Faculty of Education – The University of Oxford

A Dissertation submitted in partial completion of the MSc Education

Applied Linguistics and Second Language Acquisition

August 2021

The Use of Virtual Reality Technology in Foreign Language Education: From Teachers' Perspectives

<u>Author</u>: My Ngoc Tran

College: St Cross College

Supervisor: Dr. Sophie Booton

Word Count: 19,980 words

Abstract

In the context of the COVID-19 pandemic, in which face-to-face interaction and engagement have been difficult, language teachers are seeking a technological solution that provides students with real-life interactions so they can practise language in both classroom and home settings. The use of virtual reality (VR) technology, which allows students to immerse themselves in real-life scenarios in which they can have conversations with native speakers of the target language, is believed to be an optimal solution to this circumstance. However, research on the use of VR (e.g., ImmerseMe) in foreign language education is still in its infancy. The current study therefore aimed to empirically examine teachers' perceptions of various aspects of VR, including the areas of learning supported by VR, the benefits and drawbacks of the technology, and any implementation challenges and enabling factors. A mixed-methods approach, specifically a sequential explanatory sub-design, was employed for data collection and analysis. The study involved 105 teachers teaching at secondary and higher education levels around the world. Quantitative data was collected via an online survey and analysed using a number of statistical tests, while qualitative data was collected through focus-group discussions and analysed thematically. The results showed that teachers perceived speaking and listening to be the learning areas that benefited most from the use of VR. The clearest VR benefits were found in contextualised learning, constructivist pedagogy, and student-centred teaching approaches, as well as in students' language knowledge and confidence. Fewer benefits were found for school grades and long-term memory retention. The biggest perceived drawbacks of VR included the time-consuming nature of lesson redesign, a lack of practice time and content teaching in class, its limitation in completely replacing classroom interaction, and occasionally chaotic learning environments. Perceived challenges to VR implementation included a lack of access to hardware facilities and to the VR program, a lack of teacher training and training materials, and an insensitive voice recognition feature. It is therefore strongly advised that schools, teachers, and program organisers cooperate to maximise VR's benefits, minimise its drawbacks, and address existing challenges. This research contributes to a body of work on the effective use of VR technology in foreign language education, especially in the context of the COVID-19 pandemic.

Keywords: virtual reality, language learning, teachers' perspectives, educational technology.

Acknowledgement

Foremost, I would like to gratefully acknowledge the great contribution and guidance of Dr. Sophie Booton, my second-to-none supervisor. The year of 2020 was tough because of the COVID-19 pandemic; however, I was the luckiest person in the world to have had Sophie, who had always been responsive, supportive, and understanding of my learning situation. Special thanks to my lecturers as well for always trying to bring us the best experiences you could give. Thank you – my amazing classmates, my lovely colleagues from the Oxford Department of Education, and caring friends from St Cross College – WE MADE IT!

I cannot express enough thanks to Mr. Scott Cardwell – the founder of the VR application *ImmerseMe* – for your continuous support during my thesis. I am also extremely grateful to the research participants – 151 teachers around the world – for your valuable contribution to this study.

The completion of this project could not have been accomplished without the support and encouragement of my close friends and family in Oxford, especially Prof. Phi Anh Phan, Dr. An Tran, and Mr. David Kovac. Thank you all for having me, loving me, and being there for me when I needed you the most.

My deep gratitude also goes to the Diverling Language School team and the SparkEd team. You are the reason for my beautiful days. I also want to deeply thank the Busuu Insights & Analytics team, especially Callum Ballard and Nora Matty, for providing me with amazing opportunities for my personal growth and career development.

Lastly, to my loving family – my father (Mr. Tran The Hung), my mother (Mrs. Nguyen Thi Thu Ha), my sister and brother-in-law (Mrs. Tran Thanh Hong and Mr. Do Hai Anh), my young brother (Tran Duc Huy) and Mr. Hoang Pham – thank you for your unconditional love and support. Without you, I am not the person I am today. Thanks to you, I can pursue my dreams and live the meaningful life that I want to live.

List of Figures

List of Tables

Table of Contents

Chapter 1: Introduction

1.1. The Role of Digital Technology in Language Learning

In recent years, with the pervasiveness of the internet and the evolution of information and communications technology (ICT) that aims to support learning activities, students' experiences with new languages are no longer limited to textbooks and traditional learning methods (Lin & Lan, 2015). Teachers and learners alike now have access to technological tools that facilitate the acceleration of their language teaching and learning processes (Gander et al., 2013). Eighty-six percent of US secondary teachers agree that it is crucial to use digital tools in language classrooms (National Center for Education Statistics, 2015). In 2019, 65% of US primary and secondary teachers reported that they used digital learning tools on a daily basis; in May 2021, this number had risen to 87% (National Center for Education Statistics, 2019). In Australia, a national report in 2015 showed that 57% of K-12 students (and two thirds of high school students) used digital technology for language learning in school on a daily basis (Thomson, 2015). Given this popularity, the impacts and effectiveness of technology integration into foreign language education have been a key topic of discussion among researchers, educators, and policymakers.

In the field of educational technology, Computer-Assisted Language Learning (CALL) is one of the most prevalent fields of research examining the effects of computer applications on language teaching and learning (Gillespie, 2020). Research into CALL includes human-computer interaction, instructional technology, second and foreign language acquisition, and pedagogical concerns (Parmaxi & Zaphiris, 2017). Most CALL studies have demonstrated that computer-based programs (e.g., CornerStone) and language learning software (e.g., Duolingo, Busuu) benefit both teachers and language learners in various ways. Specifically, a large-scale meta-analysis by Means et al. (2010) examined more than 1000 empirical studies carried out between 1996 and 2008 on the effects of online learning via computers on learning outcomes. These authors found that, on average, second language learners in online learning conditions performed better than those who received face-to-face instruction.

Indeed, the majority of early literature focused on examining the benefits of computer-based technology on language teaching and learning (Burston, 2015). In recent years, with the introduction of mobile devices, including smartphones and tablets, the scope of educational technology research has been expanded. The field now includes Mobile-Assisted Language Learning (MALL), which examines the impacts of mobile devices on foreign language education (Peterson, 2011). O'Brien's (2009) comparative study demonstrated that MALL has positive impacts on students' learning outcomes, enhancing class interactivity, language confidence, and automaticity. Similarly, a large number of papers have highlighted the time- and cost-effectiveness of MALL compared to other modes of learning (Abdous et al., 2009; Stockwell, 2010). Demouy et al. (2016) further suggested that the process of teaching and learning foreign languages has been able to transcend time and space limitations with the convenience, mobility, and effectiveness of mobile learning. In conclusion, it is clear that digital technology plays an increasingly crucial role in foreign language education.

1.2. The Emergence of Virtual Reality in Foreign Language Learning

With the increasing use of computers and mobile devices in classroom settings, together with the development of wireless communication and multimedia environments, teachers began to seek educational software that would optimise learners' experiences in online environments. This resulted in the use of virtual reality (VR) technologies in language learning contexts, often known as virtual learning environments (VLEs). VR was originally developed as a gaming tool in the 1960s to allow players to perceive virtual worlds realistically (Sutherland, 1965), and only started to gain popularity among the education community in the 1990s (Cooke-Plagwitz, 2008). Currently, 96% of universities and 79% of colleges in the UK are using VR in some capacity (UKAuthority, 2019). In the US, 97% of K-12 students are familiar with the term VR and 92% would like to study with VR in schools (Muhanna, 2015). In the 2018 Australia report, seven out of ten language teachers expressed interest in using VR to simulate experiences that were relevant to the class materials (Vassallo & Warren, 2018). Indeed, VR is regularly being used by both teachers and students for various subjects and courses.

VR is a system that brings real-life experiences to its users. Specifically, it simulates topography, movement, and physics so that users in VR computer-generated settings feel immersed in their simulated environments without limitations of time and space (Mayer, 2009). VR classifications are based on the levels of interaction and immersion that are provided to users, including *desktop VR (D-VR)* and *immersive VR (I-VR)*. D-VR is primarily non-immersive. In D-VR, participants control and manipulate the VR environment using a computer screen and traditional keyboard, without the use of headsets (Lee et al., 2010). This category is often referred to as *basic VR systems* (Muhanna, 2015). I-VR, *immersive* or *enhanced VR systems*, on the other hand, are multi-modal in nature since they aim to provide users with a sense of immersion in their environment with the aid of auditory simulation, 3D visuals, and in some cases, limb proprioception (Freina & Ott, 2015). The VR-based language learning program ImmerseMe, which is used as the main tool in the current study, offers both desktop VR and immersive VR experiences. Users can select the VR mode: non-immersive, partly immersive, or fully immersive, depending on the availability of 3D glasses and physical spaces in classroom settings.

Since 2010 and the introduction of smartphones, which can be used as VR displays and modern VR glasses (e.g., Google Cardboard and Samsung Gear; see Figure 1), VR technology has become more accessible and affordable (Cochrane, 2016). From that time onwards, VR technology and its impacts on foreign language education began to receive more attention from educational researchers and practitioners, many of whom attempted to integrate VR into classroom settings. Since then, the term

VR has also appeared more frequently in the literature and has been referred to using various terms: *desktop virtual reality* (Lee & Wong, 2014); *virtual reality environments* (Lin & Lan, 2015); *screen-based virtual reality environments* (Ausburn & Ausburn, 2008); *partially or semi-immersive VR*; *fully immersive VR*; *mixed reality* (Wu et al., 2013); and *virtual reality-based environments* (Cardwell, 2017).

Figure 1: *The History of VR Glasses.*

Note. Reproduced from *How Digital and Virtual Life Trapped in Visuality and New Media Systems Affect Teaching and the Learning Process* (p.110), by G. Molna, 2016, Peter Lang Publisher. Copyright 2016 by Peter Lang Publisher.

In 2020, the use of VR technology in educational settings has been discussed extensively in the context of the COVID-19 pandemic. By the 30th of March, 2020, 87% of the world's schools had closed, forcing 1.5 billion students into home-based learning (Winthrop, 2020). This circumstance led to a surge in demand for the use of education software that can strengthen the connections between teachers and students as well as reinforcing their home learning experiences (Winthrop, 2020). VR technology is believed to be an optimal solution to distance education, and especially in foreign language education, which requires a high level of social interactivity and authenticity (Hamilton et al., 2020).

Despite their usefulness and popularity, research into VR tools and their effectiveness in foreign language education is insufficient. Specifically, the most recent systematic review by Parmaxi (2020) examined the benefits and limitations of using VR as an educational tool in language classrooms, as well as its pedagogical applications. The paper analysed and synthesised papers from 17 high-impact journals and conferences in the field of CALL and 26 scholarly manuscripts retrieved from 2015 to 2018 based on the following elements: (1) technologies used, (2) language learning settings, and (3) duration of educational activities. The researchers argued that research on VR is insufficient to yield useful insights regarding the use of VR in language teaching and learning.

1.3. Research Rationale

In the context of the COVID-19 pandemic and the increased demand for VR use in education, especially foreign language education, this study aims to empirically examine the effectiveness of VR technology from teachers' perspectives. This study used both quantitative and qualitative methods to investigate the perceptions of current users and non-users of ImmerseMe, a VR-based tool currently being used in a number of K-12 schools worldwide. Teachers in the study were asked about their expectations and experiences regarding the benefits and drawbacks of VR, as well as the areas of learning they thought were best supported by the use of VR. It is hoped that the findings will enrich our knowledge of VR use in language teaching and learning. The study also identifies the key challenges and the main factors in the effective use of VR technology in foreign language education. It is expected that the research results will provide researchers and educators with pedagogical insights to address any potential issues during the process of VR implementation and usage. The study also aims to determine whether there are differences in the perceptions of current VR users and non-users. All of these findings will lay a foundation for future research to examine how teachers' perceptions, behaviours, and decisions might evolve throughout the VR implementation process.

Chapter 2: Literature Review

2.1. Introduction

This chapter will review relevant theoretical background and empirical studies in the field of VR research, laying the foundation for the current study. The research questions for the study are presented at the end of the chapter.

2.2. Virtual Reality Research: Theoretical Background

2.2.1. Sociolinguistic Acquisition Theory

According to the sociolinguistic theory of second language acquisition (SLA), meaningful language learning occurs in conjunction with society, culture, and personally relevant life experiences (Eun & Lim, 2009). The social context and social interaction mediate the process of language learning and thus play important roles in the SLA process (Ellis, 2008). Furthermore, the contextual and non-linguistic cues from real-life conversations can reduce the level of stress in language acquisition, allowing foreign language learners to comprehend and retain knowledge more effectively (Cipresso, 2015). The implementation of VR technology into foreign language education provides foreign language learners with context-based learning environments involving cultural elements (via real-life scenarios) and contact with native speakers of the target language, conferring these social and cultural aspects of language acquisition. Serrano et al. (2011) further argued that context-based learning is not only beneficial to vocabulary acquisition, but also enhances learners' pragmatic competence, their ability to use linguistic knowledge appropriately in specific sociocultural contexts. In conclusion, existing literature in the field of sociolinguistic acquisition theory suggests that the use of VR is likely to benefit learners' social and cultural development and pragmatic competence during the language acquisition process. This idea was used as a theoretical foundation in this study in the design of questionnaires for the outcome measure *VR general benefits.*

2.2.2. Authentic Learning and Constructivist Learning Theory

As well as context-based learning opportunities, VR technology also offers authentic learning, which allows students to explore, discuss, and meaningfully construct concepts and relationships in the context of real-life problems that are relevant to the learners (Amineh & Asl, 2015). A number of studies, including Alam (2017) and Bhattacharjee (2015), have shown that authentic learning helps students to integrate novel knowledge into their existing cognitive system. McGarrell and Alvira's (2013) empirical study of 113 EFL students also indicated that students especially benefit from learning with VR, since it allows them to 'envision places, objects or even concepts they have never had access to' (p.113). This study suggested that VR-facilitated authentic experiences might foster stronger comprehension and deeper knowledge.

Solomonidou (2016) further demonstrated that authentic learning is in line with constructivist learning theory, which asserts that learners should be active in their learning environments and use their previous knowledge to acquire novel knowledge (Huang et al., 2010). Huang and Liaw (2018) suggested that in regular classroom settings, meaning is constructed in conversations between the teacher and learners. In VR settings, learners construct meaning in foreign languages through a productive interaction between learners and the learning environment, in which learners use their existing knowledge of the target language and culture to acquire linguistic and non-linguistic information. In other words, VR technology is a key driving force behind the adaptation of social constructivist theory to the classroom context (Burkle & Meredith, 2008). A number of researchers have explored and discussed the relationships between VR and constructivism, including Chen et al. (2004), Madathil et al. (2017), and Ozkan (2017). To conclude, literature with a foundation in the constructivist theory has suggested that the use of VR in foreign language education might benefit students' authentic learning. The benefits of VR in other learning approaches will be further explored in section 2.3.2, where empirical studies on VR affordances are discussed in detail.

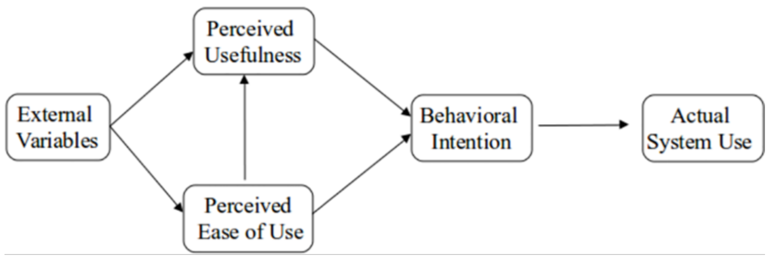
2.2.3. Technology Acceptance Framework I: Technology Acceptance Model (TAM)

Implementing a new technology in classroom settings is a complex process. Several theoretical models have been proposed to examine the process in which individuals accept or reject the use of a new technology (Ajzen, 1985; Ajzen & Fishbien, 1980). As there are no established theories that specifically examine the acceptance of VR technology in foreign language classroom settings, the current study employs two existing frameworks that are commonly used to evaluate various kinds of technology. The main aim of these frameworks, including the Technology Acceptance Model (TAM; Davis et al., 1989) and the Concern-Based Adoption Model (CBAM; Hall & Hord, 2006), is to explain users' attitudes and intentions during different phases of technology adoption in classroom settings. The following sections elaborate on the mechanisms described in each model and how they can be used as a theoretical foundation to explain teachers' perceptions at different stages of VR integration.

First, the Technology Acceptance Model (TAM) framework, introduced by Davis et al. (1989), was one of the earliest models to examine users' acceptance of a new technology. King and He's (2006) meta-analysis of 88 research papers suggests that TAM is a valid and robust model for evaluating the use of emerging technology in classroom settings. In contrast to other behavioural assessment tools usually employed for longitudinal studies, TAM can be used to design cross-sectional studies from a usability perspective. This makes it an ideal tool for the current study, which aims to examine teachers' perceptions at a single point in time.

According to the Technology Acceptance Model (TAM; see Figure 2), the *actual system use*, or the likelihood that someone will use a new technology, is derived from three main processes and four elements. The first process refers to how *external variables,* often known as explicit reasons, affect users' *perceived usefulness* and *perceived ease of use*. For instance, external reasons, such as colleague recommendation or teacher training, might affect teachers' perceptions of the technology's effectiveness in improving their teaching performance or in their evaluations of how user-friendly the new technology is (Davis et al., 1989). Specifically, Lederer et al. (2000) suggested that in the case of insufficient teacher training or lack of training materials, teachers might find the new technology ineffective and difficult to use. Other studies (Deslonde & Becerra, 2018; Tarhini et al., 2015) further highlighted that the level of difficulty when using a new technology can include both technical and non-technical aspects. Technical issues are associated with technical support, internet access, and equipment availability. Non-technical issues include teacher reluctance, lack of computer literacy, issues with classroom management, and inappropriate pedagogical approaches (Sumak et al., 2011). Any of these elements might cause teachers to hesitate when adopting a new technology, decreasing their motivation to implement it. In the context of the current study, these elements can be understood as potential challenges that teachers might encounter at the beginning of the implementation process, and thus were drawn upon in the design of questionnaires for the outcome measure *implementation challenges*.

Figure 2: *The Technology Acceptance Model (TAM).*



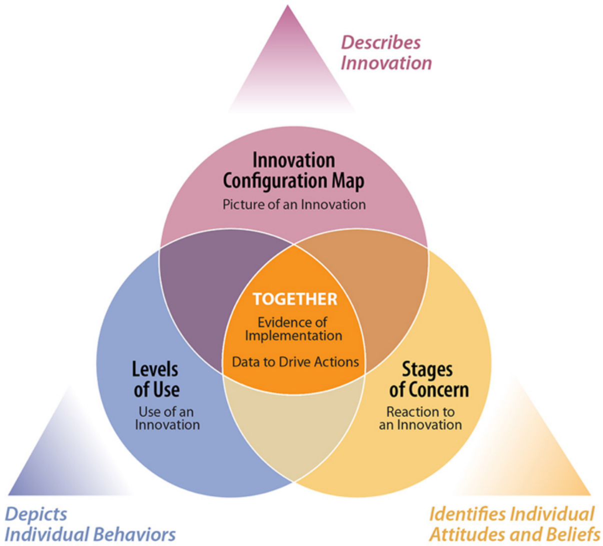
Note. Reproduced from 'User Acceptance of Computer Technology: A Comparison of Two Theoretical Models' by F.D. David, R.P. Bagozzi & P.R. Warshaw, 1989, *Management Science*, 35(2), pp. 982-1003. Copyright 1989 by INFORMS.

After going through this stage of perception formation, teachers have to decide whether to use the technology in their classrooms. This is captured in the second process, which moves from perceptions to *behavioural intention.* During this process, teachers are encouraged to explore the new tool, consider the factors that make for effective use of the tool, and assess whether the classroom conditions allow for the *actual system use* (Lederer et al., 2000). The nature of this phase suggests that as well as implementation challenges, it is also important to examine *implementation enabling factors,* since they directly affect teachers' final decisions about whether to implement the technology. More details about the factors in the effective use of technology in classroom settings will be discussed in section 2.3.5 below.

2.2.4. Technology Acceptance Framework II: Concern-Based Adoption Model (CBAM)

The Concern-Based Adoption Model (CBAM), developed by Hall and Hord (2006), consists of three diagnostic dimensions (see Figure 3) with an aim to provide techniques and tools that synthesise users' concerns during the process of technology implementation. This model goes one step further than the original TAM model, which only examines the perceptions of non-users and brand-new users and their likelihood of using a new technology, as it helps to identify specific user concerns associated with different levels of technology use (i.e., non-users and current users).

