (see Appendix 1). OPI is a valid and reliable means of assessing how well a person speaks a language which can be done through face to face or a phone interview. The test takers experience OPI in four parts: Warm-up, Level checks, Probes and Wind-down (American Council on the Teaching of Foreign Languages (ACTFL), 2018). Both participant groups underwent Pre-VR OPI and Post-VR OPI assessments (see Appendix II for sample OPI Role-play prompts). The data analysis involved employing descriptive statistics such as the mean and standard deviation, alongside the paired-sample t-test and the independent t-test (see Tables 1, 2, and 3).

During the Uptale VR intervention, individual user reports (Figure 4) were generated for each participant. These reports provide detailed information on the number of pop-up questions answered correctly, mistakes made, completion status of each scene, and heat-maps that illustrate the learners' attention focus. The Uptale user reports serve as evidence for language acquisition and offer insights into the engagement levels of the learners.

In the qualitative research phase, students participated in a survey (see Appendix III) aimed at comparing their experiences with the Mondly app and Uptale self-made VR videos. The objective was to assess how the AI features in VR applications adjust to varying levels of language proficiency and learning needs, as well as to gauge students' perceptions of these AI-driven interactions, learning engagement, and autonomy. In the analysis of the survey data, percentages in the summary chart were employed to offer a quantitative depiction of specific themes or the frequency of qualitative data. Additionally, narrative descriptions were included in this segment of data analysis.

PROCEDURE

While a previous study conducted by Shan (2023) concluded that students cannot rely solely on VR technology and that it must be integrated into a scaffolded curriculum design, the Task-Based Language Teaching (TBLT) pedagogy stands as a central framework employed in the Lehigh Chinese curriculum. TBLT emphasizes language in practical use rather than isolating language forms. Both TBLT and VR underscore the significance of immersion. VR places students in real-life scenarios to provide them with language input, while TBLT revolves around exposing learners to the target language within authentic conditions for learning. These approaches serve to enhance learner engagement in language tasks and foster learner autonomy, where learners take control of their learning to achieve functional communicative goals.

The effectiveness of incorporating VR technology into a TBLT design (Shan, 2023, p. 266) for 2nd year learners prompted the expansion of this research to the Spring 2023 Year 1 (Beginner) Chinese class. In this study, the modified Lesson 6 TBLT design comprises six subtasks/sessions, all structured to achieve a target task during the Evaluation session which is "Making requests or apologies in real-life scenarios in a more pragmatically suitable manner", following the OPI procedure. The pedagogic tasks in this study are sequenced based on the cognitive complexity of the tasks, aligning with Robinson's Cognition Hypothesis, which suggests that "pedagogic tasks be sequenced for learners largely on the basis of increases in their cognitive complexity so as to increasingly approximate the demands of real-world target tasks" (Robinson., 2001b) (Robinson, 2007a.).

Subtask 1 introduces basic vocabulary related to making appointments. Subtask 2 covers the sequence of making requests, apologetic expressions, and strategies through matching and multiple-choice exercises. Jigsaw critique in this subtask involves group discussions, selection, comparison, and reasoning about various ways to make requests. In Subtask 3, students learn to use narrative scripts with cohesive devices to reinforce making requests in appropriate pragmatic manners with friends and teachers. Subtask 4 involves watching a provided 2D video, expanding vocabulary, and practicing making requests, refusals, and apologies, including an OPI procedure (Pretest). Subtask 5 presents role-play scenarios created outside class, while some students prepare to record a 360-video using a Go-pro camera. In Subtask 6, students visit the VR lab, engaging with both Mondly VR and self-made Uptale VR videos. They respond to embedded questions in VR videos, provide written reflections through google forms, and perform a situational conversation with the instructor at the end, focusing on language proficiency and pragmatic competence.

The progression of these six pedagogic subtasks reflects a gradual increase in complexity, preparing students for the transition from conducting brief conversations to delivering extended prepared speeches and moving from rehearsed role-plays to unscripted situational interviews. As illustrated in Figures 5, both the resource-directing and resource-depleting dimensions involve a growing array of elements. Each subtask level incorporates more planning, prior knowledge, and heightened reasoning to prepare students for more intricate tasks. At the culmination, students engage in a concise situational conversation with the instructor in a provided Chinese scenario. This activity was assessed using the OPI rating scale, with the complexity of this target task reaching its zenith.

Similar to the 2nd year learners (Shan, 2023), the 1st year students' VR intervention was also divided into two stages to experience different AI-based Voice Analysis technology. In the initial stage, students interacted with a digital avatar in situational conversations through the Mondly App. They progressed through structured learning modules using Quest 2 headsets, marking their initial exposure to VR. In the subsequent stage, Uptale served as an immersive VR authoring platform, utilizing 360-degree videos or images captured by the students themselves. This platform facilitated assessment-based interactions and could be experienced either with VR headsets or in 2D via a web browser.

The second stage of Uptale's development process involved students proposing conversational scenarios, scripting dialogues, enacting these scenarios in Chinese, mastering the 360-degree camera, selecting shooting locations, rehearsing scenes, and recording them using the 360-degree camera. These recordings were then incorporated into the VR experience, enriched with interactive elements, including multiple-choice questions, audio prompts, speaking/recording cues, feedback, and grading reports. Students visited the VR lab and engaged with immersive 360-degree videos generated using the Uptale authoring platform. While watching the videos through the headset, students responded to the questions integrated into the VR Uptale videos. Subsequently, they

submitted written reflection reports through a Google form, sharing their perceptions of VR technology.