Figure 3: *Concern-Based Adoption Model (CBAM).*



Note. Reproduced from *Implementing change: Patterns, principles, and potholes* (p. 269), by G.E. Hall & S.M. Hord, 2006, Pearson Education. Copyright 2006 by Pearson Education.

In this model, the *Innovation Configuration Map* refers to a user's expectation when using the new technology. *Levels of Use* refers to the user's level of exposure to and experience with the technology. *Stages of Concern* (SoC) refers to the main concerns that users have towards the technology. Cheung and Yip (2005) argued that all three dimensions are closely intertwined. Specifically, depending on a teacher's level of exposure to the novel technology, that teacher's reactions and decisions will fall into certain SoCs or hierarchical stages, including *Evaluation*, *Information*, *Management*, and *Consequence*. *Evaluation* is the initial phase of technology acceptance and is usually experienced by new users. Mayer (2014) demonstrated that during this stage, teachers evaluate the general use of the technology in their pedagogical context and are likely to seek information about lesson design. Teachers also seek to understand how to manage students, resources, and facilities, as well as the effectiveness of the technology in terms of learning outcomes (Pritchard, 2017). During this phase, teachers are forming their perceptions of different aspects of the new technology, including the benefits, drawbacks, and areas of learning supported by the new technology (Tsai & Chai, 2012).

The next three stages, *Information*, *Management*, and *Consequence*, usually occur for teachers with higher levels of exposure to the technology. During these phases, teachers refine lesson design, learning objectives, and material distribution, as well as thinking about how to manage the class effectively and consider the consequences of the technology's implementation. Teachers' perceptions of the technology might change, depending on the technology's effectiveness and ease of use (Mayer, 2014). However, no empirical evidence has been provided for this model that shows the differences in user perceptions at different levels of usage. The current study, thus, aims to determine whether there are differences in the perceptions of non-users and current users regarding different aspects of VR technology.

Ertme (2005) further suggested that during this process of refinement, teachers might encounter three levels of barriers. First-order barriers include time, expertise, accessibility, resources, and technical support. Second-order barriers include factors leading to frustration from teachers and students, and third-order barriers relate to struggles in redesigning the curriculum and lesson systems. Compared to the TAM model, this framework provides more in-depth insights into the potential challenges that teachers might encounter, and thus can be used as a resource to design the questionnaire for the outcome measure *implementation challenges* in this study. It should be noted, however, that both models fail to justify whether teachers' perceptions of challenges and implementation enabling factors change depending on their levels of technology exposure (Tsai & Chai, 2012). Therefore, the current study will also examine whether there are differences in the perceptions of non-users and current users regarding VR implementation challenges and enabling factors.

In conclusion, existing frameworks suggest that there are many factors affecting teachers' decisions and perceptions surrounding new technology usage. Specifically, a number of external reasons (e.g., colleague recommendation, teacher training) and levels of technology usage (e.g., non-users, recent users, and long-term users) might affect teachers' perceptions of specific aspects of a technology. The current research will therefore provide empirical evidence to confirm these implications with application to VR technology specifically.

2.3. Empirical Research on Virtual Reality in Foreign Language Education

There are many aspects of VR technology upon which teachers might have an opinion, including the *areas of learning supported by the use of VR*, *general VR benefits*, and *drawbacks,* as well as the main *challenges to VR implementation*. This section will review the existing literature to identify relevant examples of each of these aspects.

2.3.1. Areas of Language Learning Supported by VR Technology

In general, while teachers widely acknowledge the benefits of VR technology for speaking skills and pronunciation, there has been a limited focus on the affordances of VR technology in relation to other skills, including writing and the receptive skills listening and reading (Huang & Liaw, 2018).

Productive Skills. A number of studies demonstrated that using VR technology in language teaching improves learners' productive skills, including speaking and writing. Alemi and Khatoony's (2020) study of 18 low-intermediate English language learners (6-12 years old) suggested that VLEs, which allow real-life interactive communication with native speakers of the target language, have significant impacts on learners' speaking confidence and thus on their overall performance. Huang et al. (2010) added that the instant responses required by the VR applications reinforce learners' encoding process, sentence production, word clarity, and idea organisation. However, this claim was not empirically tested until Yamazaki's (2018) study on the use of VR tools in learning Japanese. Yamazaki's findings showed that when 11 high school students who were learning Japanese as a second language were immersed in the 3D virtual world of Tokyo, they not only acquired vocabulary and sentence structures faster than a control group who learned Japanese through textbooks, but they also developed contextualized communicative competence, or the ability to communicate effectively in different contexts. Specifically, the quantitative analysis showed that the VLE improved learners' acquisition of incidentally encountered vocabulary, kanji pronunciation, and vocabulary interpretation. The qualitative analysis showed an enhancement in participants' acquisition of various communicative competencies specific to the context, including persuasive talk, audience awareness, and collaborative communication. In short, research appears to show that the use of VR in language teaching benefits students' communicative competence. In the current study, this premise was used to design questionnaires for the outcome measure *areas of learning supported by VR technology*.

In terms of writing, Dolgunsoz et al. (2018) conducted an empirical study on 24 Year 10 students who learned English as a Foreign Language (EFL) to determine the effects of VR experience on EFL writing skill development. The participants attended a semi-structured interview before watching a 2D traditional video and a VR video during different time periods. After completing a short writing task, students were interviewed again to discuss their experiences. These EFL learners were found to be already aware of VR technology, and the majority of them enjoyed watching VR videos. The VR experience, however, did not affect their short-term writing performance, and the students also complained about technical limitations, including low-quality video and physical discomfort. However, it should be noted that the VR technology was only implemented for two weeks; the research did not examine the long-term effects of VR on writing skills. Indeed, research on the effects of VR on students' writing is still inconclusive. Therefore, the current study seeks to empirically substantiate the benefits of VR on writing performance from teachers' perspectives.

Receptive Skills. There is a significant lack of research on the receptive skills side of VR technology (Soto et al., 2020). Tai and Chen's (2021) empirical study was among the first to examine the effects of VR programs on EFL listening comprehension. The study was conducted with 27 secondary students (7th grade) in Taiwan who were assigned randomly to an experimental VR environment or controlled environment (traditional video watching). The students in the VR environment played the VR app using a mobile-rendered head-mounted display, while the control group watched a video of the app on PC screens. Students immersed in the VR environment were found to show significantly higher listening comprehension and retention compared to the control group. The follow-up interviews indicated that students in the VR environment found the VR-assisted listening engaging and beneficial. It should, however, be noted that the students in this research may have gained comprehension from the social and contextual cues in the video images and content rather than listening to the dialogues. Therefore, the research findings, in which students in the VR environment had better comprehension than those in the control environment, may confirm that VR content simply offers participants more engagement, but does not necessarily benefit their listening skills specifically. It is clear that findings regarding the effects of VR on listening comprehension are still tentative and more empirical studies are needed.

There is no empirical evidence regarding the impact of VR on reading. This means that apart from speaking, research on the effects of VR on writing, listening, and reading has been insufficient for any conclusion to be drawn. Therefore, the current study aims to examine teachers' perceptions of the effects of VR on all four skills (i.e. speaking, writing, listening, and reading) to replicate and extend the existing literature, as well as to build a foundation for future research on the benefits of VR technology to receptive skills.

2.3.2. Benefits of VR Technology

Learning Approaches. As mentioned in section 2.2.2, existing theories suggest that the use of VR in language teaching might benefit students' learning approaches (e.g., authentic and constructivist learning). Empirical studies indicate that VLEs may facilitate other learning approaches, including active learning, interactive learning, and learning by doing, which are normally not prioritised in non-VR learning environments (Utami et al., 2021). Specifically, Kluge and Riley (2008) suggested that

VLEs, which allow students to choose the types and modes of practice when initiating conversations with native speakers, give them a sense of independence and learning autonomy. Sharma et al. (2013) further highlighted that in the long term, this type of environment can enhance students' self-initiated learning skills, independent learning capacities, and control over their process of knowledge construction. In other words, in VR-integrated environments, students' roles change from passive receivers into active constructors of knowledge (Ampatzoglou & Chatzigeogiou, 2007). These studies do, however, consider only students' perspectives. The current research aims to balance existing research by examining VR from a teacher's perspective.

Teaching Approaches. A number of researchers have shown that VLEs encourage student-centred teaching approaches instead of the traditional teacher-centred approaches, and that this is likely to improve students' experiences and their learning outcomes (Burkle & Meredith, 2008; Lukmani, 2012). Specifically, Clarke et al. (2006) demonstrated that the ideal pedagogy model for VLEs positions teachers not as knowledge holders or distributors but rather as facilitators of knowledge. This is because in VR-integrated environments, students are given opportunities to construct their own knowledge based on their existing understanding of the topic. The teacher's role thus changes from didactic and transmissive to more facilitative, communicative, and collaborative. Jensen and Konradsen (2018) agreed, arguing that the main purpose of VR-integrated teaching is not to focus on pre-specified knowledge or to develop an understanding of predetermined concepts, but rather on individual thinking and the construction of meaning, often known as constructivist pedagogy. In addition, Chuah et al. (2011) argued that VR-supported environments facilitate the integration of classroom activities distinct from textbook-based teaching. These authors elaborated that the scenario-based learning and pronunciation practice facilitated by the voice recognition technology in VR tools (e.g., ImmerseMe) is helpful in diversifying learning materials and in-class activities. In summary, existing literature indicates that VR-facilitated environments encourage the diversification of not only teaching approaches and pedagogies (e.g., student-centred approaches and constructivist pedagogy) but also learning materials and in-class activities. The current study aims to empirically confirm these benefits of VR from teachers' perspectives.

General Cognition and Language Acquisition. Early literature argued that VLEs provide contexts for the enhancement of multiple human intelligences, including verbal/linguistic, logical, auditory, spatial, interpersonal, and intrapersonal (Kohlberg, 1968). Harrow (1972) highlighted the possibilities of VR technology and its scenario-based approach provides for cognitive and affective development. However, these early studies were not empirically verified. It was not until recently that empirical studies were conducted to quantify the benefits of VLEs on students' general learning outcomes and language acquisition. For instance, Mroz's (2015) empirical study of five L2 French students aged 15-18 found that the use of a VR tool, in which students communicated in the target language to solve a social problem, led to significant language gains and use, as well as to improved critical thinking skills. Hussein and Natterdal's (2015) comparative study of 25 language teachers produced similar results, with teachers commenting that repeated practice of VR interactive tasks improved their students' English test scores. Indeed, empirical studies indicate that the use of VR in language classrooms might benefit students' language acquisition performance. In the current study, this premise was used to design questionnaires for the outcome measure *VR general benefits*.

Language Confidence and Intercultural Awareness. A meta-analysis conducted by Schwienhorst (2002) of 35 empirical studies found that the use of VR technology in language learning resulted in higher levels of language confidence and intercultural awareness compared to conventional, didactic, and textbook-based learning environments. This conclusion was supported by a number of subsequent empirical studies (Dalgarno et al., 2004; Ketelhut & Nelson, 2006), which found that practice with pre-recorded native speakers on VR tools is most beneficial to shy and introverted students, who would normally struggle to start conversations with real-life speakers. In other words, in VR-facilitated environments, students have opportunities to build their language confidence slowly and without suddenly leaving their comfort zone (Loewen & Sato, 2018). The current study aims to empirically substantiate the benefits of VR for developing students' language confidence and cultural appreciation, as seen from a teacher's perspective.

2.3.3. Drawbacks of VR Technology

In contrast to its widely recognised benefits, the drawbacks of VR technology are not well-established. Prensky (2016) proposed that the main drawbacks of VR are (1) limitation in completely replacing peer-to-peer interaction, (2) students' over-reliance on technology, (3) lesson design and time allocation, and (4) chaotic learning environments. However, so far only lesson planning issues have been empirically tested, and most recent studies have examined the drawbacks of VR at university level rather than at K-12 levels (Southgate & Smith, 2017).

Lesson Planning. Berge (2008) showed that in terms of classroom reality, the content available on VR applications does not always match the lesson content of the textbooks used by different schools. This is because existing educational VR content is typically designed for stand-alone learning experiences, with few possibilities to adapt the level of difficulty and the variety of exercises. This makes it difficult for teachers to logically incorporate VR content into their pedagogical planning. Jensen and Konradsen (2017) thus noted that if teachers want to employ VR technology, they are required to make significant adjustments to the curricula, including the possibility that they will have to produce their own VR-integrated content. In addition, Tan et al. (2016) claimed that the use of VR technology requires extended time allocations for technology setup and login, leaving teachers with significantly less time for content teaching. Fransson et al. (2020) also argued that teaching time is reduced and that lessons are disrupted in the case of sudden technical problems (e.g., system breakdown or internet connection problems). In the long term, this might affect students' learning input and outcomes. This study aims to empirically examine these lesson planning dilemmas from teachers' perspectives and propose appropriate pedagogical solutions.

Teaching and Learning Materials. The lack of teaching and learning materials has been one of the biggest challenges in the effective use of VR technology at university level (Alfalah, 2018) and in K-12 contexts (Minocha et al., 2017). These authors argued that opportunities to use VR technology are limited due to a lack of existing resources to assist learning and teaching activities in VR-driven classrooms. Blyth (2018) suggested that teachers need to design differentiated pronunciation drills and speaking practice activities for students of different levels to prepare for the VR practice. This is because VR tends to require a good level of pronunciation accuracy, and thus might be extremely challenging for weak students. Empirical studies on VR supplementary resources are still insufficient, meaning that no firm conclusions can be drawn regarding which types of teaching and learning materials are most useful in conjunction with VR use. The current study aims to address this issue by providing suggestions regarding the effective use of VR in language classrooms.

2.3.4. Key Challenges to VR Implementation in Language Learning

Since VR is a fairly new development in foreign language education, a number of challenges to VR implementation need to be addressed in order to achieve greater effectiveness and avoid potential negative experiences in the learning and teaching process. So far, two crucial implementation challenges that have been empirically examined in existing literature are (1) technology accessibility issues and (2) lack of teacher training.

Accessibility. A number of studies have indicated that the most common reason VR is beyond the reach of many schools has always been the financial difficulty of affording VR software and facilities that provide the desired immersion and interaction (Hafner et al., 2018; Sasinka et al., 2019). To provide a 360-degree virtual reality experience to every student, schools need to prepare two sets of necessary equipment, including the VR headset (e.g., Google cardboard) and the VR software installed on a digital device (e.g., a smartphone, tablet, or desktop). Utami (2021) argued that even though these devices have become much more affordable, they are still expensive per student for most schools. Moreover, the cost of both procurement and maintenance of a large number of devices has also made the mass use of this technology prohibitive (Chen et al., 2005). Chen (2016) further argued that the unavailability of hardware and software due to its high cost is not the only reason for the inaccessibility of VR technology. At many K-12 institutions, although students are provided with desktops, using VR in classroom settings can be a challenge because of poor resource organisation, poor hardware quality, and a slow installation process (Pantelidis, 2009).

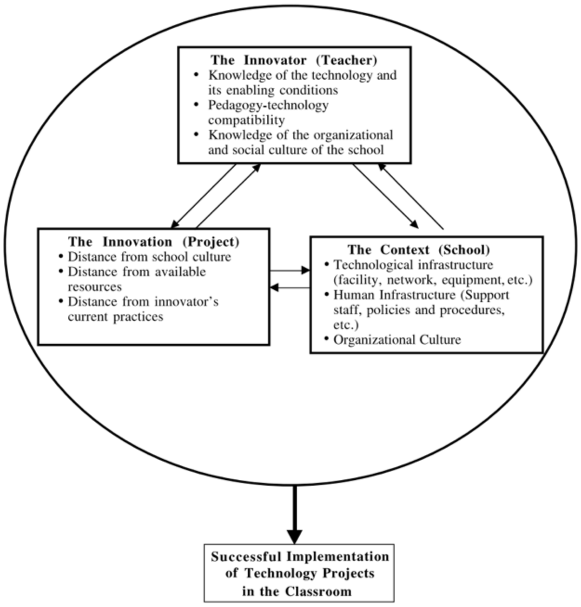
Lack of Teacher Training. For every new technology implemented in classroom settings, teacher training is one of the most important factors in implementation success (Cummins, 2000). This is especially true for the implementation of VR technology. Specifically, Oak (2018) and Kwon (2019) established that VR-integrated teaching requires advanced levels of digital literacy (e.g., software updates, profile logins) that most K-12 teachers do not have. Therefore, both program training and digital skills training are essential in VR-driven classrooms. Moreover, Parmaxi (2020) argued that teachers need support to overcome potential technical overload and computer anxiety to optimise both teachers' and students' VR experiences, further emphasising the importance of teacher training in VR implementation. In general, training programs and resources for VR use are extremely scarce for teachers of all levels, and no empirical evidence has been provided on the effectiveness of the current VR training (Utami et al., 2021). The current study aims to examine the usefulness of the training that teachers receive for the VR-based tool ImmerseMe, from there suggesting better approaches for effective VR training.

2.3.5. Factors Impacting the Effective Use of Technology

As mentioned in sections 2.2.3 and 2.2.4 (TAM and SoC frameworks), various factors affect teachers' decisions and attitudes regarding the use of the new technology, as well as their ability to use the technology effectively. While the former is addressed in the TAM and CBAM models (external variables and levels of exposure), this section will elaborate on the latter by discussing the main factors affecting the effectiveness of technology use in language classroom settings. According to Cuban (2001), a primary issue around the effective use of the new technology in educational contexts is the classroom in which the technology is used. Zhao et al.'s (2002) *Conditions for Classroom Technology Innovation* is the only relevant model to name salient factors in the effective use of technology in education and how these factors are interrelated. This model was empirically tested on 118 K-12 teachers. It delineates three segments (*the teacher*, *the innovation*, and *the context*) and nine factors required for technology implementation (see Figure 4). According to this model, the teachers need to (1) have knowledge of the technology and of the enabling factors, (2) ensure pedagogy-technology compatibility, and (3) understand the social culture of the school to adapt the technology appropriately. The school needs to (1) build the technology infrastructure, (2) build human resources through teacher training, providing adequate training materials, and (3) build an organisational culture that allows teachers to collaborate and assist each other during the implementation process. Finally, several problems need to be addressed before the technology is implemented, including (1) distance from the school culture, (2) distance from current resources, and (3) distance from teachers' practices.

In fact, a number of empirical studies have successfully exemplified the factors listed in this model. Cubukcuoglu's (2013) study of seven Turkish teachers examined teacher-related factors, demonstrating that the most necessary components of the effective use of Information Communication Technology (ICT) are teachers' digital literacy, time management, and pedagogical re-planning. It should, however, be noted that the sample size of this study is relatively small, which might reduce the reliability and generalisability of these results. Another study, conducted by Esfijani and Zamani (2020) on 180 secondary teachers, examined the main factors in the effective use of ICT in language education, demonstrated that students' experience and other student-related factors such as classroom engagement and learning autonomy can directly impact the effectiveness of teaching and thus must be addressed. Indeed, the list of factors in the effective use of technology in foreign language classrooms is inconclusive. This current study aims to empirically investigate the above factors, as well as proposing new factors relating to the effective use of VR technology.

Figure 4: *Conditions for Classroom Technology Innovation Model*



Note. Reproduced from 'Conditions for Classroom Technology Innovations', by Y. Zhao et al., 2002, *Teachers College Record*, 104(3), pp. 482-515. Copyright 2002 by Teachers College, Columbia University.

2.4. Teachers' Perspectives

The current study examined the use of VR technology from teachers' rather than students' perspectives for a number of reasons. First, Lee and Wong (2008) argued that it is crucial to take teachers' viewpoints into account in classroom-related research, since their actions directly determine the quality of teaching and learning. Palak and Walls (2009) agreed, adding that gaining insights from teachers' perspectives on the use of technology is essential, as their points of view explicitly impact their classroom planning, decisions, judgement, and the actions they perform on a minute-by-minute basis. Redmond et al. (2005) concurred with this statement, demonstrating that teachers' personal interest in using technology and their willingness to implement new technology are significant elements in the success of technology integration. Second, a number of studies have highlighted that teachers' attitudes and decisions in regard to technology implementation are strongly correlated with their students' approaches to learning and their learning outcomes (Trigwell et al., 1999). For instance, Judson (2006) argued that teachers with technology-driven mindsets are likely to adopt new technology for high-level uses such as student collaboration, critical thinking, and project-based learning. This active use of technology to enhance students' experience reflects student-centred teaching approaches, which allow students to seek meaning and gain understanding by making links within the content and constructing their own knowledge structures (Kern, 2006). This approach seems likely to enhance students' learning outcomes (Marton & Booth, 1997). In contrast, teachers with technology-resistant mindsets tend to implement technology only for low-level uses (e.g., visual aids), using technology as a reward for individual practices controlled by the teachers. These teachers tend to prefer teacher-centred approaches, for example encouraging students to memorise a given body of information for the purpose of recalling it later for assessment purposes (Palak & Walls, 2009). Slater (2018) demonstrated that this approach to learning is only beneficial for short-term assessment outcomes. Indeed, these studies suggest that teachers' decisions, viewpoints, and practices have direct impacts on the success of technology integration, thus justifying the scope of the current research.