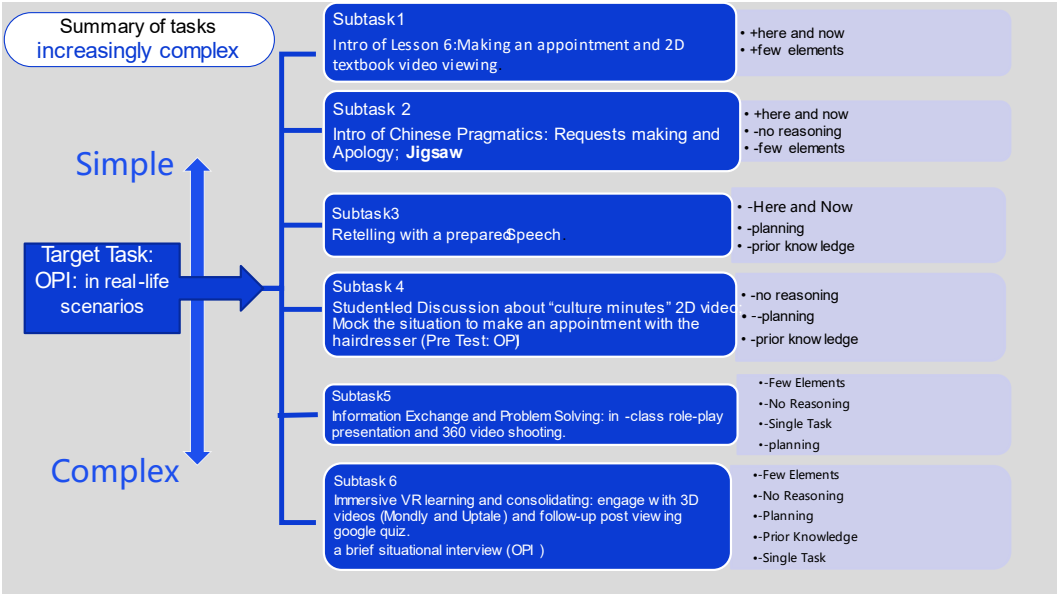

Figure 5. 1st Year lesson 6 TBLT task design

RESULTS AND ANALYSIS

The Pre-VR OPI and Post-VR OPI data, survey results and data generated from Uptale VR addressed the following two research questions.

Research Question 1: How robust is VR technology (both Mondly and Uptale) in enhancing Chinese oral proficiency and pragmatic gain between 1st year and 2nd year learners?

To answer this question, the OPI test was conducted twice for each class (i.e., before and after the VR intervention). The purpose is to measure whether robust VR technology (both Mondly and Uptale) can enhance Chinese oral proficiency and pragmatic gain between 1st year and 2nd year learners. The data for 1st year OPI was analyzed through the paired-sample t test in SPSS. As shown in Table 1, the students in post-test had significantly higher means than their pretest, $t(18) = 4.12$, 2-sided p value < .001, Cohen's d = 0.94. Table 2 also illustrates a significant improvement between pretest and post-test among 2nd year learners, $t(14) = 9.46$, 2-sided p value < .001, Cohen's d = 2.44.

Table 1. *Paired Samples t Test for the 1st Year (N = 19)*

		Mean	SD	S.E.	t	*df*	*p* value	Cohen's d
1st year	PreTest	77.68	14.50	3.33	4.12	18	< .001	0.94
	PostTest	91.58	3.53	.81				

Table 2. *Paired Samples t Test for the 2nd Year (N = 15)*

		Mean	SD	S.E.	t	*df*	*p* value	Cohen's d
2nd year	PreTest	69.60	7.39	1.91	9.46	14	< .001	2.44
	PostTest	87.87	4.10	1.06				

To better compare the OPI score improvement between 1st and 2nd year learners, an independent t test was conducted. The average improvement for 2nd year students is non-

significantly higher than for the 1st year students, t(32) = 1.05, p = .303, Cohen's d = 0.36 (see Table 3).

However, the findings suggest that 2nd year learners have higher mean than the 1st year, indicating that students with higher proficiency levels benefit more from the VR facilitated learning. This is likely because they caught up faster due to the longer time of exposure to this language. Overall, the OPI score improvement between 1st and 2nd year learners support the robustness of VR technology promoting CFL learners' proficiency.

Table 3. *Independent t-test for Change Scores*

Class		Mean	SD	S.E.	t	df	p value	Cohen's d
Change Score	1st year	13.89	14.71	3.38	1.05	32	.303	0.36
	2nd year	18.27	7.48	1.93				

Research Question 2. To what extent does the Uptale VR platform demonstrate greater adaptability to the diverse proficiency levels of CFL learners compared to the VR App Mondly? This will be evaluated using factors such as difficulty adjustments, interactive features, learner engagement, and support for learner autonomy.

DIFFICULTY ADJUSTMENTS FOR 1ST AND 2ND YEAR LEARNERS

As noted above, AI-based voice recognition technology is employed in both Mondly and Uptale, and the "speech to text" feature is widely used. When learners speak into the microphone, an instant transcript appears next to the icon. However, Mondly's speech recognition system appears to be somewhat outdated, as Figure 6 illustrates. It occasionally misinterprets characters, particularly polyphonic ones. Additionally, to pass the voice recognition, one must strictly adhere to the provided "hints." Lastly, the linguistic correctness and suitability of Mondly's "hints" can be questionable. It is evident that more input from Chinese linguistic professionals is needed in the course design, a challenge highlighted in previous studies concerning Automatic Speech Recognition (ASR) systems (Tai, 2022).

Figure 6. Mondly typos and issues in hints

On this front, Uptale performs slightly better in accurately representing characters in the transcript. Nevertheless, both platforms struggle with recognizing tones. Since Chinese is a tonal language, this leads to moments of frustration for learners at all levels during their VR experience.