2.5. An Example of a VR-based Language Learning Program: ImmerseMe

The tool chosen for this research was ImmerseMe – an interactive VR-based language learning platform that offers language lessons through one-on-one dialogues in 360-degree videos recorded in various social settings. The main purpose of choosing a sample VR tool rather than referring to all VR technologies was to help non-users, those who have never had any exposure to VR technology, imagine how VR works in language classroom environments. ImmerseMe was chosen for this study for four key reasons. First, 75,000 students and more than 1,000 teachers are currently using this platform in classroom settings, which afforded easy research recruitment. Second, ImmerseMe offers a range of learning activities – listening, speaking, reading, writing, translation, pronunciation training, dictation, and immersion – which makes the platform ideal for examining teachers' perceptions of the impacts of VR on different areas of learning. Third, this tool offers content, materials, and resources in nine languages, as well as instructional activities and lesson design guidelines, which allows us to examine the perspectives of teachers from different language backgrounds. Finally, even though the tool has been in use since 2016, research on its effectiveness and applications is still insufficient, meaning that there is ample room for further investigation.

In comparison to traditional methods, ImmerseMe places language in context by delivering dialogues with native speakers recorded at various locations in which the target language is spoken. The program also uses Google sound recognition software to recognise correct pronunciation. Correct answers allow the user to move forward in the lesson, while incorrect responses require the user to repeat the sentence until it is adequate. For instance, students can practice buying bread in a coffee shop in central Paris. The students need to speak with the cashier in French, making sure that their pronunciation is correct so the conversation can proceed. It is believed that this speech recognition feature, which enables learners to interact with virtual subjects, practice different sounds, and receive instant feedback, improves students' pronunciation and communication skills (Grant et al., 2013). However, this argument has not been empirically tested.

According to ImmerseMe (2019), the content is curriculum-led and designed by language educators, and lessons are tailored to match students' needs. For this reason, it is a suitable tool for learners of all levels who have no access to target language speakers outside of the formal instructional setting. For students at lower levels, ImmerseMe is an ideal tool for practising language accuracy. For higher-proficiency students, teachers can use ImmerseMe as a supplementary resource to help them explore native language and culture (Bajorek, 2019). While ImmerseMe has many pedagogically sound features, it also has some limitations. No written instructions or task explanations are provided to new users, which may be especially challenging for beginners and teachers with less advanced computer skills.

This research aims to assess teachers' attitudes and perceptions regarding the use of ImmerseMe in language teaching and learning to gain insights into the effectiveness and potential of VR technology in foreign language education in general. This is especially necessary in the context of the COVID-19 pandemic, in which the use of VR in language teaching has been strongly encouraged.

2.6. Research Questions

Research Question 1: What are teachers' perceptions of using ImmerseMe in language teaching? In particular, what areas of learning are supported by this VR tool, what are the general benefits and drawbacks, and what are the challenges and enabling factors relating to VR implementation?

Answering this research question requires quantitative and qualitative approaches that target teachers' general perceptions of various aspects of ImmerseMe and VR technology in general. First, quantitative analyses were conducted on the survey data to determine the perceived importance to participants of each of the outcome measures. Following this, thematic analysis was conducted on the textual data to further analyse teachers' responses.

Research Question 2: Do ImmerseMe current users and non-users differ in their perception of these aspects (areas of learning supported by VR technology, VR benefits and drawbacks, implementation challenges and implementation enabling factors)?

This second research question involves analysis of the quantitative data collected for the first research question, using MANCOVA and ANOVA tests to determine the differences in perceptions of ImmerseMe current users and non-users.

Chapter 3: Methodology

3.1. Introduction

This chapter provides an outline of the research methods used in the study. First, the chosen research design will be described. This is followed by information on the participants, including the criteria for inclusion and exclusion, the sample size, and participants' demographic background. The data collection instrument and related procedures will also be described. Subsequently, the methodological decisions regarding how to answer each research question will be discussed, with ethical considerations being addressed at the end of the chapter.

3.2. Design

This research is exploratory in nature, as it attempts to investigate the perceptions of teachers of a technology that has only recently begun to be used in foreign language education. Teachers' subjective perceptions formed the primary data for the study, and therefore an exploratory research technique was required. This study is also a natural experiment, in which two pre-existing groups, in this case ImmerseMe current users and non-users, are compared on a number of dimensions (Leatherdale, 2019). Due to the complex nature of the research, the research design involves a mixed-methods approach. More specifically, it employs a sequential explanatory sub-design in which quantitative research is followed by qualitative research (Creswell & Clark, 2007). In this case, the qualitative data is used in the subsequent interpretation and clarification of the results from the quantitative data analysis. The purpose of the mixed-methods design is not to 'cancel out' the downsides of each approach, but rather to integrate the most thorough and appropriate methods of examining the subject at hand (Tashakkori & Teddlie, 2010, p.8).

The quantitative research in this study is characterised by a within-subjects approach; that is, all participants are exposed to the same conditions and questionnaire items. The main purpose of this research design is to gain a holistic perspective on teachers' beliefs based on the quantitative data before gaining specific insights into their personal experiences and reasoning using focus-group discussions. Suh and Prophet (2018) stated that it is crucial to conduct qualitative research when examining teachers' perspectives, since it helps us to understand the ways of thinking, actions, and cognitions that underlie each decision that teachers make in their pedagogical practice. In light of this, a close-ended survey investigating teachers' perceptions of different aspects of VR was coupled with follow-up focus-group discussions, resulting in a reciprocal combination of both quantitative and qualitative paradigms. The instrument measured the perceptions of teachers (dependent variables) regarding different aspects of VR technology (independent variables).

3.3. Sampling

3.3.1. Sample Size Calculation

Before recruitment, a power analysis was conducted to calculate the ideal sample size. Using G-power software with α = .05, power = .8, numerator df = 1, effect size f = .25, and number of covariates = 4, the target sample size of this study was determined to be 128.

3.3.2. Recruitment Method

This research draws upon an opportunity sample: participants were selected from a target group to participate in the research study. The recruitment process lasted 30 days. Current users of ImmerseMe were primarily recruited via ImmerseMe's emailing list of 1,600 K-12 and higher education teachers, while non-users were recruited via both the mailing list and social media posts (Facebook, Instagram, and LinkedIn). At the end of the online survey, participants were asked to provide an email address if they wished to take part in one of the follow-up discussions. Interested participants were then contacted via email to arrange focus-group discussions.

3.3.3. Eligibility Criteria

The eligibility criteria for participants in the online questionnaire and follow-up focus-group discussions were as follows: (1) the participant needed to be a current foreign language teacher, and (2) they needed to be currently teaching at secondary and/or higher education level. Due to the COVID-19 pandemic, teachers' perceptions of e-learning and the roles of technology in language teaching may recently have shifted significantly. This research targets current teachers because their practices in relation to technology use in classrooms are likely to have recently been affected. Secondary and/or higher education levels were desirable because the ImmerseMe application is designed for Year 6 and above.

3.3.4. Participants

3.3.4.1. Online Survey

In total, 180 participants responded to the survey. Some were excluded due to incomplete responses (n = 10), ineligibility (n = 38), and late responses (n = 27). The number of valid participants for this study was $N = 105$, which was 23 participants short of the pre-determined target sample size. Due to time constraints, the data collection had to be conducted within 30 days. The target sample size would have been reached if more time was allowed. This shortage of participants may reduce the power of the results.

The participants were aged between 21 and 80 ($M = 40.30$, $SD = 12.75$). Of the participants, 83% were female and 17% were male, while 51% were current users of ImmerseMe and 49% were non-users. The average number of years of teaching experience was 14. In terms of teaching level, 83% of the teachers were currently teaching at secondary level and 17% were teaching at higher education levels. Participants resided in various parts of the world, with 41% (n = 43) in Australia, 32% (n = 35) in Vietnam, 7% (n = 8) in New Zealand, 10% (n = 11) in the United States, and 1% (n = 1) for each of the following countries: Ecuador, Iceland, Italy, Japan, Mexico, South Korea, Macedonia, and Turkey.

In general, participants agreed that digital technologies were useful ($M = 4.36$, $SD = 0.69$ on a scale of 1 to 5). Participants in this study had had considerable exposure to computer-based programs, with 51% of participants using computer-based programs on a daily basis and 40% using them more than once per week.

3.3.4.2. Focus-Group Discussion

There were a total of five focus-group discussions, involving a total number of N = 17 participants. The participants were aged between 23 and 70 ($M = 44.4$, $SD = 13.38$). Among those, 77% were female and 23% were male, while 64% (n = 11) were current users of ImmerseMe and 36% (n = 6) were non-users. Participants' average number of years of teaching experience was ten. In terms of teaching level, 77% of the participants (n = 13) were currently teaching at secondary level and 23% (n = 4) were teaching at higher education levels. Participants mainly resided in Australia (n = 7) and the United States (n = 4). Other participants resided in Vietnam (n = 2), New Zealand (n = 2), Japan (n = 1), and Venezuela (n = 1). Participants generally perceived that digital technologies were useful ($M = 4.56$, $SD = .51$). Forty-seven percent of the focus group participants used computer-based programs in classroom settings on a daily basis and only 20% of the participants used them less than once per week. Indeed, the participants in focus groups were representative of the larger sample in terms of gender, years of teaching experience, and teaching level. However, this was not the case for (1) the levels of usage (ImmerseMe current users were overrepresented, as compared to the larger sample), and (2) the country of residence (participants from Vietnam were underrepresented in the focus group sample).

These discrepancies may have led to contradictions between the quantitative and qualitative research results.

3.4. Materials

3.4.1. Online Survey

Due to the COVID-19 pandemic, the survey was conducted entirely online. It consisted of 61 items that aimed to investigate teachers' perceptions of specific aspects of VR (see Appendix G). The first part of the survey gathered participants' personal information, including age, gender, years of experience, teaching level, country of residence, perception of technology usefulness, and frequency of computer-based program use. The second part of the survey consisted of items targeting different aspects of VR technology with reference to the ImmerseMe application. The first VR aspect was *areas of learning supported by VR technology*. This covered the main learning areas that might benefit from the use of VR technology, or ImmerseMe specifically, in classroom settings (e.g., *VR improves students' writing performance*). Other aspects were: (1) benefits of VR (e.g., *VR promotes task-based learning and problem-solving*), (2) drawbacks of VR (e.g., *students might over-rely on technology and avoid face-to-face conversations*), (3) challenges to VR implementation (e.g., *unable to integrate VR technology into the curriculum*), and (4) implementation enabling factors (e.g., *long practice time*). The same items were presented to both current users and non-users of ImmerseMe.

The scales used to rate questionnaire items were adapted and developed from existing sources. Specifically, the first scale, which was created to measure four aspects including *areas of learning*, *VR benefits*, *VR drawbacks*, and *implementation challenges*, was a 5-point Likert scale in which 1 represented *strongly disagree*, 2 represented *disagree*, 3 represented *neutral*, 4 represented *agree*, and 5 represented *strongly agree*. This scale was adapted from the Technology Acceptance Model questionnaires (Venkatesh & Davis, 1996) and Mollaei and Riasati's (2013) questionnaire of technology adoption behaviours. This scale proved to be highly reliable in measuring the *areas of learning*, *VR benefits*, and *VR drawbacks* aspects, with $\alpha > .8$, but less reliable in measuring the *implementation challenges* aspect ($\alpha < .8$; see Appendix F). The metrics further showed that items under the *implementation challenges* outcome measure were not correlated, and the value of α did not change when outliers were removed. This suggests that the scale was not reliable for measuring the potential challenges of VR, which may be a limitation of the study. A plausible explanation for this low internal consistency might be that each teacher has a unique profile in relation to VR implementation challenges.

The second scale was adapted and developed from Dalgarno's (2010) modified questionnaires and Annetta et al.'s (2008) survey. This was used to rate items belonging to the *implementation enabling factors* aspect. This scale involved a 5-point Likert scale in which 1 represented *not at all important*, 2 represented *slightly important*, 3 represented *moderately important*, 4 represented *very important*, and 5 represented *extremely important*. The second scale showed excellent reliability, with $\alpha > .8$.

3.4.2. Survey Piloting

Before sending the questionnaires to all participants, a pilot study was run with five teachers who met the inclusion criteria but were not included in the study. Pilot participants were asked to provide feedback on the questionnaires in terms of length, language usage, word and sentence clarity, and other relevant aspects. The pilot study was also employed to measure the total amount of time taken to complete the survey. All of the pilot participants finished the survey within 30 minutes and provided feedback regarding ambiguous technical terms and definitions (e.g., digital technology), an unclear VR briefing, vague explanations and word clarity.

3.4.3. Focus groups

The focus groups were conducted online via Microsoft Teams. They took the form of heterogenous groups, in which participants from various backgrounds with different points of view gathering together for discussion. According to Halcomb et al. (2007), the tension that heterogeneity creates in group discussions may serve to uncover deeper insights in constructive directions. To form heterogeneous groupings, the researcher arranged participants of different backgrounds (e.g., age, years of experience, teaching levels, country of residence) into the same group and confirmed participants' availability. However, in reality, participants were rather grouped based on their time availability and locations. For instance, all of the teachers from the United States were grouped together because the US timezones are different from those of other participants' countries.

In total there were five focus group sessions, with two to five participants in each session. Each discussion lasted 30 minutes. Participants were invited to explain their survey responses and give real-life examples. The main purpose of the discussion questions was for participants who have used ImmerseMe to share their classroom experiences and non-users to share their expectations regarding the use of ImmerseMe in classroom settings. Due to time constraints, four out of six questions were chosen for each focus group, guaranteeing that each question would be answered by at least two groups. The list of questions was developed by the researcher (see Appendix N).

3.4.4. Focus Group Piloting

Before the scheduled focus group sessions, a pilot was run on three participants who had previously piloted the survey and were not included in the study. The pilot was intended to check the clarity of the questions and measure the length of the discussion. The pilot participants understood the questions clearly and the timing estimate was accurate, at approximately 30 minutes, therefore no change was made to the focus group instrument.

3.5. Procedure

The survey was administered via Qualtrics (https://www.qualtrics.com/uk/) and took around 30 minutes to complete. Participants were asked to read the information sheet, which outlined information about data privacy and confidentiality, before signing the consent form. They were then asked to provide personal information (e.g., gender, years of experience, teaching levels). Before beginning the main questions on different aspects of VR technology, participants were provided with a brief explanation of VR technology and the ImmerseMe application, along with a demonstration video showing the main functions of ImmerseMe. At the end of the survey, participants were asked to provide their email address if they wished to take part in a follow-up discussion to discuss their responses. These respondents were then contacted and asked to sign the focus group consent form before taking part in the discussions.

Microsoft Teams was used to conduct the online discussions, with video recording and automatic transcription features being used. Two hours before the discussion, participants received a reminder to attend the meeting via the sent link. Participants were required to turn on the camera and videos were recorded with consent. At the beginning of the focus group session, the researcher welcomed the participants, provided guidelines, and started the discussion. At the end of each session, video recordings and transcripts were uploaded automatically via Microsoft Stream, which was then used for data analysis.

3.6. Data Analysis

3.6.1. Research Question 1

To answer RQ1, which aimed to determine teachers' perceptions of different aspects of VR technology (i.e., areas of learning, VR benefits, VR drawbacks, implementation challenges, and implementation enabling factors), two analyses were run using the Statistical Package for the Social Sciences (SPSS) version 22.0. First, repeated measures ANOVAs were conducted to determine whether there were differences between item means. It is important to note that the homogeneity of the variance (sphericity) must be tested as a condition for repeated measures ANOVAs. In cases in which the ANOVA result is significant, a second analysis, a pairwise comparison with Bonferroni correction, was conducted to determine which item means were significantly different from the others. In the case of the outcome measure *VR benefits,* instead of comparing all items, items under each sub-scale (teaching approaches, learning approaches, and linguistic acquisition) were compared to each other. This is because this outcome measure had the highest number of items (14 items) and thus, the number of pairs when comparing all items exceeded 150.

3.6.2. Research Question 2

To determine whether there were differences in the perceptions of VR-based tools between current and non-current users of ImmerseMe, a MANCOVA was conducted followed by ANCOVAs. Before running the test, t-tests and chi-square tests were used to check whether there were differences between the two groups (ImmerseMe users and non-users) in terms of any demographic characteristics. Subsequently, a MANCOVA test was conducted to determine whether there were differences between the two groups for each outcome measure after controlling for these demographic characteristics. Finally, individual ANCOVA tests were run for each outcome measure to determine where the differences lay. It should be noted that data was missing for participants on the following key variables: age ($n = 2$), years of experience ($n = 2$), and teaching level ($n = 1$). Since the covariate tests (see Chapter 4) showed that age and years of experience are controlled covariates for the statistical analyses, the missing data on age and years of experience was managed with listwise deletion, where the missing cases were dropped from the analyses.

3.6.3. Qualitative Research

3.6.3.1. Thematic Analysis

Thematic analysis was chosen for the qualitative analysis with a goal to identify repeated themes and patterns in the dataset that were important and useful. These themes were then used to address the research questions and shape the discussion about certain issues (Braun & Clarke, 2006, 2013). In this research, thematic analysis was used to analyse the transcriptions of the focus-group discussion sessions. After data familiarisation, patterns and themes were identified both deductively and inductively. Specifically, a deductive orientation refers to the top-down coding and theme development that is primarily directed by existing concepts and ideas (Crabtree & Miller, 1999). In this case, some themes were developed deductively from the existing literature and from the questionnaire used in the current study. In contrast, an inductive orientation refers to bottom-up theme development that is directed by the content of the data; that is, any topics that have not been included in previous or current research (Charmaz, 2006).

3.6.3.2. Transcription and Coding Protocol

The transcription was automatically executed by Microsoft Teams and downloaded for formatting and proofreading. The processes of coding and theme development were conducted by two raters, the main researcher and a second rater. A spreadsheet that contained columns of all relevant quotes from the transcription and rows of theme names, was created. The researcher then counted the number of quotes that contained individual themes. For instance, for the outcome measure *learning areas supported by VR technology*, there were 12 quotes that mentioned the *speaking and conversational skills* theme. The lists of themes were continuously reviewed during the process of coding, with more themes added inductively where necessary. Once finished, the raw table of quotes and established themes, along with explanations of the themes, was sent to the second rater, who then reviewed the themes and counted the relevant quotes. Subsequently, a discussion between the researcher and the second rater took place to discuss the suitability of the themes and any mismatching quote counts. The percentage agreement between the two raters was 81%, indicating a high level of inter-rater reliability.

3.6. Ethical Procedures

The data collection, transfer, and storage for this research were conducted in accordance with the University of Oxford Policy on Data Collection (University of Oxford Central University Research Ethics Committee, 2016). After receiving ethics approval (See Appendix E), participants were contacted via email or social media (see Appendix A). Participants were then provided with an information sheet that outlined data privacy and other study information (see Appendix D) prior to participation. Participants were then asked to give consent before completing the online survey and/or attending the focus group discussions. Following data analysis, focus-group video recordings and questionnaire data were stored on a password-protected laptop and will be kept for at least three years after the study completion.

Chapter 4: Results

In this chapter, the research questions will be addressed through analyses of the quantitative data derived from the questionnaires and the qualitative data derived from the focus-group discussions. Research question 1, which examined teachers' perceptions of specific aspects of VR technology, was investigated using ANOVAs and t-tests. A thematic analysis was conducted on the textual data to further support the quantitative results for research question 1. Research question 2, which aimed to determine whether there are differences in the perceptions of ImmerseMe current users and non-users, was explored using MANCOVA and individual ANCOVA tests.

4.1. Research Question 1

What are teachers' perceptions of using ImmerseMe in language teaching, in terms of areas of learning supported by VR, VR benefits and drawbacks, implementation challenges, and implementation enabling factors?

4.1.1. Testing Assumptions

First, normality tests were carried out for items 10 to 61 on the questionnaire to evaluate the assumptions required for ANOVA analyses. (Items 1 to 9 collected participants' demographic information and thus were not included in the table.) The results below demonstrated that the distribution of most dependent variables was normally distributed, with skewness values between -.5 and .5 and kurtosis values between -1 and 1 (see Table 1), except for items 51, 57, and 58, which had skewness and kurtosis values of >1. This suggests that the mean values of these items were not normally distributed, potentially reducing the power of the test. The histograms demonstrated that there were no significant outliners on any of the 51 items.

Table 1: *Normality Tests and Descriptive Statistics for Questionnaire Items 10-47* (see Appendix H for the full version)

Outcome Measure	Subscale	Item	*M*	*SD*	Skewness	Kurtosis
Areas of Learning Supported by VR Technology	**Linguistic forms**	42 (complex and grammatically accurate language)	3.52	.819	-.168	.74
		43 (grammatical and lexical skills)	3.53	.733	-.026	-.242
		44 (vocabulary use)	3.64	.807	-.251	.243
		45 (better written performance)	3.32	.834	-.164	.707

	Productive skills	46 (better speaking abilities)	4.04	.675	-.234	-.140
	Receptive skills	47 (improve listening skill)	4.03	.683	-.364	-.713
Benefits of VR Technology	Learning approaches	10 (task-based learning)	4.06	.715	-.56	.54
		11 (experiential and contextualised learning)	4.27	.578	-.098	-.484
		12 (meet students' different needs)	4.00	.756	-.539	.248
		13 (autonomous annd active learning)	4.18	.644	-.401	.329
	Teaching approaches	14 (integration of language activities)	4.08	.700	-.445	.236
		15 (constructivist teaching)	4.01	.683	-.377	.314
		16 (learner-centred mode of instruction)	4.12	.700	-.343	-.299
		17 (rich and natural interaction environments)	4.18	.659	-.411	.164
	Language Acquisition	18 (speaking skills)	4.02	.743	-.456	.067
		19 (general language skills and knowledge)	4.05	.681	-.427	.418
		20 (language confidence)	4.15	.659	-.170	-.694
		21 (language performance and grades)	3.79	.836	-.088	-.234
		22 (self-correcting skills)	3.84	.732	-.334	.058
		23 (long-term memory retention of learnt knowledge)	3.84	.678	.021	-.421
Drawbacks of VR Technology	Educational value	25 (educational value)	2.77	1.017	.471	-.310
		26 (distracting)	2.46	.938	.499	-.464
	Learning Autonomy	27 (over-reliance on technology)	2.75	1.113	.395	-.651
		28 (students follow script)	3.20	.909	-.019	-.795
	Lesson Planning	29 (time-consuming)	3.32	1.065	-.387	.477
		30 (insufficient practice)	2.99	.951	.154	-.729
	Classroom environment	31 (uncontrollable class)	2.76	.991	.492	-.318
Implementation Challenges	Technical-related issues	33 (unable to cope with technical issues)	2.57	.966	.326	.496
		34 (ineffective teaching and learning)	3.31	1.008	-.207	-.656
		35 (insufficient facilities)	3.49	1.044	-.179	-.969
		36 (no access to VR tools)	3.50	1.062	-.389	-.644
	Lesson Design	37 (unable to integrate VR into curriculum)	2.58	1.013	.465	-.348
		38 (replanning the lesson design)	3.08	1.066	-.009	-.777
	Teacher Training	39 (lack of training materials)	3.08	1.021	-.099	-.844
		40 (lack of teacher training)	3.42	1.004	-.344	-.713
	Technology	49 (teachers' and students' digital literacy)	4.17	.765	-.697	-.208

		50 (teacher and student comfort with the use of VR)	3.98	.759	-.371	-.189
		51 (functioning facilities)	4.52	.695	-1.490	2.127
Implementation Enabling Factors	**Curriculum**	52 (adaptive curricula)	3.96	.784	-.177	-.761
		53 (flexible lesson structure)	4.01	.803	-.359	-.543
		54 (long practice time)	3.64	.856	.121	-.748
		55 (consistent practice)	4.00	.797	-.233	-.812
		56 (class management)	3.90	.894	-.799	-.713
	Student contribution	57 (students' level of engagement)	4.26	.721	-1.216	3.355
		58 (students' learning autonomy)	4.18	.782	-.946	1.596
		59 (students' concentration)	4.20	.642	-.428	.386
		60 (students' strong theoretical understanding of grammar)	3.14	.975	.025	-.073
		61 (students' pronunciation)	3.71	.895	-.546	.025

Another assumption for which checking is required before running repeated measures ANOVAs is the sphericity test. If $p > .05$ then sphericity is assumed. If $p < .001$, sphericity is not assumed, and the Greenhouse-Geisser (when $\varepsilon < .75$) or Huynh-Feldt (when $\varepsilon > .75$) correction is used to make an adjustment to the degrees of freedom (Howell, 2002). The results of the sphericity test are reported for each outcome measure.

4.1.2. Areas of Learning Supported by VR Technology

4.1.2.1. Quantitative Analysis

Sphericity was not assumed for this outcome measure ($p < .001$). The repeated measures ANOVA results demonstrated that there was a statistically significant difference between the means [$F(4,461) = 31$, $p < .001$]. As there were differences in the means of items, a pairwise comparison was conducted to determine which items' means were significantly higher than the others. The comparison consisted of multiple paired t-tests with a Bonferroni correction. The total number of pairs was 15 for this outcome measure, making the adjusted alpha level $p < .003$. Results can be found in Table 2 below. Due to space constraints, only pairs with statistically significant differences ($p < .003$) were listed. See Appendix I for all pairs.

Table 2: *Paired T-Test Results for the Outcome Measure Areas of Learning*

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q42 – Q46	-.519	.875	.085	-.067	-.035	-6.103	105	.000*
Q42 – Q47	-.509	.784	.076	-.660	-.359	-6.692	105	.000*

Q43 – Q45	.208	.700	.068	.073	.342	3.053	105	.000*
Q43 – Q46	-.509	.796	.077	.073	-.356	-6.591	105	.000*
Q43 – Q47	-.500	.694	.067	-.634	-.366	-7.423	105	.000*
Q44 – Q45	.321	.763	.074	.174	.468	4.329	105	.000*
Q44 – Q46	- .396	.801	.078	-.055	-.242	-5.093	105	.000*
Q44 – Q47	- .387	.738	.072	-.529	-.245	-5.398	105	.000*
Q45 – Q46	-.717	-848	.082	-.880	-.554	-8.705	105	.000*
Q45 – Q47	-.708	.084	.078	-.862	-.553	-9.056	105	.000*

Table 2 shows that statistically significant differences ($p < .003$) were found. Specifically, items 46 (better speaking abilities) and 47 (improve listening skill) scored significantly higher than all other items (42 to 45). In other words, teachers perceived speaking and listening as the two areas of learning supported the most by VR technology. Items 46 and 47, however, were not significantly different from each other. This means that the benefits of VR technology for speaking and listening were perceived as similar by teachers. Table 2 also shows that participants rated item 45 (better written performance) significantly lower than items 43 (develop self-correcting grammatical and lexical skills) and 44 (use vocabulary more flexibly; see Table 1 for item means). In other words, the teachers believed that VR-based programs benefit learners' grammatical and lexical skills and vocabulary usage more than their writing performance. To conclude, these findings indicate that teachers perceive the levels of VR benefits as follows: speaking = listening > vocabulary usage = self-correcting grammar and lexicon > writing performance.

4.1.2.2. Qualitative Analysis

Seven themes emerged from the qualitative data for the outcome measure *areas of learning supported by VR technology,* all of which had already been identified in the literature (see Table 3). Teachers mentioned the benefits of VR for speaking and conversational skills the most, from 10 to 12 times in the discussions. This finding aligns with the quantitative results, which showed that speaking was rated as the area of learning that benefited most from VR. Specifically, non-users and current users of ImmerseMe agreed that the speaking practice VR offers is particularly beneficial for beginner levels. Teacher H, a current user of ImmerseMe, said that 'the ability to slow the speaking pace [of the on-screen characters] for each sentence is fantastic for those who want to hear every word and repeat [the sentence]' (Discussion 2). Non-user C added that 'students may lose their initial inhibitions and speak up while interacting with VR programs compared to real humans from scratch' (Discussion 4).

Apart from speaking skills in general, teachers also concurred that VR technology specifically benefits students' pronunciation. Teacher A commented that 'having immediate feedback and having a VR speaker who knows what's going on and can accept or reject answers based on whether or not they make sense, that part would be fantastic' (Discussion 2). Similarly, current user K stated that 'if they don't respond correctly, it doesn't move on, and […] they have to be careful about their pronunciation and their understanding. […] there are lots of different activities that they can do, I think VR really brings textbooks to life' (Discussion 2). Teacher D further mentioned that 'it is fantastic to […] get that practice of pronunciation with someone who is a native speaker and can understand them [students]' (Discussion 5). Teachers were also impressed with the fact that VR exposes students to different learning sources rather than only textbooks and teacher input, facilitating students' learning experiences and improving their communication skills.

However, the data shown in Table 3 suggest that teachers pay less attention to the benefits of VR for receptive skills, including listening and reading, which were only mentioned once or twice during the discussions. Neither current users nor non-users elaborated on how VR might benefit receptive skills, suggesting that the idea of using VR for listening and reading is not well-recognised and might still appear unclear to teachers.

Table 3: *Thematic Analysis of Outcome Measure Areas of Learning Supported by VR*

Overarching themes	Keywords/Sample quotes	First Rater Quote Count	Second Rater Quote Count
Linguistic Forms (Lexicon and Grammar)	**E.g.:** "Vocabulary can be every naturally learned. For instance, pointing to an object and you can hear the word, say a word and the object is highlighted, play a video game and voice recognition"	1	1
Speaking and Conversational Skills	"speak up", "allowing the repetition practice of pronunciation, "immediate practice", "speaking fluency", "the ability to slow the speaking pace", "improve engagement and communication"	12	10
Listening	"great for listening", "practice repeatedly", "improving listening skills" **E.g.:** "It's a great platform for speaking and listening. We (teachers) often struggle to get real-life audio for the kids to hear and they (students) can practice that repeatedly with VR"	2	2
Pronunciation	"improve pronunciation", "students repeat themselves", "practice pronunciation with native speakers", "more contact with native sounds" **E.g.:** "Having the consonant feedback between the speech recognition of ImmerseMe, I think it is being extraordinary to improve pronunciation"	6	6
General Performances and Learning Outcomes	"provide support for standardised testing", "enhance general learning" **E.g.:** "It (ImmerseMe) gives them the ability to practice skills that they can use beyond just a school test"	3	3
Reading	**E.g.:** "Students' reading skills would also improve"	1	1

4.1.3. Benefits of VR Technology

4.1.3.1. Quantitative Analysis

Distinct from the analysis for outcome measure *areas of learning*, which only examined the statistically significant differences between item means to determine which items were perceived as more important than the others, two separate analyses were conducted for the outcome measure *benefits of VR technology*. Firstly, a similar analysis to the above was conducted to determine which items were rated higher than the others under each subscale (learning approaches, teaching approaches, and language acquisition).

Sphericity for this outcome measure was not assumed ($p < .001$). The repeated measures ANOVA results demonstrated that there was a statistically significant difference between the means of all items [$F(10,1066) = 9.55$, $p < .001$]. A pairwise comparison was then conducted to determine which items' means were significantly higher than the others within each subscale. The numbers of pairs for each subscale were 6, 6, and 15 respectively, making the adjusted alpha levels .008, .008, and .003. Results can be found in Table 4 below. Due to space constraints, only pairs with statistically significant differences (p < adjusted α) were listed. See Appendix J for all pairs.

Table 4: *Paired T-Test Results for Each Subscale of Benefits of VR*

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q10 – Q11	-.217	.647	.063	-.342	-.092	-3.451	105	.001*
Q11 – Q12	.274	.655	.064	.147	.400	4.299	105	.000*
Q18 – Q21	.226	.637	.062	.104	.349	3.661	105	.000*
Q19 – Q21	.255	.691	.067	.122	.388	3.796	105	.000*
Q19 – Q22	.208	.672	.065	.078	.337	3.179	105	.002*
Q19 – Q23	.208	.658	.064	.081	.334	3.248	105	.002*
Q20 – Q21	.358	.733	.071	.217	.500	5.037	105	.000*
Q20 – Q22	.311	.667	.065	.183	.440	4.805	105	.000*
Q20 – Q23	.311	.638	.062	.188	.434	5.025	105	.000*

Regarding the subscale *learning approaches* (items 10-13), statistically significant differences ($p < .008$) were only found in two pairs of items: items 10 and 11 and items 11 and 12 (see Table 4). Specifically, item 11 (experiential and contextualised learning) had a significantly higher mean than both item 10 (task-based learning) and item 12 (meet students' different needs) (contextualised > task-based = meet needs). No significant differences were found between item 13 and the others, indicating that teachers perceived experiential and contextualised learning as the most important VR benefit in terms of learning approaches. Regarding the subscale *teaching approaches*, no statistically significant differences were found, indicating that teachers perceived VR benefits to teaching approaches as equal. Regarding the subscale *linguistic acquisition*, statistically significant differences ($p < .003$) were found. Specifically, items 19 (language skills and knowledge) and 20 (language confidence) had higher means than items 21 (language performance and grades), 22 (general self-correcting skills), and 23 (long-term memory retention). This indicates that teachers perceived VR benefits to language knowledge and confidence as greater compared to its benefits on school grades, self-correcting skills, and long-term memory retention (language knowledge = language confidence > class grades = self-correcting skills = memory retention). The data in Table 1 further suggests that teachers perceived language confidence ($M = 4.15$) as the most common benefit of VR in terms of linguistic acquisition.

The second analysis aimed to determine which subscale of *benefits of VR* was perceived as most important by the participants. In this analysis, the means of all items were averaged, with descriptive statistics provided in Table 5.

Table 5: *Descriptive Statistics for the Subscale Benefits of VR*

Item Pair	**Mean**	**Std. Deviation**	**N**
Learning approaches (OM1)	4.1274	.54253	105
Teaching approaches (OM2)	4.0967	.56122	105
Linguistic acquisition (OM3)	3.9481	.58026	105

As shown in Table 5, all three subscales have high means (~4.00 on a 1-to-5 scale), and thus all are perceived as being supported by VR. This finding confirms that VR technology benefits learning approaches, teaching approaches, and language acquisition. To further test the differences between these means, repeated measures ANOVAs were conducted. The results demonstrated that there was a statistically significant difference between the means of the three items [$F(2,182) = 10$, $p < .001$. A pairwise comparison was then conducted to determine which subscale scored significantly higher than the others. The number of pairs was 3, so the adjusted alpha level was .017. Results can be found in Table 6 below.

Table 6: *Paired T-Test Results for the Benefits of VR Subscale Comparison*

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
OM1 – OM2	.03066	.40781	.03961	-.04788	.10920	.774	105	.441
OM1 – OM3	.17925	.35779	.03475	.11034	.24815	5.158	105	.000
OM2 – OM3	.14858	.50579	.04913	.05118	.24599	3.025	105	.003

As shown in Table 6, there were statistically significant differences between the means of two pairs: OM1 (learning approaches) and OM3 (linguistic acquisition), and OM2 (teaching approaches) and OM3 (linguistic acquisition) ($p < .017$). OM1 and OM2 had higher means than OM3, suggesting that teachers perceived the benefits of VR benefits to learning and teaching approaches to be approximately equal, and significantly higher than the benefits to linguistic acquisition.

4.1.3.2. Qualitative Analysis

Seven themes emerged from the qualitative data on the outcome measure *benefits of VR technology*. Four of these themes had already been described in the literature, while the other three, *content enrichment*, *convenience and affordability*, and *students' experiences*, were developed inductively based on the data (see Table 7). The most positively perceived VR benefit was on students' experiences, with more than 35 mentions, followed by the benefits to language confidence and intercultural awareness, with 18 mentions. Benefits to learning and teaching approaches were also mentioned frequently in the discussion (8-13 times); however teachers barely mentioned VR benefits to students' linguistic knowledge or skills.

Many teachers agreed that VR brings novelty and authenticity to the classroom, which might boost students' interest, engagement, and confidence. For instance, teacher M said that 'VR shows students that they can participate in a full conversation with a native speaker without having their anxiety go through the roof […]. It really comforts them in there [VR environments]' (Discussion 4). Other teachers pointed out that the authenticity offered by VR is specifically beneficial to students from rural areas, many of whom lack genuine interactions with native speakers of the target language. Teacher B, a current user of ImmerseMe based in the United States, said that:

'They [students] like the immersive part because a lot of my students have never been abroad, or have never left the state of Virginia, or maybe even the Hampton Roads area. So it was their first time really seeing a grocery store in France, greeting someone in an informal or formal way. I think overall it was a really positive experience' (Discussion 5).

Teacher P, another current user of ImmerseMe based in New South Wales, Australia, added that 'to be able to listen and interact with native speakers in real-life scenarios, particularly in the community that we come from, where there are students who have maybe never gone more than a couple of hours from this small country town, is extraordinary' (Discussion 5). Teachers also strongly believed that exposure to authentic scenarios allows learners to increase their cultural awareness. Teacher V, a current Russian teacher based in the US, commented that:

‘We were able to introduce some different accents of Russian speaking as well, from different cities. It allowed us to diversify because our textbooks have very Russia-centered content […] for example accents from Moscow, St. Petersburg. […] It is the era of diversity, inclusion, and equity, so we needed to show students that there are different parts of Russia, like Islamic parts of Russia or non-Slavic parts of Russia. ImmerseMe really helped with that’ (Discussion 3).

In conclusion, the qualitative findings suggested that teachers are most aware of VR’s benefits to student’s experiences (via the elements of novelty and authenticity). The benefits to learning and teaching approaches were also discussed among teachers (see Table 7), but discussion around VR benefits to linguistic acquisition centred around how VR improves language confidence with no focus on language abilities or skills. This finding aligns with the quantitative results, which demonstrated that language confidence is perceived as the most common benefit of VR in terms of linguistic acquisition.

Table 7: *Thematic Analysis of Outcome Measure Benefits of VR Technology*

Overarching themes	**Keywords/Sample Quotes**	**First Rater Quote Count**	**Second Rater Quote Count**
Learning approaches	“individualised”, “situational”, “differentiation”, “safe bubble”, “explore at their own pace”, “collaborative working”, “benefit solo study”, “personalised”, “body movements”, “targeting students’ needs”, “active and engaging”, “cater for all learners” **E.g.:** “It is very individualised. Each student can work at their own pace in their own time”	13	13
Teaching approaches	“peer-to-peer teaching”, “classroom engagement”, “save time for teachers to generate lasting knowledge”, “scaffolding support”, “reinforce language teaching”, “authentic teaching”, “elicit concepts for new words”, “authentic environment” **E.g.:** “It helps with monolingual language teaching (teaching without students’ L1) because teachers can accurately elicit concepts for new words and phrases by showing them to the students in a virtual environment”	8	8
Students’ Learning Attitudes and Motivation	“stimulate students’ interest”, “students prone too feel comfortable and ready to learn”, “good for students who get nervous during the pandemic”, “there is no hiding in the virtual world”, “without anxiety when have conversations with native speakers”	9	9

Language Confidence and Intercultural Awareness	"belief in themselves (students)", "students learning culture along with language", "provide culturally real environment", "authenticity and connection to the real world", "non-threating learning experience", "no worries about others react to their input", "access to native speakers", "providing cultural awareness and dialects" **E.g.:** "What I see with my students is the change in terms of perception and agency and in terms of their own beliefs in themselves"	18	16
Content Enrichment	"enrich lesson content", "practice and engage with textbook content", "authentic materials", "allowing to renovate materials" **E.g.:** "VR as a program was great because teachers could use it to enrich the lesson content"	6	6
Convenience and Affordability	"super accessible", "not expensive", "can be done in a safe environment" **E.g.:** "It works brilliantly on a laptop or an iPad or even a phone because you can still look around without the VR kits"	4	4
Students' Experiences	"exciting new way of learning", "novelty", "away from the classroom pressure and feel the nicer vibe", "totally immersed into the environment in separation to the classroom", "positive students' experiences"	35	33

4.1.4. Drawbacks of VR Technology

4.1.4.1. Quantitative Analysis

Sphericity for this outcome measure was not assumed ($p < .001$). The repeated measures ANOVA results demonstrated that there was a statistically significant difference between the means [$F(5{,}569) = 14$, $p < .001$]. A pairwise comparison was then conducted to determine which items' means were significantly higher than the others. The total number of pairs was 21 for this outcome measure, thus the adjusted alpha level was $p < .002$. Results can be found in Table 8 below. Due to space constraints, only pairs with statistically significant differences ($p < .002$) were listed. See Appendix K for all pairs.