It is worth noting that more 2nd year learners (15.4%), as shown in Table 4, disagreed with some of the hints provided by the Mondly app compared to their 1st year counterparts (0%). This suggests that 2nd year learners possess a greater capacity for critiquing the suggestions offered by Mondly and points to a deficiency in the Mondly course design in

terms of linguistic support. It fails to account for the variety of vocabulary and grammar structure choices at varying proficiency levels. Additionally, Mondly's course scaffolding doesn't effectively convey the differentiation between beginner, intermediate, and advanced levels. Moreover, more 2nd year learners (46.2%) consider Uptale's customized videos to be suitably challenging, as they reinforce classroom learning. In contrast, Mondly's videos are perceived as less challenging by 2nd year learners. In terms of the difficulty level for 1st and 2nd year learners, 1st year learners generally perceive the videos as somewhat more challenging (66.7%/55.6%) compared to 2nd year learners (38.5%). Prior research also indicates that when engaging in activities in virtual environments, learners sometimes struggle to maintain focus on the learning tasks. Beginners might face a higher cognitive load, as they need to manage the "dual input of spoken messages accompanied by simultaneously displayed text," all while participating in real-time interactions during immersive exercises (Tai, 2022, p. 15).

Table 4. *Difficulty Level Perceived by 1st and 2nd Year Learners*

	1st Year	2nd year
Mondly	somehow challenging (66.7%) No one disagree the "hints"	Somehow challenging (38.5%) Disagree with the "hints" (15.4%)
Uptale	somehow challenging (55.6%) more appropriate level (11.1%)	somehow challenging (38.5%) more appropriate level (46.2%)

LEARNERS' ENGAGEMENT THROUGH INTERACTIVE FEATURES

VR interactions could foster learners' engagement in the immersive environment, promoting a state of language flow. Table 5 shows that students valued the immersive features offered by both Mondly and Uptale, such as hints, multiple-choice questions, immediate feedback, and scene replays. One learner stated: "I like the multiple choice asking about what you just watched." While some 1st year students enjoyed interacting with the avatars in Mondly, Uptale lacks virtual characters to enhance the sense of presence. However, the majority of 2nd year learners (61.5%) enjoyed responding to self-created VR interactions: "Being able to stand in the middle of a virtual space is great and allows for a much more dynamic viewing and learning experience."

Table 5. *Mondly vs Uptale: Engagement & student Preferred Interactions*

	1st Year	2nd year
Mondly	with Avatar (77.8%) Hints (66.7%) Voice Recognition less forgiveness (27.8%)	with Avatar (53.8%) Hints (76.9) Voice recognition less forgiveness (69.2%)
Uptale	Respond to self-built VR (27.8) Multiple Choice (83.3%) Scene "Replay" (77.8%) immediate feedback (61.1%)	Respond to self-built VR (61.5%)Multiple Choice (100%) Scene "Replay" (92.3%) immediate feedback (53.8%)
VR videos *vs* 2D TB ones	immediate response (66.7%) Better Engage (55.6%) self-shot videos replicate real life scenarios better (50%) uptake videos connect to TB topics better (44.4%)	immediate response (76.9%) Better Engage (61.5%) self-shot videos replicate real life scenarios better (50%) lack of distraction (46.2%)

Without Pinyin	not a big issue, still I still can comprehend fully in the immersive environment (61.1%)	not a big issue, still I still can comprehend fully in the immersive environment (88.9%)
Props and Stages in uptale	No, it doesn't matter. The props and stage setting is fine. (50%)	No, it doesn't matter. The props and stage setting is fine. (69.2%)

As shown in Table 5, in the VR environment students exhibited higher levels of engagement compared to watching traditional textbook videos. They perceived the VR experience as a more distraction-free environment (44.4 % for 1st year and 46.2% for 2nd year), engaging more deeply with the materials through pop-up style questions and spontaneous feedback (>50%). Self-shot videos were also seen as more accurate in replicating real-life scenarios (50%). As one learner stated:

> *"I like how it can help simulate real conversations much better than just listening on an app, especially since having things such as different accents present is important to be able to listen through in real life."*

Engagement has been defined by various authors in different ways. In this study engagement is defined as "what occurs during a task or experience, a result of the interaction between learners and the characteristics of the task and the supporting environment" (Bodzin, Junior, Hammond, & Anastasio, 2020, p. 349). It consists of behavioral, emotional, and cognitive engagement. Studies indicate that students' learning can benefit from the synergy of these three facets of engagement, as proposed by Fredricks, Blumenfeld, and Paris (2004), and virtual reality has the potential to enhance all of them (Shernof, 2012). In the VR learning environment of this study, students adhered to the rule of participating in a learning task (behavioral). The novelty and immersive nature of the VR headset, along with feelings of presence, a high-definition and realistic interface, student-directed skits, increased curiosity, attention, satisfaction, and enjoyment, leading to continuous engagement (emotional). This, in turn, fostered "thoughtfulness and willingness to have deeper understanding of the content, to exert effort in completing a real-life task or problem-solving" (cognitive) (Li, Zheng, & Lajoie, 2023, p. 1533). This combined engagement potentially facilitated a state of flow characterized by focused concentration and enjoyment in the learning process, ultimately improving fluency (Philp & Duchesne, 2016). Learners noted: "*I like being able to hear the narration voice in Chinese alongside the questions. I also think the interactive nature makes it easier to focus on what they are saying in the skit*"; "*I thought it was very engaging and since it was such an exciting method of learning it was very easy to stay engaged*"; and "*I like how it feels like you're actually there*."

The findings are also supported by data analytics in the form of heatmaps, viewpoint tracking, and area of interest displays. Heatmaps on Uptale, as shown in Figure 7, illustrate the relative intensity of learners' attention in specific areas of the experience. This is captured by an invisible cursor in front of the VR users' eyes or at the center of the page for web users. Warm colors (red) indicate a high level of activity/interest, while cold colors (blue) represent lower levels of activity/interest. Green indicates some interest/attention. Heatmaps are generated for each scene of an experience and can be found in the data report page as a specific tab.