Table 8: *Paired T-Test Results for Outcome Measure Drawbacks of VR Technology*

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q25 – Q28	-.425	1.242	.121	-.664	-.185	-3.520	105	.001*
Q25 – Q29	-.547	1.374	.133	-.812	-.283	-4.100	105	.000*
Q26 – Q28	-.736	.989	.096	-.926	-.545	-7.664	105	.000*
Q26 – Q29	-.858	1.099	.107	-1.070	-.647	-8.041	105	.000*
Q26 – Q30	-.528	1.053	.102	-.731	-.326	-5.166	105	.000*
Q26 – Q31	-.302	.968	.094	-.488	-.116	-3.212	105	.002*
Q27 – Q28	-.453	1.243	.121	-.692	-.213	-3.750	105	.000*
Q27 – Q29	-.575	1.265	.123	-.819	-.332	-4.685	105	.000*
Q28 – Q31	.434	1.163	.113	.210	.658	3.841	105	.000*
Q29 – Q30	.330	.983	.095	.141	.519	3.459	105	.001*
Q29 – Q31	.557	1.079	.105	.349	.764	5.313	105	.000*

Table 8 shows a number of statistically different pairs ($p < .002$). Specifically, items 28 (follow the script) and 29 (time-consuming) had significantly higher means than items 25 (educational value), 26 (distracting), and 27 (technology over-reliance). Table 8 also demonstrates that item 29 (time-consuming) had a significantly higher mean than items 30 (insufficient practice time) and 31 (harder to control class). However, items 28 and 29 were not significantly different from each other, suggesting that the two biggest drawbacks of VR technology in language learning are (1) that it is time-consuming for lesson planning and (2) that it makes students follow a script, and they may therefore struggle to generate conversations themselves.

4.1.4.2. Qualitative Analysis

Six themes emerged from the qualitative data for the outcome measure *drawbacks of VR technology*. Four of these already existed in the literature, while the last two, *students' physical and mental health* and *lack of cultural diversity*, were developed inductively (see Table 9). In general, teachers concurred that there are several drawbacks when using VR technology in classroom settings. The biggest drawback, mentioned ten times during the discussions, was that VR could not be used to completely replace classroom interaction. For instance, teacher D, who had had limited experience with ImmerseMe, argued that:

'It's walled off in terms of how students might use it. In terms of getting [students] to communicate and using target language in a lesson, it [content on VR] is a chunk on its own. And so I think using VR as a reward is a good thing but working it into lesson planning and using it to replace face-to-face interaction, it could be a problem. It's a solution to a problem that doesn't really exist. We can get [students] to talk to each other if they're face to face' (Discussion 2).

Similarly, teacher P, a non-user of ImmerseMe, stated that ‘if [students] can only have 10-15 minutes talking to a real person on the screen, it’s not really interactive. It’s not a proper conversation. I’d rather have them talking to each other.’ (Discussion 1).

Other teachers believed that using VR technology has negative impacts on learning environments, an idea that was mentioned seven times during the discussions. Teacher N said, ‘I've tried some classes when everyone was using it simultaneously, which really does not work, and it makes students frustrated’ (Discussion 1). On the same note, teacher B, based in New Zealand, also commented that ‘it is too noisy to use VR for 27 students, it’s only useful as a homework tool’. This suggests that class size can be a factor that affects the effective use of VR tools in classroom settings.

Finally, the emergent theme *lack of cultural diversity* shows that teachers may believe that ImmerseMe does not offer cultural diversity in its scenarios. For instance, teacher B, who was a new user of ImmerseMe, stated that ‘the scenarios tend to be very generic, rather than language- or culture-specific. It does not allow for cultural differences within a language’ (Discussion 3). This seems to contradict the previous statement by teacher V (in the *VR benefits* section), suggesting that the content developed for certain languages might be more culturally aware than for others.

To conclude, the qualitative results in this area seem to contradict the quantitative findings, which had shown that the biggest drawbacks of VR technology were (1) its time-consuming nature and (2) that students over-rely on scripts and struggle to generate conversations themselves. Throughout the discussions, issues with lesson design were only discussed briefly (with three mentions). Instead, teachers focused on how the use of VR could not replace peer-to-peer interaction, its lack of cultural diversity and how it had negative impacts on classroom environments.

Table 9: *Thematic Analysis of Outcome Measure Drawbacks of VR Technology*

Overarching themes	Keywords/Sample Quotes	First Rater Quote Count	Second Rater Quote Count
Educational value	“total unsuccess”, “not lending in”, “rigid use”, “walled off”, “not a proper conversation”, “not everyone engages”, “cannot replace human interactions”, “too generic” **E.g.:** “It (ImmerseMe) doesn’t lend in classroom settings very well so we can’t use it much”; “It (ImmerseMe) is a solution to the problem that doesn’t exist”	12	10
Learning autonomy	“over-relying”, “reluctant”, give up too easily” **E.g.:** “They (students) might get over-rely on technology to practice, they might only feel like practicing when that availability happens”	2	2
Lesson Planning	“require new ideas”, “complicated lesson planning”, “require extra time”, “highly skilled task” E.g.: “Creating lessons with VR will require a lot of new ideas from teachers”	3	3

Learning environment	"does not work simultaneously", "not interactive enough", "noise management", "students rocking, spinning and twisting around the room", "too noisy"	7	7
Students' Mental and Physical Health	"feeling ill", "getting dizzy", "become anxious", "students freak out" **E.g.:** "They (students) might feel a little ill from the glasses"	5	5
Lack of Cultural Diversity	"too generic", "not cultural specific", "not enough real culture"	2	2

4.1.5. Implementation Challenges

4.1.5.1. Quantitative Analysis

Sphericity for this outcome measure was not assumed ($p < .001$). A repeated measures ANOVA test was then conducted for the outcome measure *implementation challenges.* The results demonstrated that there was a statistically significant difference between the means [$F(6,610) = 18$, $p < .001$]. A pairwise comparison was then conducted to determine which items' means were significantly higher than the others. The total number of pairs was 28 for this outcome measure, thus the adjusted alpha level was $p < .002$. Results can be found in Table 10 below. Due to space constraints, only pairs with statistically significant differences ($p < .002$) were listed. See Appendix L for all pairs.

Table 10: *Paired T-Test Results for Outcome Measure Challenges to VR Implementation*

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q33 – Q34	-.745	1.096	.106	-.956	-.534	-7.001	105	.000*
Q33 – Q35	-.925	1.165	.142	-1.027	-.642	-6.497	105	.000*
Q33 – Q36	-.934	1.354	.132	-.1.195	-.673	-7.101	105	.000*
Q33 – Q38	-.509	1.340	.130	-.767	-.251	-3.915	105	.000*
Q33 – Q39	-.509	1.396	.136	-.778	-.241	-3.758	105	.000*
Q33 – Q40	-.858	1.390	.135	-1.126	-.591	-6.359	105	.000*
Q34 – Q37	.726	1.444	.140	.448	1.005	5.178	105	.000*
Q35 – Q37	.906	1.299	.126	.656	1.156	7.181	105	.000*
Q35 – Q38	.415	1.241	.121	.176	.654	3.443	105	.001*
Q35 – Q39	.415	.994	.097	.224	.607	4.300	105	.000*
Q36 – Q37	.915	1.448	.141	.636	1.194	6.505	105	.000*
Q37 – Q38	-.491	1.189	.116	-.720	-.262	-4.247	105	.000*
Q37 – Q39	-.491	1.221	.119	-.726	-.255	-4.137	105	.000*
Q37 – Q40	-.840	1.266	.123	-1.083	-.596	-6.828	105	.000*
Q39 – Q40	-.349	-.873	.085	-.517	-.181	-4.115	105	.000*

Table 10 shows a number of significantly different pairs ($p < .002$). Specifically, item 33 (inability to cope with technical issues) had a significantly lower mean than all of the other items except for item 37 (inability to integrate VR into the curriculum), for which no difference was found. In other words, the inability to deal with technical issues and to integrate VR into the curriculum were the two challenges that least concerned teachers when using VR technology. As also shown in Table 10, items 35 (lack of functioning facilities) and 36 (no access to VR programs) had significantly higher means than all other items. This suggests that lack of access to hardware facilities and to the VR program are currently the two biggest challenges to VR implementation. Finally, Table 1 also shows that items 39 (lack of training materials) and 40 (lack of teacher training) generally had high means ($M = 3.08$ and $M = 3.42$), only lower than item 35 ($M = 3.49$) and 36 ($M = 3.50$) mentioned above (training materials < teacher training < functioning facilities < no access to VR). To conclude, these findings confirm that the main challenges to the effective use of VR technology in the classroom are a lack of access to the VR program and to the required VR facilities, as well as a lack of teacher training.

4.1.5.2. Qualitative Analysis

Six themes emerged for the outcome measure *implementation challenges*. These were mostly consistent with the literature, except for the theme *functioning software*, which referred to the functionality of the voice recognition feature (see Table 11). This theme was mentioned the most during the discussion (25 mentions), and thus was obviously seen as the biggest barrier to effective use of VR. Many teachers agreed that the voice recognition feature of ImmerseMe did not work well, which was the main reason for students' frustration and loss of interest. This technical barrier was mainly recognised by current users of ImmerseMe, who have had unpleasant experiences with this feature. Teacher A, a long-term user of ImmerseMe, said, 'The most frustrating thing is the technical glitches. We didn't know what mistake it was to move on, it might have been a full stop, a comma or accents.' (Discussion 5). Similarly, teacher D, a recent user of ImmerseMe, added that:

'Since the system did not work well, I didn't know what mistakes students were making, whether they were punctuation or grammar or pronunciation mistakes. It wasn't terribly clear, and so my major worry would be students making mistakes and they don't understand what mistakes they were making and then that going unnoticed entirely' (Discussion 2).

Other main challenges to VR implementation in classroom settings include teacher training (15 mentions) and functioning facilities (12 mentions). In terms of teacher training, teacher N, a recent user of ImmerseMe, commented that 'most of my stuff has been through the experience of using and trial and error in the classroom' (Discussion 1). Teacher D added that her time using VR in Spanish classes was not ideal: 'I walk into the room and throw the kids on the tool and not panic too much because I have got enough teaching experience and I know that the kids will find their way. It can be tough if you haven't done a lot of teaching before' (Discussion 1). These comments suggest that a lack of high-quality teacher training is another major barrier that needs to be addressed. In terms of functioning facilities, most of these concerns came from teachers in rural areas, who claimed that the facilities did not allow them to install and use VR tools effectively. In particular, teachers M, B, and K, who are based in rural regions in Australia and the US, complained that the computers were 'too slow' and 'not functioning well', and that the available technology was quite 'primitive' and thus the teaching with VR did not go well.

In conclusion, on one hand, the qualitative findings support the quantitative results, which showed that lack of functioning facilities, access to the VR program, and teacher training were the main challenges in the effective use of VR. On the other hand, the qualitative findings introduced another crucial concern of ImmerseMe current users: the functionality of the voice recognition features.

Table 11: *Thematic Analysis of Outcome Measure Challenges to VR Implementation*

Overarching themes	Keywords/Sample Quotes	First Rater Quote Count	Second Rater Quote Count
Functioning Software (Voice Recognition)	"technical glitches", "wasting time waiting for voice recognition", "get frustrated", "strong feeling of giving up because of glitches", "the computer doesn't understand students' voices", "slow sound processing" **E.g.:** "We didn't know what the mistake was to move onto the next one, it might have been a full stop or a comma out of place or might have been the accent. We're not sure"	25	25
Functioning Facilities	"slow functioning on iOS", "lack of functioning computers", "really slow computers", "students' computer settings froze", "primitive technology in state schools" **E.g.:** "Our school doesn't have enough computers for the kids. They have to share with other schools so using VR can be challenging, unless the other school wants it too"	12	12
Internet Connection	"troubling internet connection", "slow internet", "really poor internet connection", "slow WIFI network speed", "spinning blue circle for the whole lesson", "took too long" **E.g.:** "Lack of fast internet was the main issues for my students, They didn't enjoy the program because it took too long to open"	8	8
Teacher Training	"basically no training", "test it out and find your way", "the kids will find their way", "lack of hands-on training", "a lot of teaching other teachers", "feel lost"	15	15

Lesson Design	"time-consuming re-designing process", "things might go off plan because of technical issues", "longer lesson", "less time for teaching", "a lot of pre-session planning", "not suitable for test-oriented culture"	10	9
Affordability	"not affordable", "expensive tools", "out budget cannot afford glasses", "costly", "cost barriers" **E.g.:** "We don't have the budget to purchase glasses for all students. The computers are slow, but we can't afford the ones that allow high quality VR experiences"	9	9

4.1.6. Implementation Enabling Factors

4.1.6.1 Quantitative Analysis

Sphericity for this outcome measure was not assumed ($p < .001$). The repeated measures ANOVA results demonstrated that there was a statistically significant difference between the means [$F(8,872) = 29$, $p < .001$]. As there were differences in item means, a pairwise comparison was conducted to determine which items' means were significantly higher than the others. The total number of pairs was 78 for this outcome measure, making the adjusted alpha level $p < .001$. As the number of possible pairs was high (>50 pairs), a pre-assessment was conducted to select the most important paired t-tests to be run. First, the items were reorganised according to their mean values (see Table 12). According to this order, the item with the lowest mean (item 60) was compared to the item with the second-lowest mean (item 54). If a statistically significant difference was not found, the item with the lowest mean (item 60) was compared to the item with the third-lowest mean (item 61), and so on, until a pair showing a statistically significant difference was encountered. It was considered likely that subsequent comparisons to items with higher means would also be statistically significant (unless the standard deviations were very different).

Table 12: *Means of Implementation Enabling Factors Items (After Reorganisation)*

Implementation Enabling Factors	Mean
60. Student's strong theoretical understanding of grammar	3.14
54. Long practice time	3.64
61. Student's pronunciation of learnt vocabulary	3.71
56. Class management	3.90
52. Adaptive curricula	3.96
50. Teacher and student's comfort with the use of VR	3.98
55. Consistent practice	4.00
53. Flexible lesson structure	4.01
49. Teacher and student's digital literacy	4.17

58. Student’s learning autonomy	4.18
59. Student’s concentration	4.20
57. Student's level of management	4.26
51. Well-functioning facilities	4.52

The results of the paired t-tests can be found in Table 13 below. Due to space constraints, only pairs with statistically significant differences ($p < .001$) were listed. See Appendix M for all pairs.

Table 13: *Paired T-Test Results for Outcome Measure Implementation Enabling Factors*

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q60 – Q54	-.495	1.066	.104	-.702	-.289	-4.759	105	.000*
Q54 – Q52	-.324	.791	.077	-.477	-.171	-4.197	105	.000*
Q61 – Q49	-.457	1.047	.102	-.660	-.254	-4.473	105	.000*
Q56 – Q51	-.619	1.023	.100	-.817	-.421	-6.203	105	.000*
Q52 – Q51	-.562	.876	.086	-.731	-.392	-6.571	105	.000*
Q50 – Q51	-.543	.797	.078	-.697	-.389	-6.980	105	.000*
Q55 – Q51	-.524	.962	.094	-.710	-.338	-5.581	105	.000*
Q53 – Q51	-.514	.856	.084	-.680	-.349	-6.155	105	.000*
Q49 – Q51	-.352	.784	.077	-.504	-.201	-4.604	105	.001*
Q58 – Q51	-.343	.853	.083	-.508	-.178	-4.119	105	.000*
Q59 – Q51	-.324	.753	.073	-.470	-.178	-4.406	105	.000*
Q57 – Q51	-.267	.737	.072	-.409	-.124	-3.706	105	.000*

As shown in Table 13, there were statistically significant differences between item 60 (theoretical understanding of grammar) and item 54 (long practice time) and all other items with higher means ($p < .001$), illustrating that teachers perceive students’ theoretical understanding of grammar to be the least important factor in the effective use of VR in classrooms. In other words, students’ grammatical knowledge was not perceived to play an important role in enabling VR-driven teaching. Table 13 further shows that the first statistically significant difference found in the comparison between item 54 and the rest is Q54-Q52, indicating that item 54 (long practice time) had a significantly lower mean than item 52 (adaptive curricula) upward. Statistically significant differences were also found in the comparison of item 61 (pronunciation) and item 49 (digital literacy) upward. For the remaining analyses, the only pair that showed a statistically significant difference was the comparison between item 51 (functioning facilities) and every other item. This means that items 52 to 57 were not significantly different from each other, but were significantly different from item 51. This suggests that teachers perceived functioning facilities (computers and hardware) to be the most important factor in the effective use of VR technology in foreign language education. Another insight from this table is that factors related to students' contributions (student engagement, student learning autonomy, and student concentration) had high means compared to other factors.

4.1.6.2. Qualitative Analysis

Of the seven themes that emerged from the qualitative data, most were consistent with the literature, except for *regular practice* and *improvement in voice recognition accuracy*. In contrast to the quantitative results, which found that having functioning facilities was the most important factor in the effective use of VR, other technical factors were not mentioned frequently during the discussions. Instead, teachers focused on curriculum-related factors (6 - 8 mentions).

Most teachers agreed that rebuilding lesson planning and curriculum was crucial in the integration of VR, and that teachers should be flexible during the adaptation process. Teacher C, who had had more than five years of experience with ImmerseMe, suggested that:

'What really matters is to be considered in embedding the VR dialogues into the current curriculum. So being purposeful in selecting dialogues to be practised, and choosing a communication focus that is usually inaccessible via traditional teaching. My endgame is that by the time I introduce ImmerseMe, the students already know how to pronounce the words correctly to avoid problems with the voice system' (Discussion 4).

In other words, teacher C was recommending the use of VR as a supplementary resource to practice previously taught knowledge rather than as a learning tool in its own right. Teacher M concurred with teacher C, suggesting that preparing the students with sufficient knowledge and drill exercises before the VR practice was essential to their learning success and their general experiences with VR. Teacher B added another factor that teachers should consider while preparing lesson plans. She said:

'Being mindful of the teaching sequence to make the lessons more interesting and more engaging. Sometimes I'll start teaching in reverse with spelling mode then go back to pronunciation, then go back to writing, then conversations or translation. That way, students are not overwhelmed and the lesson plan is not repetitive.'

In short, apart from functioning facilities and a good internet connection, the qualitative analysis suggests that teachers should also consider curriculum-related factors, such as sufficient knowledge of learning content, adaptive curriculum, and a flexible teaching sequence to ensure that their VR-integrated classroom is effective and well-managed.

Table 14: *Thematic Analysis of Outcome Measure Implementation Enabling Factors*

Overarching themes	Keywords/Sample Quotes	First Rater Quote Count	Second Rater Quote Count
Teacher and Student's Computer Literacy	"effective demonstration", "digital capability", "basic IT skills for kids", "good digital literacy" **E.g.:** "Both students and teachers need good enough IT skills. If the teacher is comfortable to learn, the students will adapt. If the students have good digital literacy, which they usually do, but the teacher is not flexible, the platform is not used anyway".	3	3
Regular Practice	"regular usage in class", "finding the right balance in terms of classroom practice" **E.g.:** "I think it is important to get into the habit of regular usage in class"	2	2
Functioning Facilities	"hardware problems", "technical problems", "cost of software", "functionality of computers"	3	3
Internet Access and Connection	"no interruption", "good Internet", "long wait for uploading", "strong connectivity" **E.g.:** "It is definitely crucial to make sure that everyone had access to Internet and no interruption with their Internet"	3	3
Curriculum-related factors	"how to carry out curriculum", "big lesson planning issues", "what can be achieved in terms of curriculum and learning outcomes", "being purposeful when building curriculum", "right amount of time and activities" **E.g.:** "I think it is extremely important for teachers to consider how VR affects the curriculum. Why are we using it? What can be achieved additionally from VR that cannot be achieved in a normal classroom?"	8	6
Class Management	"appropriate breakout spaces", "keep quiet when others speak", "be structural", "big class size", "class organisation" **E.g.:** "An appropriate environment and small class size are needed to make this work. Make sure students don't run around and bump into each other"	5	5
Improvement in Voice Recognition Accuracy	"program's intelligibility level", "adaptive", "communicate effectively with the system" **E.g.:** "It would work for students to practice pronunciation if they can actually communicate without having to do several attempts because of the system"	2	2

4.2. Research Question 2

Do ImmerseMe current users and non-users differ in their perception of these aspects?