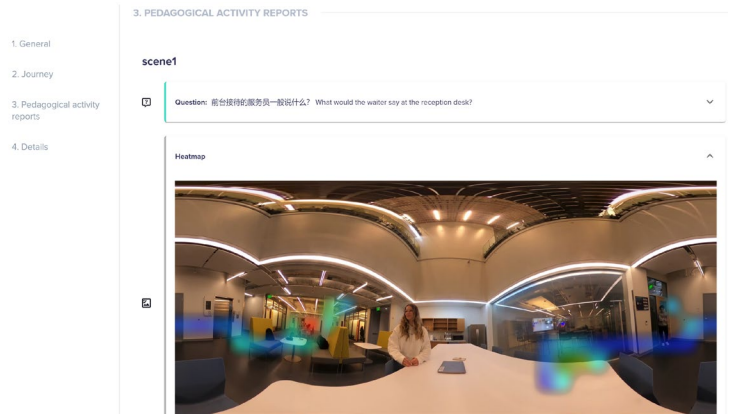

Figure 7. Heatmaps on uptale user report

It's noteworthy that the absence of Pinyin support did not significantly impede students' comprehension of dialogue in the immersive VR environment. This finding holds significance for Chinese language learning, where Pinyin is often heavily relied upon in regular classroom teaching. Additionally, it's interesting to observe that students focused more on language performance and content rather than the props and stage settings in the scenes. These findings are noteworthy for future research, although they are not the focus of this paper.

PROMOTION OF LEARNERS' AUTONOMY

In the VR learning environment, "learners have autonomous control to intentionally select a pathway that personalizes their learning to accomplish the task" (Bodzin, Junior, Hammond, & Anastasio, 2020, p. 349). This learning autonomy, self-regulation skills, and strategic planning positively contribute to students maintaining "a high level of cognitive engagement throughout the problem-solving process", leading to success (Li, Lajoie, Zheng, Wu, & Cheng, 2021, p. 4). The feeling of autonomy through Uptale is seen from the perspective of both the student using the VR technology and those who shot the 360 videos.

Within the Uptale VR intervention, learners have the unique capability to 'replay' scenes, affording them the flexibility to restart their immersive experience as necessary. This feature empowers learners to tailor their learning pace to suit individual needs. A significant number of learners express the belief that this Uptale VR technology provides them with a sense of control over their learning journey and allows for effective self-monitoring of their performance.

On the other hand, students who shot the 360 videos reflected their unique experience on learner's autonomy in the survey. In general, students at both proficiency levels appreciate the opportunity to create their own 360 videos as indicated in Table 6. Students found that shooting the video is perceived as less anxiety-inducing compared to traditional classroom presentations, despite the longer preparation and shooting time required. A significant number of students find the experience enjoyable and pleasant, expressing their interest in shooting more videos around the campus. The following was derived from the collected student survey responses: "It's fun to make the videos and way less pressure than presenting in class. Also fun to show my video in front of the class but not everyone may agree with me on that" and "*The 360 shooting takes pressure off speaking in class*." It appears that 2nd year learners are more comfortable with on-campus video shooting, possibly because they are more familiar with the school environment and experience less stage fright.

Table 6. *Learners' Autonomy: Perception of 360 Video Producers*

	1st Year	2nd year
Longer Time commitment (more than 2 hours)	41.7%	55.6%
Making the video produces less anxiety than presenting in front of the class (you can rehearse between scenes, and re-shoot the video)	63.6%	66.7%
I was comfortable shooting the video around the campus	54.5%	55.6%
Shooting the videos was a fun experience overall. I want to do more.	45.5%	55.6%
It enhanced my OWN learning of Mandarin Chinese.	definite (63.6%) somehow (27.3)	definite (22.2%) somehow (44.4%)

DISCUSSION

The current study yielded two primary findings that answered the research questions and expanded the literature on the comparison study between established VR apps (Mondly) and customized VR platforms (Uptale). The findings indicate that CFL learners achieve higher levels of oral proficiency when VR technology is integrated into the curriculum, particularly with the more adaptable nature of Uptale VR. As a result, it becomes imperative to explore its incorporation into Task-Based Language Teaching (TBLT) instructional designs. Moreover, there is potential to consider Uptale VR technology as the alternative assessment tool for the Simulated Oral Proficiency Interview (SOPI) instead of the traditional Oral Proficiency Interview (OPI).

ADAPTIVE LEARNING FOR DIVERSE PROFICIENCY LEVELS AND LEARNING NEEDS

As mentioned earlier, the 1st year learners are at the ACTFL Novice Low to Novice High proficiency level, while 2nd year learners are at the ACTFL Novice High to Intermediate Low level. Mondly offers a limited choice of only three proficiency levels, whereas Uptale provides an authoring platform that allows for the customization of interactions in the virtual world. As evident in Figure 8, learners at both proficiency levels believe that Uptale's customized videos complement classroom materials more effectively. This is primarily due to Uptale's robust "Text to Speech" technology, which empowers instructors to tailor interactions, scaffolding them to cater to students' varying proficiency levels, learning styles, and individual needs. For instance, instructors can input their own "rubrics" or prompts, which may differ between 1st year and 2nd year learners.