4.2.1. Demographic Differences

A number of tests were run to determine whether there were statistically significant demographic differences between ImmerseMe current users and non-users, and thus which covariates should be controlled for in the MANCOVA and ANCOVA analyses. T-tests were run on the continuous variables, including age, years of experience, perception of digital technology usefulness, and frequency of computer-based program usage. Chi-square tests were run on the categorical variables, including gender and teaching level. The results showed that the differences in age and years of experience between the two groups were statistically significant. Specifically, current users have higher age mean ($t = 5.31$, $df = 99$, $p < .001$) and higher average number of years of experience ($t = 4.31$, $df = 103$, p < .001; also see Table 15). These two variables are controlled for in the following MANCOVA and ANCOVA tests.

Table 15: *Demographic Statistics for ImmerseMe Current Users and Non-Users*

Variable	Group	Mean	Std. Deviation	N
Age	Current users	45.94	11.13	53
	Non-users	34.08	11.27	48
Years of Experience	Current users	18.35	10.22	55
	Non-users	10.06	9.40	50
Perception of Technology Usefulness	Current users	4.38	.680	55
	Non-users	4.33	.653	51
Computer-based Program Usage Frequency	Current users	4.31	.858	55
	Non-users	4.24	1.050	51

4.2.2. MANCOVA Results

Descriptive statistics for each group (ImmerseMe current users and non-users) can be found in Table 16 below.

Table 16: *Descriptive Statistics for ImmerseMe Current Users and Non-Users*

Outcome Measure	Group	Mean	Std. Deviation	N
Areas of Learning	Current users	3.55	.58	53
	Non-users	3.79	.57	48
	Total	3.67	.58	101
General benefits	Current users	4.02	.47	53
	Non-users	4.00	.53	48
	Total	4.01	.50	101

Drawbacks	Current users	2.69	.63	53
	Non-users	3.10	.66	48
	Total	2.89	.67	101
Implementation Challenges	Current users	3.08	.48	53
	Non-users	3.21	.61	48
	Total	3.14	.55	101
Implementation Enabling Factors	Current users	3.98	.44	53
	Non-users	3.97	.59	47
	Total	3.98	.51	100

The MANCOVA result demonstrates that the covariates of age and years of experience are related to the combined dependent variables (*areas of learning, VR benefits*, *VR drawbacks*, *implementation challenges,* and *implementation enabling factors*). The effect of age was found to be significant [$F(5,92) = 3.05$, $p < .05$, *Wilks' Λ* =.858], while the effect of years of experience was not found not to be significant [$F(5,92) = 1.732$, $p > .05$, *Wilks' Λ* = .914]. However, when controlling for both covariates a statistically significant difference was found, with $F(5,92) = 3.02$, $p < .05$, *Wilks' Λ* = .859.

4.2.3. ANCOVA Results

To avoid false statistical inferences, the alpha level was adjusted by dividing the standard alpha by the number of comparisons (Chen et al., 2017). The resulting adjusted α was .013. Results can be found in Table 17 below.

Table 17: *ANCOVA Results*

Outcome Measure	**Type 3 Sum of Squares**	***df***	**Mean Square**	***F***	***p***	**Partial Eta Squared**
Areas of Learning	2.065	1	2.065	6.166	.015	.060
General Benefits	.029	1	.029	.115	.736	.001
Drawbacks	1.925	1	1.925	4.924	.029	.048
Implementation Challenges	.001	1	.001	.004	.947	.000
Implementation Enabling Factors	.041	1	.041	.153	.696	.002

In terms of *areas of learning supported by VR* ($p = .015$) and *drawbacks of VR technology* ($p = .029$), there was a statistically significant difference between the perceptions of current users and non-users after controlling for age and years of experience at the standard significance level ($p = .05$). Specifically, regarding the outcome measure *areas of learning supported by VR*, the direction of the trend in difference was current users; meanwhile, the trend in difference of the latter outcome measure (*drawbacks of VR technology*) was non-users. However, when correcting for multiple comparisons in which the adjusted alpha level was $p = .013$, the difference was no longer significant. In terms of *benefits of VR technology* (*p* = .736), *implementation challenges* (*p* = .947), and *implementation enabling factors* (*p* = .696), there is no statistically significant difference in the perceptions of ImmerseMe current users and non-users after controlling for age and years of experience.

In conclusion, according to the MANCOVA results, there is a difference in the perceptions of ImmerseMe current users and non-users. However, ANCOVA findings with the adjusted alpha level suggested that there was no statistically significant difference in the perceptions of current users and non-users after controlling for participants' age and years of teaching experience. In other words, there is no evidence of differences between current users and non-users of ImmerseMe in regard to their perceptions of the aspects of VR investigated in the current study (*areas of learning*, *VR benefits*, *VR drawbacks*, *implementation challenges,* and *implementation enabling factors*).

Chapter 5: Discussion

5.1. Discussion of Findings

5.1.1. Research Question 1

Areas of Learning. According to the quantitative and qualitative results, the teachers in this study perceived that speaking skill is the area of learning best supported by the use of VR technology. Teachers elaborated that VR technology helps students to lose their initial inhibitions and shyness, allowing them to practice language production for social use. This is especially beneficial for beginner learners who tend to struggle in initiating social conversations. Teachers added that in VR-facilitated environments, students can listen to and repeat sounds produced by native speakers, therefore their pronunciation is also significantly improved. These findings are consistent with the existing literature (e.g., Alemi & Khatoony, 2020; Yamazaki, 2018), which acknowledges the crucial roles of VR technology in improving speaking and conversational skills. A plausible explanation as to why VR benefits speaking more than any other areas of learning may be that the initial purpose of VR in language classroom settings was to provide communication-driven environments to address one of the biggest problems in language learning: that is, lack of real-life communication. This, however, does not mean that the applications of VR are only limited to speaking and developing conversation skills.

In terms of listening, while the quantitative results demonstrated that teachers perceived both speaking and listening to be supported by the use of VR, in the follow-up discussions they rarely mentioned or elaborated on how VR benefits listening skill and comprehension. A plausible explanation for this might be that since communication requires both listening and speaking, it seems obvious that VR benefits both skills. However, teachers might have struggled to explain how exactly VR-driven environments benefit listening comprehension more than traditional methods (e.g., textbook listening exercises). This suggests a research gap in how engagement and interaction via VR tools might have effects on listening comprehension, and a lack of research comparing the effectiveness of VR in improving listening skills with other methods (e.g., listening exercises, audiobook).

In terms of writing, both the quantitative and qualitative findings suggest that the use of VR has clearer benefits to students' vocabulary and grammar use than to their writing. This supports the previous finding by Dolgunsoz et al. (2018), which demonstrated that VR did not affect short-term writing performance. It is, however, still inconclusive whether VR use can have long-term effects on student's writing. Longitudinal studies should be conducted to examine the long-term benefits of VR on writing performance. It is important to note that the language use and scenarios in VR settings are conversational and socially based, and therefore it is unrealistic to expect VR to benefit academic writing. Future researchers should be mindful of this when designing their studies. In terms of reading, this study found that the effects of VR on reading skills and comprehension were not clear to teachers compared to other resources (e.g., textbooks). This finding underscores the research gap in the potential benefits of VR to reading skills.

Benefits of VR. According to teachers, the use of VR in language classrooms mainly benefits contextualised learning as compared to other learning approaches. In other words, the teachers in this study believed that VR engages students in active learning and assists them to create meaning out of the knowledge they are building. The qualitative results further support this finding: teachers commented that the most impressive benefit of ImmerseMe is how it can improve students' experiences with the authentic culture of the target language, allowing them to learn by seeing and doing. These findings, indeed, empirically support a number of more theoretical works (e.g., Ampatzoglou & Chatzigeogiou, 2007; Kluge & Riley, 2008), that had previously argued that real-life language exposure in VR-facilitated environments allows students to construct and contextualise the new knowledge in their own ways.

In terms of language acquisition, teachers indicated that VR technology has clearer benefits for students' language knowledge and confidence than for their class grades, long-term memory, or self-correcting skills. The quantitative findings are supported by the qualitative results, in which teachers mostly emphasised the benefits of VR on students' language confidence rather than on any other specific skills or language abilities. This is consistent with previous findings that in VR-facilitated environments, students, especially ones at lower ability levels, have opportunities to build their language confidence without suddenly stepping out of their comfort zone (Ketelhut & Nelson, 2006; Loewen & Sato, 2018). This study also found benefits of VR in relation to developing students' intercultural awareness, which had not been empirically tested in previous studies. In short, VR is believed by teachers to be an ideal tool for improving students' communicative competence, learning experiences, and language confidence through its provision of social contexts for daily practice. VR has the potential to address some of the more long-standing issues in foreign language education, including out-of-context learning, students' lack of social knowledge and confidence in dealing with real-life situations, and students not making use of knowledge learnt in school contexts. VR, so far, seems to be an optimal solution to that problem, especially in the context of the COVID-19 pandemic in which face-to-face interaction is limited. According to teachers, the use of VR is not test-oriented, since its effects on all language skills and performance remain unclear.

In terms of teaching approaches, the teachers in this study believed that VR technology allows for the practice of constructivist pedagogy and student-centred approaches, in which students are placed at the centre and teachers only facilitate classroom activities to benefit students' learning. This finding supports theoretical works in the existing literature (e.g., Clarke et al., 2006; Jensen & Konradsen, 2018) that had previously argued that in VR-integrated classrooms, teachers' roles change from didactic and transmissive as knowledge holders into communicative and collaborative as facilitators. The teachers in the current study also agreed that the use of VR technology allows for the integration of different in-class activities and teaching materials. This result justifies Chuah et al.'s (2011) argument that VR-related materials diversify teaching resources, enabling the practices of more creative teaching approaches. It is, however, noted that most teachers who took part in this research are either current users of VR tools (e.g., ImmerseMe) or potential users who are interested in adopting the technology; therefore, their responses in terms of the benefits of VR for teaching practices might be different to those with more technology-resistant mindsets. As argued in section 2.4.1, teachers with technology-driven perspectives tend to use technology for higher-level uses (e.g., student collaboration or project-based learning), and thus their approaches to teaching are already more student-centred compared to those who have resistant views on technology adoption (Kern, 2006). Future research should aim for a broader sample of teachers with more diverse representations of mindsets toward technology to increase the reliability of the findings.

Drawbacks of VR. The quantitative findings support existing empirical literature (Fransson et al., 2020; Tan et al., 2016). The two biggest perceived drawbacks of VR are that it is time-consuming to redesign lessons around it and a lack of practice time and content teaching due to the extra time allocation for technology setup and to solve technical problems. These findings also empirically substantiate the third VR drawback proposed by Prensky (2016): lesson designing and time allocation. The qualitative findings further support Prensky's (2016) other drawbacks, which were (1) limitation in replacing peer-to peer interaction, (2) student's overreliance on technology, and (3) chaotic learning environments. Some teachers indicated that VR technology, in the long run, cannot replace face-to-face interaction and peer learning. In this view, VR should only be used to diversify learning resources and as a supplementary resource for individual practice. Teachers also commented that in most classroom settings, where the class size is large (25-27 students), it is impossible to manage the use of VR. The noise level is uncontrollable and the learning environment is chaotic as a result. Finally, while some teachers claimed that ImmerseMe did not offer enough cultural diversity, others disagreed. This might be because the content developed for some language courses is more or less culturally inclined than for others, or that teachers have different expectations of the courses in terms of cultural diversity. It is important to note that during the discussions, teachers rarely mentioned problems with lesson planning and curriculum, though these had emerged as important in the quantitative findings. One plausible explanation for this is that teachers might have perceived the lesson redesign as a factor that needs to be addressed rather than a drawback as such. This will be covered in more detail in the following section.

Barriers to VR Implementation. According to the teachers, the inability to deal with either technical issues or VR integration into the existing curriculum were the two least concerning challenges in VR implementation. However, the participants in this research were all either current users of ImmerseMe or were interested in using the technology, and therefore were likely to already have a good level of digital literacy. This suggests that the result cannot be generalised to the broader teacher population. Empirical studies on a wider-ranging sample of teachers are thus needed to further explore this finding.

The current study confirms VR implementation challenges that had been examined in previous studies. Specifically, the results demonstrated that the two biggest challenges to the effective use of VR are (1) lack of access to hardware facilities and to the VR program and (2) lack of teacher training and training materials. The former finding is consistent with previous studies (Chen, 2016; Sasinka et al., 2019) that had previously found that high cost and poor access to functioning facilities were two of the main barriers to the mass use of VR technology. The latter finding is also consistent with previous empirical studies (Kwon, 2019; Oak, 2018) that found that digital skills training and program-specific training are essential for the success of any technology adoption. These findings also align with the challenges listed in the TAM model (Davis et al., 1989) when describing potential challenges that teachers encounter at the beginning of technology implementation: technical support, equipment availability, and teacher training. The results are also consistent with the first-order barriers listed in the Stages of Concern (SoC) segment of the CBAM framework (Hall & Hord, 2006) as probable challenges during the refinement process of technology usage, namely expertise, training, and resources. This suggests that poor accessibility and teacher training remain the biggest concerns for both non-users and current users.

In addition, current users of ImmerseMe further suggested a crucial challenge that they experienced during the adoption and usage process, that is, the functionality of the voice recognition feature. Most teachers argued that that the voice recognition feature did not work well in classroom settings due to excessive noise levels. This problem typically becomes the main reason for both teachers' and students' frustration and loss of learning interest. This challenge to the effective use of VR aligns with the SoC second-order barrier, which highlights how technical failures can lead to dissatisfaction and thus hinder technology use. However, empirical studies examining the voice recognition features in VR applications are still limited. Future research could investigate this issue based on existing literature on other applications (e.g., smartphones or general language learning apps) in order to optimise teaching and learning experiences with VR technology.

Implementation Enabling Factors. To make the most effective use of VR, it is important to address the implementation challenges mentioned above, including providing access to hardware facilities and VR programs as well as teacher training. The quantitative results for the *implementation enabling factors* also indicated that the existence of functioning facilities is perceived by teachers as the most important factor in effective VR use. These findings empirically confirm the factors under the sub-group *the context (school)* in Zhao et al.'s (2002) *Conditions for Classroom Technology Innovation* model, which proposed that the main factors required to make effective use of a new technology from the school's perspective include (1) technology infrastructure, (2) human resources and teacher training, and (3) an organisational culture that allows the sharing of and collaboration in pedagogical practices.

Apart from teacher training and facilities, the qualitative findings of this study suggested additional factors that are important in the use of VR. These were mainly focused on curriculum or lesson design. One explanation for the difference in focus might be that functioning facilities, internet connections, and teacher training are out of a teacher's control, and that in the discussions teachers focused their discussion on the factors that they could more easily control. Teachers may also perceive the redesign of curriculum and lesson planning as a factor in effective VR use rather than as a drawback or challenge. If this is true, it may be that the reconstruction of curriculum is already seen as a part of their responsibility as teachers, with or without the adoption of new technology, and thus the adjustment is a normal obligation rather than an added burden.

Teachers in this study recommended that (1) the curriculum be structured in an adaptive way, in case of any technical issues arising during teaching, (2) teachers should ensure that students have sufficient knowledge before VR practice, and (3) teachers should build flexible teaching sequences to avoid lesson repetition and cognitive overload for students. These findings are consistent with the factors under the sub-group *the innovator (teacher)* in Zhao et al.'s (2002) *Conditions for Classroom Technology Innovation* model. This model proposed several factors required to make effective use of a new technology from teachers' perspectives, including (1) knowledge of the technology and (2) pedagogy-technology curriculum compatibility. This finding also aligns with the third-order barriers proposed in the SoC model (Hall & Hord, 2006), which view the redesign of curriculum and lesson planning as the last barrier to effective use of technology in classroom settings. Finally, according to teachers, students' theoretical understanding of grammar is the least important factor in the effective use of VR in language classrooms. This finding suggests that the nature of the VR-facilitated environment is not ideal for grammar-based test and exam preparation.

5.1.2. Research Question 2

When compared to the standard alpha, there were several differences in ImmerseMe users' and non-users' perceptions regarding the areas of learning supported by VR and the drawbacks of VR use. Specifically, the direction of the trend in difference for *areas of learning supported by VR* was toward current users, while that for *drawbacks of VR technology* was toward non-users. A plausible explanation might be that ImmerseMe users not only developed their understanding of the technology functions but also gained hands-on experience with VR use, and therefore were more confident in specifying the areas of learning best supported by VR technology. On the contrary, non-users might be more resistant to VR than current users because they find the idea of using VR for language teaching new and challenging.

However, when compared with the adjusted alpha, this difference no longer existed. In other words, there was no evidence of differences in the perceptions of VR non-users and current users regarding all five outcome measures (*areas of learning supported by VR*, *VR benefits*, *VR drawbacks*, *implementation challenges*, and *implementation enabling factors*) after controlling for teacher age and years of teaching experience. This finding is inconsistent with existing theories, including the TAM framework (Davis et al., 1989) and the SoC model (Hall & Hord, 2006), which suggest that there might be differences in the perceptions of users depending on how external factors and the level of exposure to the technology (non-usage, recent usage, or long-term usage) affect users' mindsets and decisions. A plausible explanation for this contradictory finding is that the current study did not differentiate users based exactly on these categories (non-users, new users, and long-term users), which might have affected the reliability of the results. Specifically, there is a possibility that most current users in this study were quite new to the technology, and therefore their mindsets were not substantially different from those of non-users. This suggests that more empirical studies should be conducted to compare the perceptions of teachers regarding VR technology in consideration of the distinction between different levels of usage. It is also noted that VR is still a fairly new technology; thus, the resources that would normally help users overcome the existing challenges and thus change their perceptions are still limited.

5.2. Future Research and Pedagogical Implications

5.2.1. Future Research

Areas of Learning Supported by VR. It is suggested that future research focus more on areas other than speaking. For example, exploratory studies should be conducted to examine how the content engagement offered by VR tools (e.g., ImmerseMe) might assist listening and reading comprehension. Comparative studies should also be undertaken to compare the effects of VR with other teaching methods (e.g., listening/reading exercises and audiobooks or textbooks) on listening and reading skills. Regarding writing, longitudinal studies are needed to determine the long-term effects of VR on writing performance. These studies should take the demands of different writing styles into consideration.

Benefits and Drawbacks of VR Technology in Language Education. As there has already been considerable empirical research regarding the benefits of VR, future research should go beyond questions being asked in this study by focusing on specific pedagogical practices that can optimise the benefits of VR and minimise its drawbacks. For instance, this study showed that VR allows for different teaching approaches (e.g., student-centred). Future studies should explore in what ways and under what conditions teachers can best practice different teaching approaches in VR-integrated classroom settings. In addition, the findings of this study suggest that the time-consuming nature of lesson redesign is one of the biggest drawbacks of VR use in classroom settings; therefore, future research should provide teachers with research-driven pedagogy to assist their redesign of curriculum and lessons.

Voice Recognition Feature. The results of this study demonstrated that teachers and students were frustrated during their lessons with VR because of the poor functioning of the voice recognition feature. Future studies should examine the main issues and solutions for deal with this feature in order to optimise teaching and learning experiences with VR technology.

Differences in Teachers' Perceptions. Another potential avenue for future research is to explore the differences in perception between users with different levels of experience. Even though the study results demonstrated that there are no differences in the perceptions of non-users and current users regarding all explored aspects of VR, this result is tentative for a number of reasons (see section 5.3 below). It is therefore advised that more empirical studies be conducted to evaluate this result in consideration of the distinction between users with different levels of experience.

Generalisability. Finally, it should be noted that one of the key issues with this study is that the results cannot easily be generalised to the wider teacher population due to the nature of the study participants. Because of the recruitment process, which targeted members of an ImmerseMe email list, participants were more likely to have technology-driven mindsets than the general teacher population. Therefore, empirical studies using a broader sample are recommended to improve the generalisability of the findings.

5.2.2. Pedagogical Implications

In general, speaking and listening were perceived by the teachers in this study as the learning areas that benefited most from the use of VR. This suggests that prospective teachers should consider setting up initial VR practice of speaking and listening skills before exploring how VR might benefit other skill sets (e.g., reading and writing). The findings also indicate that VR-integrated materials and exercises (e.g., pronunciation practice) are beneficial for individual practice, and therefore are ideal for self-practice activities in home settings, especially in the context of the COVID-19 pandemic.

There are still a number of challenges that need to be addressed in order to implement and use VR effectively. It is advised that schools, teachers, and program organisers collaborate to optimise VR teaching and learning experiences. From the school's perspective, it is crucial to provide teachers and students with access to required hardware and software facilities and a stable internet connection to avoid any disruptions to VR implementation. Schools should also provide digital skills training and VR program-specific training to make sure that teachers are well-equipped for the use of VR in classroom settings. From the teachers' perspectives, it is important to discuss and collaborate to redesign the curriculum and carry out lesson planning, enrich teaching resources, and prepare backup plans to cope with any technical issues that might arise. It is also key that teachers show an adaptive mindset when integrating VR technology into the curriculum, and that they carefully build flexible teaching sequences to avoid lesson repetition and student cognitive overload. Teachers should also work on improving their own digital literacy.

Finally, to optimise student learning experiences, teachers and program organisers should target voice recognition features in VR programs. Teachers should prepare students, and especially lower-level learners, with speaking drills and pronunciation practice, to prepare them for VR activities. Program organisers should improve the accuracy of the feature, ensuring that they include different accents and incorporate diverse cultural elements into the scenarios. There are also other ways in which VR tools can be better designed to support teachers' VR use in classroom settings. For instance, VR programs could include drill exercises to practice individual sounds with the voice recognition feature before starting on whole conversations.

5.3. Research Limitations

Direction of causation. The first limitation of this research is that the direction of causation cannot be determined. For instance, it may be either that the use of ImmerseMe causes teachers to perceive VR in a certain way, or that existing perceptions lead teachers to use VR, or both.

Insufficient sample size. The sample for this study was insufficient by 20 participants, and this affects the validity of the research findings. According to Faber and Fonseca (2014), a small sample size might lead to Type II error, i.e., the null hypothesis is incorrectly accepted and no difference between the study groups is found. In this case, there is the possibility that the result of research question 2, in which no significant differences were found between the perceptions of non-users and current users, was affected by the effects of insufficient sample size. Replication studies with a sufficient sample size are strongly encouraged.

Sampling. As mentioned in section 3.3.4.2, the qualitative sample was not representative of the whole sample population in terms of levels of usage and country of residence. This might have caused contradictions between quantitative and qualitative research results. It should also be noted that most of the teachers who took part in this research were either current users of VR tools (e.g., ImmerseMe) or potential users who were interested in adopting the technology; therefore, their responses and opinions cannot be generalised to the broader population.

Data collection. There was a methodological limitation during the process of qualitative data collection. Participants were grouped based on their time availability and locations rather than demographic backgrounds. This inevitably affected the heterogeneity of the discussions and the diversity of opinions and perceptions, since teachers from the same location probably share similar teaching experiences, school infrastructure conditions, and classroom culture and beliefs. This unavoidably had an impact on the nature of the discussions, limiting the possibilities for teachers from diverse backgrounds and levels of use to share their real-life experiences and have meaningful and constructive discussions. Future studies should aim for more heterogenous groupings, as this may trigger more productive discussions.

Recruitment Process. A further methodological limitation involved the recruitment process. The two groups of participants (non-users and current users) were not recruited from the same sources. The ImmerseMe current users were exclusively recruited via the ImmerseMe email list, while the non-users were recruited via both the mailing list and social media platforms. The participants recruited via social media tended to be the same age as the researcher, and therefore were likely to be fresh graduates who were inexperienced secondary teachers. This might have affected the covariate test results and the findings (MANCOVA and ANCOVA analyses) relating to research question 2.

Poor scale reliability. The final limitation of this study involves the poor reliability of the Likert rating scale for the outcome measure *implementation challenges* ($\alpha < .8$). This value suggests that the scale might not give the same results if the survey questionnaires were to be taken again under similar conditions. This may have affected the consistency of results related to the outcome measure *implementation challenges*. In addition, since the items have low internal consistency, they are likely not to correlate highly with each other. This means that the MANCOVA and ANCOVA results for research question 2, in which the items in this outcome measure were averaged together to make one scale, might have been affected in terms of consistency.

5.4. Conclusion

Regardless of the popularity and potential of VR in the language classroom, research into its use in foreign language education is still in its infancy. This study aimed to empirically examine various aspects of VR use in language learning and teaching, including the areas of learning supported by VR, VR's benefits and drawbacks, implementation challenges, and important factors in the effective use of VR. In general, the findings of this study are highly consistent with findings from earlier research. In terms of areas of learning, the results show that teachers perceived the clearest VR benefits to be for speaking and listening skills as compared to other skills. It is therefore advised that prospective teachers set up initial VR practice for speaking and listening before exploring VR's benefits for other skill sets (e.g., reading and writing) to increase the possibility of positive learning outcomes. In terms of VR benefits, the use of VR in language classrooms appears to benefit contextualised learning, in which the VR application (e.g., ImmerseMe) provides students with experiences of the authentic culture of the target language, assisting them to create meaning from the knowledge they are gaining. VR is also believed by teachers to allow the practice of constructivist pedagogy and student-centred approaches, both of which benefit students in the long term. Regarding language acquisition, the teachers in this study saw clearer benefits of VR technology for students' language knowledge and confidence, as compared to their class grades, long-term memory retention, and self-correcting skills. This suggests that VR might not be the most ideal tool for exam preparation; instead, it can help teachers to address common problems in language classrooms, including out-of-context learning and students' lack of pragmatic competence, authentic knowledge, and confidence in conducting real-life conversations. The biggest drawbacks of VR perceived by the teachers in this study included the time-consuming nature of lesson redesign, a lack of practice time and content teaching in class, its limitation in completely replacing classroom interaction, and occasionally, chaotic learning environments. The perceived challenges to VR implementation included lack of access to hardware facilities and to the VR program, lack of teacher training and training materials, and an insensitive voice recognition feature that led to student and teacher frustration with the program. It is thus strongly advised that schools, teachers, and program organisers cooperate to maximise the benefits of VR, minimise its drawbacks, and address the existing challenges. Schools should provide teachers with access to required facilities and with high-quality teacher training. Teachers themselves should collaborate to work on curriculum redesign, lesson planning, and teaching sequences, while improving their own digital literacy so they can deal with any technical issues that might arise. Finally, teachers and program organisers should work together on improving the functionality of the voice recognition feature to optimise students' learning experience. These pedagogical recommendations will contribute to more effective use of VR technology in foreign language education, especially in the context of the COVID-19 pandemic.

References

Abdous, M., Camarena, M., & Facer, B. R. (2009). MALL technology: Use of academic podcasting in the foreign language classroom. *ReCALL, 21*(1), 76–95. https://doi:10.1017/S0958344009000020

Accuosti, J. (2014). *Factors affecting education technology success* [Conference Paper]. Proceeding of the ASEE 2014 Zone I Conference (pp. 11–17). https://www.asee.org/documents/zones/zone1/2014/Student/PDFs/112.pdf

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behaviour*. Prentice-Hall.

Ajzen, I. (1985). *From intentions to actions: A theory of planned behaviour*. Springer.

Alam, M. M. (2017). Constructivism and the classroom curriculum. *The International Journal of Indian Psychology, 5*(1), 20–29. https://doi:10.25215/0501.103

Alemi, M., & Khatoony, S. (2020). Virtual Reality Assisted Pronunciation Training for Young EFL learners. *Teaching English with Technology, 20*(4), 59–81.

Alfalah, S. (2018). Perceptions toward adopting virtual reality as a teaching aid in information technology. *Education and Information Technologies, 23*(6), 2633–2653. https://doi.org/10.1007/s10639-018-9734-2.

Amineh, R. J., & Asl, H. D. (2015). Review of constructivism and social constructivism. *Journal of Social Sciences, Literature and Languages, 1*(1), 9–16.

Ampatzoglou, A., & Chatzigeorgiou, A. (2007). Evaluation of object-oriented design patterns in game development. *Information and Software Technology, 49*(7), 445–454.

Annetta, L., & Minogue, J., & Holmes, S., & Cheng, M. (2009). Investigating the impact of video games on high school students' engagement and learning about genetics. *Computers & Education, 53*(3), 74–85.

Ausburn, L. J., & Ausburn, F. B. (2008). New Desktop Virtual Reality Technology in Technical Education. *Journal of Educational Technology, 4*(4), 48–61.

Bajorek, J. P. (2019). L2 Virtual Reality-Based ImmerseMe Software: Speaking, Interactivity, & Time. *Computer Assisted Language Instruction Consortium (CALICO), 37*(3), 321–330.

Berge, Z. (2008). Changing instructor's roles in virtual worlds. *Quarterly Review of Distance Education, 9*(2), 407–415.

Bhattacharjee, J. (2015). Constructivist Approach to Learning– An Effective Approach of Teaching Learning. *International Research Journal of Interdisciplinary & Multidisciplinary Studies, 1*(5), 65–74.

Blyth, C. (2018). Immersive technologies and language learning. *Foreign Language Annals, 51*(1), 225–232. https://doi.org/10.1111/flan.12327.

Boschman, F. B., McKenney, S., & Voogt, J. (2015). Exploring teachers' use of TPACK in design talk: The collaborative design of technology-rich early literacy activities. *Computers & education, 82*(3), 250–262. https://doi.org/10.1016/j.compedu.2014.11.010

Burkle, M., & Meredith, S. (2008). Building bridges between University and Industry: theory and practice. *Education & Training, 50*(2), 199–215.

Burston, J. (2015). Twenty years of MALL project implementation: A meta-analysis of learning outcomes. *ReCALL, 27*(1), 4–20. https://doi:10.1017/S0958344014000159

Cardwell, S. (2018, March 19). *ImmerseMe: VR-based language learning*. ImmerseMe Blog. https://immerseme.co/blog

Chen, C. J. (2018). The interplay of tasks, strategies and negotiations in Second life. *Computer Assisted Language Learning, 31*(8), 960–986.

Chen, C. J., Toh, S. C., & Fauzy, W. M. (2004). The theoretical framework for designing desktop virtual reality-based learning environments. *Journal of Interactive Learning Research ,15*(2),147–167.

Chen, C. J., Toh, S. C., & Ismail, W. M. F. W. (2005). Are Learning Styles Relevant to Virtual Reality?. *Journal of Research on Technology in Education, 38*(2), 123–141.

Chen, Y. L. (2016). The Effects of Virtual Reality Learning Environment on Student Cognitive and Linguistic Development. *Asia-Pacific Education Research, 25*(4), 637–646. https://doi.org/10.1007/s40299-016-0293-2

Chuah, K., Chen, C. J., & Teh, C. S. (2011). *Emotional affordances in virtual reality-based learning environments*. LAMBERT Academic Publishing.

Cipresso, P. (2015). Modeling behavior dynamics using computational psychometrics within virtual worlds. *Frontier Psychology, 6*(4), 1725–1754.

Clarke, J., Dede, C., Ketelhut, D. J., & Nelson, B. (2006). A design-based research strategy to promote scalability for educational innovations. *Educational Technology, 46*(3), 27–36.

Cochrane, T. (2016). Mobile VR in education: from the fringe to the mainstream. *International Journal of Mobile Blended Learn, 8*(2), 44–60.

Cooke-Plagwitz, J. (2008). New directions in CALL: an objective introduction to second life. *CALICO Journal, 25*(3), 547–557.

Creswell, J. W., & Plano, V. L. (2007). *Designing and Conducting Mixed Methods research*. Sage Publications.

Cuban, L. (2001). *Oversold and underused: Computers in the classroom*. Harvard University Press.

Cubukcuoglu, B. (2013). Factors enabling the use of technology in subject teaching. *International Journal of Education and Development using ICT*, 9(3), 450–478.

Cummings, A. D. (2005). *Administrative and pedagogical uses of computers in foreign language classrooms: A survey of Spanish teachers' beliefs and practices* [Doctoral dissertation, University of Iowa]. ProQuest Dissertations & Theses Global.

Cummins, J. (2000). *Language, Power and Pedagogy*. Multilingual Matters.

Dalgarno, B., & Lee, M. J. (2010). What are the learning affordances of 3-D virtual environments?. *British Journal of Educational Technology, 41*(1), 10–32.

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. *Management Science, 35*(2), 982–1003.

Davis, F.D. (1989) "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology" MIS Quarterly 13(3), pp. 319-340.

Dawley, L., & Dede, C. (2013). Situated Learning in Virtual Worlds and Immersive Simulations. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of Research on Educational Communications and Technology* (pp. 723–734), Springer.

Demouy, V., Jones, A., Kan, Q., Kukulska-Hulme, A., & Eardley, A. (2016). Why and how do distance learners use mobile devices for language learning?. *The EuroCALL Review, 24*(1), 10–24. https://doi.org/10.4995/eurocall.2016.5663

Dolgonsoz, E., Yildirim, G., & Yildirim, S. (2018). The effect of virtual reality on EFL writing performance. *Journal of Language and Linguistic Studies, 14*(1), 278–292.

Ellis, R. (2008). Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics, 18*(1), 4–22. https://doi.org/10.1111/j.1473-4192.2008.00184.x

Ertmer, P. A. (2005). Teacher pedagogical beliefs: The final frontier in our quest for technology integration?. *Educational Technology Research and Development, 53*(4), 25–39.

Esfijani, A., & Zamani, B. E. (2020). Factors influencing teachers' utilisation of ICT: the role of in-service training courses and access. *Research in Learning Technology, 28*(1), 4-24. https://doi.org/10.25304/rlt.v28.2313

Eun, B., & Lim, H. S. (2009). A Sociocultural View of Language Learning: The Importance of Meaning-Based Instruction. *TESL Canada Journal, 27*(1), 12–26. https://doi.org/10.18806/tesl.v27i1.1031

Faber, J., & Fonseca, L. M. (2014). How sample size influences research outcomes. *Dental press journal of orthodontics, 19*(4), 27–29. https://doi.org/10.1590/2176-9451.19.4.027-029.ebo

Fransson, G., Holmberg, J., Westelius, C. (2020). The challenges of using head mounted virtual reality in K-12 schools from a teacher perspective. *Education and Information Technologies, 25*(4), 3383–3404. https://doi.org/10.1007/s10639-02

Freina, L., & Ott, M. (2015). *A Literature Review on Immersive Virtual Reality in Education*. State of the Art and Perspectives Publications.

Gander, F., Proyer, R. T., Ruch, W., Wyss, T. (2013). Strength-based positive interventions: further evidence for their potential in enhancing well-being and alleviating depression. *Journal of Happiness Studies, 14*(5), 1412–1459. https://doi:10.1007/s10902-012-9380-0

Gillespie, J. (2020). CALL research: Where are we now? *ReCALL, 32*(2), 127–144.

Grant, S. J., Huang, H., & Pasfield-Neofitou, S. E. (2013). Language learning in virtual worlds: The role of foreign language and technical anxiety. *Journal of Virtual Worlds Research, 6*(1), 1–9.

Hafner, P., Ducker, J., Schlatt, C., & Ovtcharova, J. (2018). *Decision support method for using virtual reality in education based on a cost-benefit-analysis* [Conference Paper]. 4th International Conference of the Virtual and Augmented Reality in Education (pp. 103–112). https://doi.org/10.1016/j.jksuci.2014.03.023.

Halcomb, E. J., Gholizadeh, L., DiGiacomo, M., Phillips, J., & Davidson, P. M. (2007). Literature review: Considerations in undertaking focus group research with culturally and linguistically diverse groups. *Journal of Clinical Nursing, 16*(6), 1000–1011. https://doi.org/10.1111/j.1365-2702.2006.01760.x

Hall, G. E., & Hord, S. M. (2006). *Implementing change: Patterns, principles, and potholes*. Pearson Education.

Hamilton, D., McKechnie, J., Edgerton, E. (2020). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computer Education, 8*(1), 1–32. https://doi.org/10.1007/s40692-020-00169-2

Harrow, A. (1972). *A Taxonomy of the Psychomotor Domain*. David McKay.

Howard-Jones, P., Ott, M., Leeuwen, T., & De Smedt, B. (2015). The potential relevance of cognitive neuroscience for the development and use of technology-enhanced learning. *Learning, Media and Technology, 40*(3), 131–151.

Huang, H. M., & Liaw, S. (2018). An Analysis of Learners' Intentions Toward Virtual Reality Learning Based on Constructivist and Technology Acceptance Approaches. *The International Review of Research in Open and Distributed Learning*, 19(2), 544–589.

Huang, H. M., Rauch, U., & Liaw, S. S. (2010). Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers and Education*, *55*(2), 1171–1182.

Hussein, M., & Natterdal, C. (2015). *The benefits of virtual reality in education: A comparison study*. [Unpublished bachelor thesis]. Chalmers University of Technology, Gothenburg.

Jensen, L., & Konradsen, F. (2018). A review of the use of virtual reality head-mounted displays in education and training. *Education and Information Technologies*, *23*(4), 1515–1529. https://doi.org/10.1007/s10639-017-9676-0

Judson, E. (2006). How Teachers Integrate Technology and Their Beliefs About Learning: Is There a Connection?. *Journal of Technology and Teacher Education, 14*(3), 581–597.

Jung, I. (2005). ICT-Pedagogy Integration in Teacher Training: Application Cases Worldwide. *Educational Technology & Society, 8*(2), 94–101.

Kern, R. (2006). Perspectives on Technology in Learning and Teaching Languages. *TESOL Quarterly, 40*(1), 183–210. https://doi:10.2307/40264516

Kim, J. (2002). The effects of teacher training as constructivist on teacher behaviour and student achievements. *Educational Development Research Institute, 18*(2), 425–452.

King, W., & He, J. (2006). A meta-analysis of the Technology Acceptance Model. *Information & Management, 43*(5), 740–755.

Kluge, S., & Riley, E. (2008). *Teaching in Virtual Worlds: Opportunities and Challenges*. Pearson Education.

Kohlberg, L. (1968). Early Education: A Cognitive-Developmental View. *Child Development, 39*(5), 1013–1062.

Kwon, C. (2019). Verification of the possibility and effectiveness of experiential learning using HMD-based immersive VR technologies. *Virtual Reality, 23*(1), 101–118.

Lai, C. (2017). *Autonomous Language Learning with Technology: Beyond the Classroom.* Bloomsbury.

Lawrence, G., & Ahmed, F. (2020). *Avatar teaching and learning: Examining language teaching and learning practices in Virtual Reality Environments*. IGI Global.

Lederer, A. L., Maupin, D. J., Sena, M. P., and Zhuang, Y. (2000). The Technology Acceptance Model and the World Wide Web. *Decision Support Systems, 29*(3), 269–282.

Lee, E. A., & Wong, K. W. (2008). *A Review of Using Virtual Reality for Learning*. Springer.

Lee, I., & Wong, K. (2014). Bringing Innovation to EFL Writing: The Case of a Primary School in Hong Kong. *The Asia-Pacific Education Researcher, 23*(2), 159–163.

Levy, M., & Kennedy, C. (2005). Learning Italian via Mobile SMS. In A. Kukulska-Hulme, & J. Traxler (Eds.), *Mobile Learning a Handbook for Educators and Trainers* (pp. 76-83). Routledge.

Lin, T., & Lan, Y. J. (2015). Language Learning in Virtual Reality Environments: Past, Present, and Future. *Journal of Educational Technology Society*, *18*(4), 486–497.

Loewen, S., & Sato, M. (2018). Instructed second language acquisition and English language teaching. *Theory, research, and pedagogy, 4*(1), 1–19.

Lu, H. M., Lou, S. J., Papa, C., & Chung, C. C. (2011). *Study on influence of adventure game on English reading confidence, motive and self-efficacy*. Springer.

Madathil, K. C., Frady, K., Hartley, R., Bertrand, J., Alfred, M., & Gramopadhye, A. (2017). An empirical study investigating the effectiveness of integrating virtual reality-based case studies into an online asynchronous learning environment. *Computers in Education Journal, 8*(3), 1–10.

Mangina, E. (2017). *3D learning objects for augmented/virtual reality educational ecosystems* [Conference Paper]. 23rd International Conference on virtual system & Multimedia (pp. 1–6). https://ieeexplore.ieee.org/xpl/conhome/8340209/proceeding

Marton, F., & Booth, S. A. (1997). *Learning and awareness*. Lawrence Erlbaum Associates Publishers.

Mayer, R. E. (2014). *Cognitive theory of multimedia learning.* Cambridge University Press.

Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). Cambridge University Press.

McGarrell, H., & Alvira, R. (2013). Innovation in techniques for teacher commentary on ESL writers' drafts. *OLBI Working Papers, 5*(2), 37-55. http://dx.doi.org/10.18192/olbiwp.v5i0.1117

Means, B., Toyama, Y., Murphy, R., Bakia, M., & Jones, K. (2010). *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. U.S. Department of Education Publications.