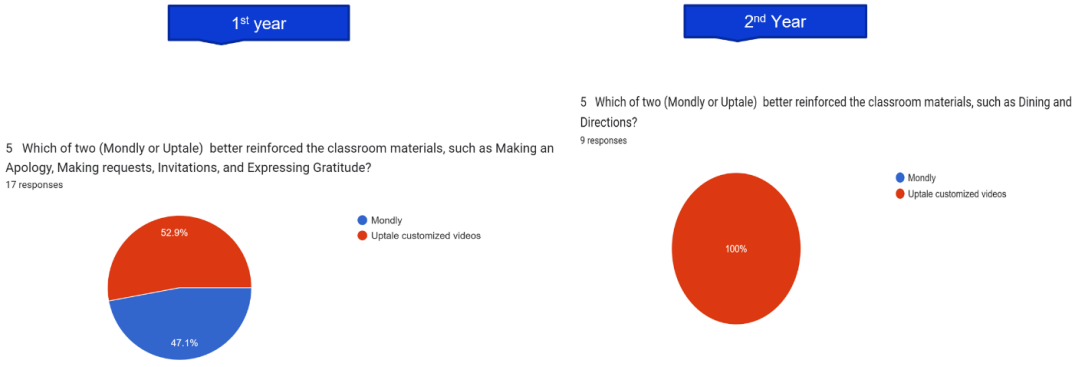

Figure 8. Adaptive learning for reinforcing the class materials

An example of how scaffolded task design is reinforced in the Uptale customized VR experience can be observed in Figure 9. Following the observation of a role-play scene, the video pauses, and questions appear. Students are then prompted to practice answering similar questions twice within the VR video. For instance, they might be asked, “要是你是小平，你怎么和老师约时间？” ("How would you make an appointment with Teacher Wang to ask her questions?") The first time their response is assessed through voice recognition, and the second time they respond to a multiple-choice question. Following the VR viewing activity, students engage in a Google Form quiz, which serves to recap and reinforce their knowledge of making a similar request. As one learner stated: “The questions that appear throughout the dialogue help summarize what's happening in the dialogue”. By the time students reaches the instructor for the OPI interview, they are thoroughly prepared and possess greater confidence to complete the final target task, which involves problem-solving.

Throughout this VR-facilitated, goal-oriented process, students remain focused, continuously challenged at their respective proficiency levels, and receive feedback concurrently. These conditions significantly enhance the likelihood of students reaching a heightened state of flow, which is considered the "ultimate stage of engagement" for developing fluency (Philp & Duchesne, 2016). As other research has indicated, individuals “tend to experience flow when engaged in activities that appropriately challenge their skill levels” (Shernof, 2012, p. 350), and this, coupled with the provision of opportunities and feedback for skill enhancement, further supports this endeavor (Shernof, Csikszentmihalyi, Schneider, & Shernof, 2003).

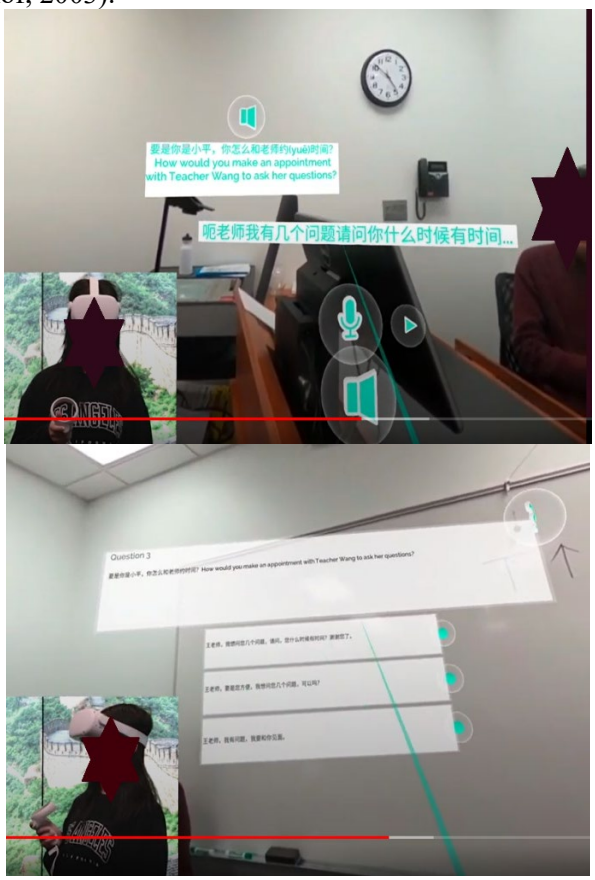

Figure 9. The scaffolded task design in uptale videos: Voice recognition then multiple-choice question

ADVANTAGES OF EMPLOYING UPTALE FOR SOPI COMPARED TO TRADITIONAL OPI

In 1986 the Center for Applied Linguistics (CAL) developed the Simulated Oral Proficiency Interview (SOPI) for Chinese (Yang & Ke, 2021). SOPI is "a type-mediated test designed to be surrogate for OPI in situations where a face-to-face interview is not possible or desirable" (Stansfield & Kenyon, 1992). Based on previous research, both SOPI and OPI test formats were proved to be effective and reliable methods to measure student's language proficiency (Kuo, 1997).

As shown in this project, VR primarily functions as a robust rehearsal tool, having the potential to evolve into an assessment tool, such as SOPI, for evaluating Chinese oral proficiency to provide an alternative method to traditional face-to-face procedures. In contrast to the conventional SOPI format where test items are recorded on the test tape and printed in the test booklet, the Uptale VR platform offers more advanced voice analysis tools, making SOPI more feasible. These tools include real-time speech-to-text and batch transcription provided by Microsoft Azure Cognitive Services. The Uptale VR platform allows the generation of a complete transcript and audio replay, as illustrated in Figure 10. The Longer Transcript with analyzing metrics for language proficiency is presented in Figure 11. This capability establishes it as an alternative tool for assessing oral proficiency or conducting speech analysis.

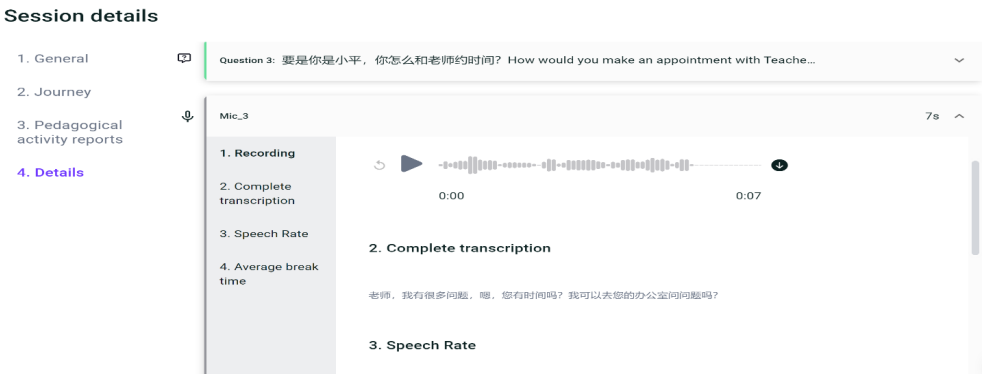

*Figure 10. A*udio replay and transcript

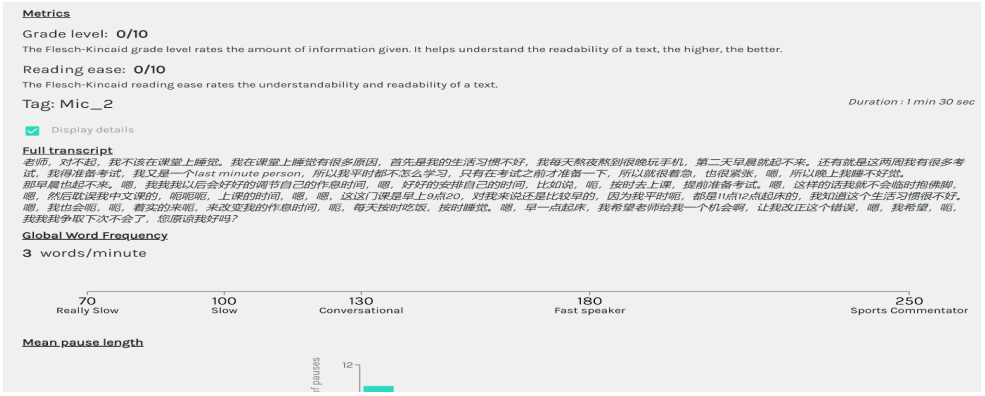

Figure 11. Longer Transcript with analyzing metrics for language proficiency

Additionally, Uptale offers more natural situational prompts for the Oral Proficiency Interview (OPI), eliminating the need to place students in a testing room for interviews with instructors. All OPI questions can be integrated into the VR video, creating a relaxed,

distraction-free environment with a more objective and standardized testing process for assessing Chinese language oral proficiency and pragmatic skills. Furthermore, the incorporation of 3D objects and magic tags in the VR interactions design can further enhance learners' immersion and interactivity which might improve learners test performance.

Regarding the actual test performance in OPI and Uptale SOPI, as well as the examinees' perceptions of the comparison between these two test formats, further research is required for a thorough investigation. This can be considered as the next focus for a research project.

LIMITATIONS

This research is subject to several limitations. Firstly, the small sample size and the non-randomized sampling method of CFL learners may restrict the applicability of the descriptive findings to the broader student population. For example, the non-significant change for the first and second-year students in the independent t test (see Table 3) is possibly because the sample size for the comparison is not big enough. With a larger sample size in the future research, the improvement between classes may turn significant.

Secondly, the study heavily depends on self-reported data collected through surveys, raising the potential for response bias or recall bias. Thirdly, the exclusion of a subset of desktop browser LiVR users could potentially influence their perceptions of VR differently. Lastly, the study didn't run the statistical analysis based on participants' prior VR experiences or their cultural backgrounds (heritage vs. non-heritage) due to the small sample size, which may impact their overall acceptance of VR technology.

SUGGESTIONS ON EXPLORING OTHER LANGUAGE SKILLS USING VR TOOLS

One of the drawbacks of Mondly is its lack of provision for written practice, hindering the development of reading and writing skills. In contrast, Uptale allows for the integration of reading exercises, where learners read Chinese characters while listening to them. Some researchers have utilized the VR interface to train academic writing structures, but the actual assignments still need to be completed on paper (Barrett, Pack, & Quaid, 2021). As VR technology advances, we have a great opportunity to boost the reading and writing skills of advanced Chinese learners. To meet this need, we should integrate VR seamlessly into the curriculum. With VR progressing quickly, researchers must be prepared to embrace this challenge soon.

SUGGESTIONS ON AVOIDING COGNITIVE OVERLOADING

Previous studies have indicated that excessive stimulation in Immersive Virtual Reality (IVR) may "overwhelm and distract learners" (Zinchenko, et al., 2020, p. 2). In this research, certain first-year learners may have felt overwhelmed by the language input, resulting in their preference for listening to the narration without simultaneously reading the text. Furthermore, the textual content vies for the learners' constrained visual processing capacity, leading to their expending needless effort in " reconciling these dual information streams" (Zinchenko, et al., 2020, p. 3).Therefore, as suggested by Liu, Yu, Liu, Pi, & Cui (2023), educators should "prioritize content optimization", the design of elements that pique students' curiosity, the creation of user-friendly interfaces, and "smooth device operation" experiences, particularly to prevent cognitive overload among entry-level learners. During the development of the Uptale VR experience, it is crucial to "employ well-designed audio-visual cues to direct users' attention, as these are key to achieving optimal user immersion and spatial presence" (Sarker, 2017). Balancing the use of cues for various student proficiency levels in VR-enhanced learning can significantly contribute to their success.

CONCLUSION

In conclusion, this study emphasizes the pivotal role of virtual reality technology in creating an advantageous environment for Chinese language acquisition. Firstly, it pioneers a comparative study between the widely used VR App Mondly and the customized VR platform Uptale, specifically focusing on AI-based voice recognition analysis features. Secondly, both VR technologies make positive contributions to Chinese oral proficiency and pragmatics by improving OPI score significantly, with Uptale displaying superior adaptability across diverse proficiency levels. The tailored features of Uptale, such as dynamic difficulty adjustments and enhanced interactive elements, enhance its effectiveness in meeting the varied needs of CFL learners. The study reveals that well-designed audio-visual cues and pedagogical interactions in VR environments elevate user immersion and engagement. Lastly, the research explores the feasibility of utilizing "Uptale" for Oral Proficiency Interviews (OPI) assessments, presenting an alternative to traditional face-to-face interviews. These insights illuminate the potential of VR platforms, particularly Uptale, to provide personalized and adaptive learning experiences in CFL education.

ACKNOWLEDGEMENT

Special thanks to the Center for Innovation in Teaching and Learning (CITL) for funding this project and to Lehigh's Data Visualization & XR Learning Lab for their contributions and technical expertise.

REFERENCES

American Council on the Teaching of Foreign Languages (ACTFL). (2018). *OPI Examinee Handbook* https://www.languagetesting.com/pub/media/wysiwyg/PDF/opi-examinee-handbook.pdf

Aubrey, S. (2017). Inter-cultural contact and flow in a task-based Japanese EFL classroom and Gayle Gregory & Martha Kaufeldt, “The motivated brain: Improving student attention, engagement, and perseve. *Language Teaching Research Vol. 21(6)* , 719-720.

Barrett, A. J., Pack, A., & Quaid, E. D. (2021). Understanding learners’ acceptance of high-immersion virtual reality systems: insights from confirmatory and exploratory PLS-SEM analyses. *Computers & Education 169*, 13.

Bloom, B. S. (1956). *Taxonomy of Educational Objectives, Handbook I: The Cognitive Domain.* New York: David McKay Co Inc,.

Bodzin, A., Junior, A., Hammond, T., & Anastasio, D. (2020). Investigating Engagement and Flow with a Placed-Based Immersive. *Journal of Science Education and Technology (2021 30:*, 347–360.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School Engagement: Potential of the Concept, State of the Evidence. *Review of Educational Research Vol. 74, No. 1,*, 59-109.

Gayle Gregory & Martha Kaufeldt. (2015). “The motivated brain: Improving student attention, engagement, and perseverance,” Alexandra, VA: ASCD (2015): 150-161. *Alexandra, VA: ASCD* , 150-161.

Hua, C., & Wang, J. (2023). Virtual reality-assisted language learning: A follow-up review (2018–2022). *Frontier Psychology. 14:1153642.*, 1-13, pg1.

Kaplan-Rakowski, R., & Gruber, A. (2019). Low-immersion versus high-immersion virtual reality: Definitions, classification, and examples with a foreign language focus. *Proceedings of the Innovation in Language Learning International Conference*.

Kun-Wan Philip Choong. (2011). “Task Complexity and Linguistic Complexity: An Exploratory Study Teachers College,” ,. *Columbia University Working Papers in TESOL & Applied Linguistics, Vol. 11, No. 1*, pp. 1-28.

Kuo, J. (1997). Assessing the Assessments: The OPI and the SOPI. *Foreign Language Annals Volume: 30 Issue 4 ISSN: 0015-718X Online ISSN: 1944-9720*.

Li, S., Lajoie, S. P., Zheng, J., Wu, H., & Cheng, H. (2021). Automated detection of cognitive engagement to inform the art of staying engaged in problem-solving. *Computers & Education volume 163*.

Li, S., Zheng, J., & Lajoie, S. P. (2023). The relationship between cognitive engagement and students' performance in a simulation-based training environment: an information-processing perspective. *Interactive Learning Environments; Vol. 31 Issue 3,*, 1532-1545.

Liu, Z., Yu, P., Liu, j., Pi, Z., & Cui, W. (2023). How do students' self-regulation skills affect learning satisfaction and continuous intention within desktop-based virtual reality? A structural equation modelling approach. *British Journal of Educational Technology, 54*, 667–685.

Uptale. (2023). *VR for Education - Design Immersive Learning Experiences for Students - Uptale*. https://www.uptale.io/en/education/

Parmaxi, A. (2020). Virtual reality in language learning: a systematic review and implications for research and practice. *Interactive Learning Environment*.

Philp, J., & Duchesne, S. (2016). Exploring engagement in tasks in the language classroom. *Annual Review of Applied Linguistics 36*, 50-72.

Robinson, P. (2007a.). Criteria for grading and sequencing pedagogic tasks. In P. Robinson, *Investigating Tasks in Formal Language Learning* (pp. 7-27). Clevedon, UK: Multilingual Matters.

Robinson, P. (2005). Cognitive complexity and task sequencing: studies in a componential framework for second language task design. . *International Review of Applied Linguistics, 43,*, 1-32. .

Robinson., P. (2001b). Task complexity, task difficulty, and task production: Exploring interactions in a componential framework. *Applied Linguistics, 22*, 27 – 57.

Rothman, D. (2021). Transformers for Natural Language Processing : Build Innovative Deep Neural Network Architectures for NLP with Python, Pytorch, TensorFlow, BERT, RoBERTa, and More. *Birmingham : Packt Publishing, Limited*, Preface.

Sarker, B. (2017). Decoding the user experience in mobile virtual reality narratives. *S. Lackey & J. Chen (Eds.), International Conference on Virtual, Augmented and Mixed Reality*, 437–452.

Shan, L. (2023). The Use of VR in Task-based Teaching for Mandarin Chinese. *AMPS procedings series 31*, 252-264.

Shernof, D. J. (2012). Engagement and positive youth development: Creating optimal learning environments. (A. P. Association, Ed.) *K. R. Harris, S. Graham, & T. Urdan (Eds.), The APA educational psychology handbook*, 195–220.

Shernof, D. J., Csikszentmihalyi, M., Schneider, B., & Shernof, E. S. (2003). Student engagement in high school classrooms from the perspective of fow theory. *School Psychology Quarterly, Vol. 18, No. 2*, 158–176.

Stansfield, C. W., & Kenyon, D. M. (1992). Research on the Comparability of the Oral Proficiency Interview and the Simulated Oral Proficiency Interview. *Center for Applied Linguistics, Washington, DC, USA Vol. 20, No. 3*, 347-364.

Tai, T.-Y. (2022). Impact of mobile virtual reality on EFL learners’ listening comprehension. *Language Learning & Technology*, 13.

Tai, T.-Y., Chen, H. H.-J., & Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Comput. Assist. Lang. Learn. 35,*, 892–917.

Yang, L., & Ke, C. (2021). Proficiency and pragmatic production in L2 Chinese study Abroad. *ScienceDirect System 98 (2021): 102475 www.elsevier.com/locate/system*.

Zinchenko, Y. P., Khoroshikh, P. P., Sergievich, A. A., Smirnov, A. S., Tumyalis, A. V., Kovalev, A. I., . . . Golokhvast, K. S. (2020). Virtual reality is more efficient in learning human heart anatomy especially for subjects with low baseline knowledge. . *New Ideas in Psychology, 59*, 2-3.

APPENDIX I

OPI/SOPI Holistic rating scale: 5-point, assessing communicative function, linguistic accuracy, and appropriateness (Yang&Ke, 2021)

0 Not ratable

Making no response; responding with utterances that are extremely hard to comprehend, or with language that has no connection to the scenario.

1 Not acceptable

Responding with language that fails to communicate target functions; with utterances that may potentially offend the hearer, or with inappropriate choice of pragmatic strategies for the context.

2 Problematic

Responding with language that somewhat communicates target functions; with pragma-linguistic errors that cause misunderstandings, but of a less serious nature; with certain language so strange or unexpected that causes awkwardness for the hearer.

3 Acceptable

Responding with language that communicates target functions; with utterances that are mostly appropriate for the

context, but may contain grammatical/lexical errors that do not interfere much with appropriateness.

4 Good

Responding with language that fully communicates target functions; with utterances that are pragmatically appropriate for the specified context, but may still sound non-native-like in terms of register, length, intonation, etc.

5 Near-native

Responding with clear and appropriate utterances that fully communicate target functions, close to native responses.

APPENDIX II

Sample of OPI Role-play Prompts

This is a test of your ability to express yourself in Chinese. Chinese will be used throughout the exercise. The interview will be recorded and last between 2 and 3 minutes. During the interview, we will discuss two scenarios from the list below. Your personal opinions and points of view will not in any way affect your grade. However, your ability to state and support opinions will be evaluated. At some point during the interview, I may also ask you to participate in a role-play situation. I will introduce the role play in Chinese (the English instructions are provided below) and you and I will act out the situation in Chinese. The length of the role play will depend on our interaction.

Your participation in the interview is important and will help you show the level of your Chinese oral proficiency and pragmatic competency.

Do you have any questions before we get started?

1 要(yào)是(shi)你在课(kè)堂(táng)上睡(shuì)觉(jiào)，你怎么向老师道(dào)歉(qiàn)?（if you sleep in class, how would you apologize to the teacher? Try to use direct apology +explanation or promise of forbearance ）

2 如(rú)果(guǒ)你(nǐ)有(yǒu)问(wèn)题(tí)，你(nǐ)怎(zěn)么(me)和(hé)老(lǎo)师(shī)约(yuē)时(shí)间(jiān)？ How would you make an appointment with your teacher to ask questions? （Hints: want statement +conventional indirect strategy ）

3 要是你是王老师(yàoshinǐshìwánglǎoshī)，你得去开会(nǐděiqùkāihuì)，你说什么(nǐshuōshénme)？ if you are Teacher wang and you need to rush to a meeting, how would you end the conversation politely with a student?

4 如(rú)果(guǒ)你(nǐ)希(xī)望(wàng)妈(mā)妈(mā)帮(bāng)你(nǐ)复(fù)习(xí)中(zhōng)文(wén)，你(nǐ)怎(zěn)么(me)和(hé)妈(mā)妈(mā)约(yuē)时(shí)间(jiān)？ How would you make an appointment with your mom if you need her help on your Chinese study?

5 你(nǐ)想(xiǎng)请(qǐng)你(nǐ)的(de)朋(péng)友(you)陪(péi)你(nǐ)一(yì)起(q)去喝奶茶，你(nǐ)怎(zěn)么(me)说(shuō)？ How would you invite your friends to go out with you to drink Boba tea? Hint“陪 accompany” and try to include “pre-invitation” strategy.)

APPENDIX III

Survey: Perceptions of Virtual Reality Technology enhancing Chinese learning experience.

(High-immersion VR)

1. Which interaction did you prefer while using the Mondly simulations?
2. Which interactions did you prefer while using the Uptale VR simulation?
3. Which of the following best describes the difficulty level of the Mondly scenario?
4. Which of the following best describes the difficulty level of the Uptale (student-shot) scenario?
5. Which of two (Mondly or Uptale) better reinforced the classroom materials, such as Making an Apology, Making requests, Invitations, and Expressing Gratitude?
6. The most challenging feature in the Mondly and Uptale VR simulations experience was _____.
7. Is the use of props and stage important to your learning and engagement with the video as could be seen in the Mondly VR experience?
8. Was the lack of Pinyin Hints (phonetic tool) a problem for students that do not know many characters?
9. Scholarship on language teaching has shown that an immersive experience provided by virtual reality can yield a better learning experience compared to a 2D textbook video. What features of the VR immersive experiences enhanced your learning of Mandarin Chinese?

Below are the three questions designed for those who shot 3D videos in this class.

1. How long did it take you to shoot the video compared to a regular classroom skit presentation?
2. Which of the following best describe your experience making the 3D video?
3. Preparing for shooting the 3D video enhanced my OWN learning of Mandarin Chinese.

Below are the open-ended questions:

1. Are there any features you like about the 360-video shooting or VR simulation design in your language learning experience? (Ex: the interactions such as multiple choice and voice recognition, how it differs from learning with a traditional 2D video, etc.)

2. Are there any suggestions you would like to make to further enhance your language learning experience? For example, did you prefer a specific way of answering questions during the VR experience (multiple choice, voice recognition, etc.)?