Merchant, Z., Goetz, E. T., Cifuentes, L., Keeney-Kennicutt, W., & Davis, T. J. (2014). Effectiveness of virtual reality-based instruction on students' learning outcomes in K-12 and higher education: A meta-analysis. *Computers & Education, 70*(2), 29–40.

Minocha, S. (2015). The state of virtual reality in education: Shape of things to come. *International Journal of Engineering Research, 4*(11), 596–598.

Minocha, S., Tudor, A. D., & Tilling, S. (2017). Affordances of mobile virtual reality and their role in learning and teaching. In L. Hall, T. Flint, S. O'Hara, & P. Turner (Eds.), *Proceedings of the 31st international BCS human computer interaction conference* (pp. 345–367). ScienceOpen. https://doi.org/10.14236/ewic/HCI2017.44.

Mollaei, F., & Riasati, M. (2013). Teachers' Perceptions of Using Technology in Teaching EFL. *International Journal of Applied Linguistics & English Literature*, 2(2) 13–22.

Molna, G. (2016). *How Digital and Virtual Life Trapped in Visuality and New Media Systems Affect Teaching and the Learning Process*. Peter Lang Publisher.

Muhanna, A. M. (2015). Virtual reality and the CAVE: Taxonomy, interaction challenges and research directions. *Journal of King Saud University Computer and Information Sciences*, *27*(3), 344–361.

Mroz, A. (2015). The development of second language critical thinking in a virtual language learning environment: A process-oriented mixed-method study. *CALICO Journal*, *32*(4), 528–553.

National Centre for Education Statistics, (2015). Virtual and Augmented Reality effects on K-12, higher and tertiary education students' 21st century skills. U.S Department of Education. https://nces.ed.gov/pubs2015/2015144.pdf

National Center for Education Statistics, (2019). Virtual Reality and Its Application in Education: Survey. U.S Department of Education. https://www.mdpi.com/2078-2489/10/10/318/htm

Nelson, B., & Ketelhut, D. (2007). Scientific Inquiry in Educational Multi-user Virtual Environments. *Educational Psychology Review, 19*(8), 265–283.

Oak, J. W. (2018). Educational needs for virtual reality-based program development: Focusing on practice of medication nursing. *Asia Life Science, Supplement, 15*(1), 891–899.

O'Brien, G., Mary, T. (2009). Virtual Immersion: The Role of CAVE and PC Technology. *CALICO Journal,* 26(2), 337–362.

Palak, D., & Walls, R. (2009). Teachers' beliefs and technology practices: A mixed methods approach. *Journal of Research on Technology in Education, 41*(4), 417–441.

Pantelidis, V. (2009). Reasons to use virtual reality in education and training courses and a model to determine when to use virtual reality. *Themes in Science and Technology Education, 2*(1), 59–70.
Papin, K. (2019). Can 360 virtual reality tasks impact L2 willingness to communicate?. In T. Peppi, J. Juha, B. Linda, & T. Sylvie (Eds.), *Future-proof CALL: language learning as exploration and encounters* (pp. 243–248). Routledge.
Parmaxi, A. (2020). Virtual reality in language learning: a systematic review and implications for research and practice. Interactive Learning Environments, 6(1), 1–13.
Parmaxi, A., & Zaphiris, P. (2017). Web 2.0 in computer-assisted language learning: A research synthesis and implications for instructional design and educational practice. *Interactive Learning Environments, 25*(6), 704–716.
Peterson, M. (2011). Towards a research agenda for the use of three-dimensional virtual worlds in language learning. *CALICO Journal, 29*(1), 67–82. https://doi.org/10.11139/cj.29.1.67-80
Prensky, M. (2016). *Education to better their world: Unleashing the power of 21st-century kids*. Teachers College Press.
Prensky, M. (2006). Listen to the natives. *Educational Leadership, 63*(4), 8–13.
Pritchard, A. (2017). *Ways of learning: Learning theories for the classroom* (4th ed.). Routledge.
Redmond, P., Albion, P., & Maroulis, J. (2005). Intentions and Reality: Preservice ICT Integration during Professional Experience. *Proceedings of SITE 2005: Society for Information Technology & Teacher Education International Conference*, *USA, 116*(17), 1566–1571.
Sasinka, C., Stachon, Z., Sedlak, M., Chmelik, J., Herman, L., Kubicek, P., Sasinkova, A., Dolezal, M., Tejkl, H., Urbanek, T., Svatonova, H., Ugwitz, P., & Jurik, V. (2019). Collaborative immersive virtual environments for education in geography. *International Journal of Geo-Information, 8*(3), 1–15.
Schwienhorst, K. (2002). The State of VR: A Meta-Analysis of Virtual Reality Tools in Second Language Acquisition. *Computer Assisted Language Learning, 15*(2), 221–239.
Serrano, R., Llanes, A., & Tragant, E. (2011). Analyzing the Effect of Context of Second Language Learning: Domestic Intensive and Semi-Intensive Courses vs. Study Abroad in Europe. *International Journal of Educational Technology and Applied Linguistics, 39*(1), 4–35. https://10.1016/j.system.2011.05.002
Sharma S., Agada, R., & Ruffin, J. (2013). Virtual reality classroom as ana constructivist approach. *ReCALL, 3*(4), 1–15.
Slater, M. (2018). Immersion and the illusion of presence in virtual reality. *British Journal of Psychology, 109*(3), 431–433.
Solomonidou, C. (2016). Constructivist design and evaluation of interactive educational software: a research-based approach and examples. *Ανοικτή Εκπαίδευση: το περιοδικό για την Ανοικτή και εξ Αποστάσεως Εκπαίδευση και την Εκπαιδευτική Τεχνολογία, 5*(6), 340–375.

Soto, B., Ocampo, J. T., Colon, D., Alejandro, L. O. (2020). Perceptions of ImmerseMe Virtual Reality Platform to Improve English Communicative Skills in Higher Education. *International Journal of Interactive Mobile Technologies, 14*(4), 56–78.

Southgate, E., & Smith, S. P. (2017). *Designing and conducting research using immersive technologies in schools: Seven observations* [Conference Paper]. IEEE Virtual Reality Workshop on K-12 Embodied Learning through Virtual & Augmented Reality. https://doi.org/10.1109/KELVAR.2017.7961564.

Stepp-Greany, J. (2002) Student Perceptions on Language Learning in a Technological Environment: Implications for the New Millennium. *Language Learning & Technology, 6*(2), 165–180.

Stickler, U., & Hampel, R. (2010). CyberDeutsch: Language Production and User Preferences in a Moodle Learning Environment. *CALICO Journal, 28*(1), 49–73.

Stockwell, G. (2010). Using Mobile Phones for Vocabulary Activities: Examining the Effect of the Platform. *Language Learning and Technology*, *14*(2), 95–110.

Suh, A., & Prophet, J. (2018). The state of immersive technology research: A literature analysis. *Computers in Human Behaviour, 86*(3), 77–90.

Sumak, B., & Hericko, M., & Pusnik, M. (2011). A meta-analysis of e-learning technology acceptance: The role of user types and e-learning technology types. *Computers in Human Behaviour, 27*(6), 2067–2077.

Tai, T. Y., & Chen, H. H. J. (2021). The Impact of Immersive Virtual Reality on EFL Learners' Listening Comprehension. *Journal of Educational Computing Research*, *30*(2), 185–203. https://doi.org/10.1177/0735633121994291

Tan, S., O'Halloran, K. L., & Wignell, P. (2016). Multimodal research: Addressing the complexity of multimodal environments and the challenges for CALL. *ReCALL, 28*(3), 253–273. https://doi.org/10.1017/S0958344016000124

Tarhini, A., Teo, T., & Tarhini, T. (2015). A cross-cultural validity of the E-learning acceptance measure in Lebanon and England: a confirmatory factor analysis. *Education and Information Technologies, 21*(5), 1269–1282. https://doi:10.1007/s10639-015-9381-9

Teo, T. (2011). Factors influencing teachers' intention to use technology: Model development and test. *Computers & Education, 57*(4), 2432–2440.

Thomson, S. (2015). *Policy Insights: Australian Students in Digital World*. Australian Council for Educational Research Publications.

Trigwell, K., Prosser, M. (2004). Development and Use of the Approaches to Teaching Inventory. *Educational Psychology Review, 16*(4), 409–424. https://doi.org/10.1007/s10648-004-0007-9

Tsai, C. C., & Chai, C. S. (2012). The third-order barrier for technology-integration instruction: Implications for teacher education. *Australasian Journal of Educational Technology, 28*(6), 186–201. https://doi.org/10.14742/ajet.810

UK Authority. (2019). *VR and AR attract education sector interest*. UK Authority Office. https://www.ukauthority.com/articles/vr-and-ar-attract-education-sector-interest/

Utami, F., Andika, W. D., & Sumarni, S. (2021). *Introduction to Sea Animals with Augmented Reality Based Flashcard for Early Childhood* [Conference Presentation]. 4th Sriwijaya University Learning and Education International Conference, Indonesia.

Venkatesh, V., & Davis, F. D. (1996). A Model of the Antecedents of Perceived Ease of Use Development and Test. *Decision Sciences, 27*(3), 451–481.

Vogel, G. E., & Bowers, C., & Clint, B. (2006). Using Virtual Reality with and without Gaming Attributes for Academic Achievement. *Journal of research on technology in education, 39*(4), 105–119.

Winthrop, R. (2020, March 31). COVID-19 and school closures: What can countries learn from past emergencies? *Centre for Universal Education*. https://www.brookings.edu/research/covid-19-and-school-closures-what-can-countries-learn-from-past-emergencies/

Wu, H., & Lee, S. W., & Chang, H., & Liang, J. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education, 62*(2), 41–49.

Yamazaki, K. (2018). Computer-assisted learning of communication (CALC): A case study of Japanese learning in a 3D virtual world. *ReCALL, 30*(2), 214–231. https://doi.org/10.1017/S0958344017000350

Yip, D., & Cheung, D. (2005). Teachers' Concerns on school-based assessment of practical work. *Journal of Biological Education, 39*(3), 156–162.

Zhang, W. (2014). On college oral English teaching in the base of virtual reality technology. *Applied Mechanic Materials, 68*(7), 2427–2430.

Zhao. Y., Pugh, K., Sheldon, S., & Byers, J. L. (2002). Conditions for classroom technology innovations. *Teachers College Records, 104*(3), 482–515.

Appendix A: Participant Outreach Email

CALL FOR PARTICIPANTS

Dear teachers,

If you are a foreign language teacher and currently teaching secondary to higher education levels, you are invited to take part in a research which investigates teachers' perceptions of Virtual Reality (VR-based) technology in language learning. With your contribution, the research aims to shed light on practical pedagogical applications for the future use of VR technology in language education.

There are two main phases in this research:

Phase 1: an online survey for all current teachers: 30 minutes long

The online survey is attached in this link. The survey will ask questions to examine your perceptions regarding the benefits, drawbacks, and major barriers of using Virtual Reality (VR-based) tools in language teaching, as well as the areas of learning that are expected to benefit from this application. You do not need to have used VR technology in the classroom to take part, as we are interested in the views of all teachers.

At the end of the survey, you will be asked if you want to take part in a focus group session to further discuss your answers in the survey.

Phase 2: A focus group discussion: 30 minutes long

In the discussion, researchers will refer to your answers in the survey and further discuss your selection in a group of 3-5 participants.

Thank you so much for your great contribution to this research!

Appendix B: Online Survey Consent Form

CUREC Approval Reference: ED-CIA-21-156

General Information

This study examines teachers' perceptions of the benefits and drawbacks of Virtual Reality (VR-based) technology, and the main barriers to effective use of VR programs at different schools. With your contribution, the research aims to provide insights for useful pedagogical applications for the future use of VR technology in language education.

We appreciate your interest in participating in this online survey. You have been invited to participate as you are:

1. Currently a **foreign language teacher**
2. Currently **teaching secondary to higher education levels**

The survey will ask questions to examine your perceptions of the benefits, drawbacks and barriers of using Virtual Reality (VR-based) tools in language teaching, as well as the areas of learning that are expected to benefit from this application. You **do not** need to have used VR technology in the classroom to take part, as we are interested in the views of all teachers.

Completing the survey should take about **30 minutes**. No background knowledge is required. The data will be used for the purpose of this study by the principal researcher and authorized people and will be stored on a University of Oxford OneDrive account or a password-protected laptop. The raw data will not be shared with any third parties.

You may ask any questions before deciding to take part by contacting the researcher (details below).

The **Principal Researcher** is **xxxxx** who is attached to the **Department of Education** at the **University of Oxford.** This project is being completed under the supervision of **xxxxx**.

Do I have to take part?

No. Please note that participation is **voluntary**. If you do decide to take part, you may withdraw at any point for any reason before submitting your answers by **pressing the 'Exit' button/ closing the browser.**

How will my data be used?

The data we will collect that could identify you will be:

- Your age
- Your gender
- Years of teaching experience

- Educational taught level
- Countries where you are teaching

You will also have the option to provide your email address if you would like to participate in a future focus group. Your IP address **will not** be stored. We will take all reasonable measures to ensure that data remain confidential.

The responses you provide will be stored in a password-protected electronic file and may be used in academic publications. Research data (including consent records) will be stored for three years after publication or public release.

Who will have access to my data?

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the study. The University will process your personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest. Further information about your rights with respect to your personal data is available from https://compliance.admin.ox.ac.uk/individual-rights.

We would also like your permission to use the data in future studies, and to share data with other researchers (e.g., in online databases). Data will be de-identified before it is shared with other researchers or results are made public.

This survey will be written up for a Master of Science (MSc) degree.

Who has reviewed this study?

This project has been reviewed by, and received ethics clearance through, the Department of Education Research Ethics Committee (Approval Reference: ED-CIA-21-156)

Who do I contact if I have a concern or wish to complain?

If you have a concern about any aspect of this study, please speak to xxxxx via her email address xxxxx or the supervisor xxxxx and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible via **Social Sciences & Humanities Interdivisional Research Ethics Committee:**

Email: ethics@socsci.ox.ac.uk

Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

Please note that you may only participate in this survey if you are 18 years of age or over.

☐ I certify that I am 18 years of age or over

If you have read the information above and agree to participate with the understanding that the data (including any personal data) you submit will be processed accordingly, please check the relevant box below to get started.

☐ Yes, I agree to take part

Appendix C: Focus Group Consent Form

PARTICIPANT CONSENT FORM

CUREC Approval Reference: ED-CIA-21-156

Please read through the **information sheet attached** before agreeing to participate (if you wish to)

1	I confirm that I have read and understand the information sheet for the above study. I have had the opportunity to consider the information, ask questions and have had these answered satisfactorily.	☐
2	I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason, and without any adverse consequences or penalty.	☐
3	I understand that research data collected during the study may be looked at by authorised people outside the research team. I give permission for these individuals to access my data.	☐
4	I understand that this project has been reviewed by, and received ethics clearance through, the Department of Education Research Ethics Committee.	☐
5	I understand who will have access to personal data provided, how the data will be stored and what will happen to the data at the end of the project.	☐
6	I understand how this research will be written up and published.	☐
7	I understand how to raise a concern or make a complaint.	☐
8	I consent to being video recorded	☐
9	I understand how videos will be used in research outputs	☐
10	I agree to the use of anonymised quotes in research outputs	☐
11	I agree to take part in the study	☐
12	I agree that my personal contact details can be retained in a secure database so that the researchers can contact me about future studies.	☐

13 I agree for research data collected in this study to be given to researchers, including those working outside of the UK and the EU, to be used in other research studies. I understand that any data that leave the research group will be anonymised so that I cannot be identified. ☐

Name of Participant ____________________ Date __________________

Appendix D: Participant Information Sheet

PARTICIPANT INFORMATION SHEET

CUREC Approval Reference: ED-CIA-21-156

1. *Why is this research being conducted?*

In recent years, the use of technology in foreign language classroom has become increasingly prevalent. Technology can be used either as a platform for obtaining language input and interaction (from learners' perspectives) or as a tool for knowledge distribution (from teachers' perspectives). Many studies primarily focused on educational technology from students and parents' perspectives, yet little research justified this phenomenon from teachers' points of view. It was demonstrated that teachers encountered a wide range of technology-related issues that directly affected their teaching during the process of technological implementation. Some researchers provided an explanation for the lack of research on teacher's practice, arguing that hidden situations in each classroom setting challenge the validity of empirical research. This study, therefore, involves both quantitative and qualitative research to examine teachers' perceptions regarding the benefits and drawbacks of Virtual Reality (VR) technology, and the main barriers to effective use of VR programs at different schools.

All participants, who are current foreign language teachers, would be asked to take part in an online survey, which include questions about the benefits, drawbacks, barriers and areas of learning of VR-based tools. A subset of participants would then be asked to voluntarily take part in one of several focus-group discussion sessions to further discuss their responses in the survey.

2. *What will happen to me if I take part in the focus group sessions?*

If you are happy to take part in the focus group discussion, you will be interviewed in a group of 3-5 participants via a Teams call. Each session will last 30 minutes, and the researcher will ask questions in relation to your answers in the survey. You will then be asked about your opinions of the factors that might impact the effective implementation of VR-based tools in language classrooms. During the session, if a participant were to become upset, they would be reassured and allowed to exit the meeting. With your consent, we would like to video record you so we can have an accurate record of your thoughts.

3. *Why have I been invited to take part?*

You have been invited because you are:

1. A current foreign language teacher
2. Currently teaching secondary and/or higher education levels

4. *Do I have to take part?*

No. The participation is voluntary and you can ask questions before deciding to take part or not.

5. *Are there any potential risks in taking part?*

One risk that might occur when taking part in this study is breaching of confidential identifiable data. To reduce this potential risk, participants' names will not be collected and all data will be stored in a password-protected electronic file under University of Oxford OneDrive account.

6. *Are there any benefits in taking part?*

There will be no direct or personal benefit to you from taking part in this research.

7. *What happens to the data provided?*

The information you provide during the study is the research data. Any research data from which you can be identified (age/gender/years of experience/taught level/countries where the participant is teaching at) is known as personal data.

Personal data will be stored in a University of Oxford OneDrive account or a password-protected laptop. The data will be retained for 3 years after publication. Other research data (including consent forms) will also be stored for at least 3 years after publication or public release of the work of the research.

The principal researcher and the supervisor will have access to the research data. Responsible members of the University of Oxford may be given access to data for monitoring and/or audit of the research.

We would like your permission to use direct quotes in any research outputs.

We would like your permission to use anonymised data in future studies, and to share data with other researchers (e.g., in online databases).

8. *Will the research be published?*

The research will be written up as a Master's dissertation. On successful submission of the dissertation, it may be deposited both in print and online in the University archives to facilitate its use in future research. If so, the thesis will be openly accessible.

9. *Who has reviewed this study?*

This study has been reviewed by, and received ethics clearance through, the Department of Education Research Ethics Committee (Approval Reference: ED-CIA-21-156)

10. *Who do I contact if I have a concern about the study or wish to complain?*

If you have a concern about any aspect of this study, please speak to xxxxx via her email address: xxxxx or the supervisor: xxxxx and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible via Social Sciences & Humanities Interdivisional Research Ethics Committee:
Email: ethics@socsci.ox.ac.uk
Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

11. *Data Protection*

The University of Oxford is the data controller with respect to your personal data, and as such will determine how your personal data is used in the study.
The University will process your personal data for the purpose of the research outlined above. Research is a task that is performed in the public interest.
You will have the rights under General Data Protection Regulation (GDPR)
Further information about your rights with respect to your personal data is available from http://www.admin.ox.ac.uk/councilsec/compliance/gdpr/individualrights/.

12. *Further Information and Contact Details*

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

xxxxxx
Department of Education
15 Norham Gardens, Oxford OX2 6PY
University email: xxxxxx

Appendix E: CUREC Approval

30 March 2021 10:10

Dear xxxxx and xxxxx,

The Use of Virtual Reality (VR) based Technology in Foreign Language Education: From teachers' perspectives - approval reference ED-CIA-21-156

The above application has been considered on behalf of the Departmental Research Ethics Committee (DREC) in accordance with the procedures laid down by the University for ethical approval of all research involving human participants. I am pleased to inform you that, on the basis of the information provided to DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly, approval has been granted.

Please continue to follow all current guidance issued by CUREC during the pandemic, notably COVID-19: CUREC guidance on research involving human participants, and on internet-based research: https://researchsupport.admin.ox.ac.uk/governance/ethics/coronavirus

If relevant, please also check the CUREC website for their best practice research guides: https://researchsupport.admin.ox.ac.uk/governance/ethics/resources/bpg

Should there be any subsequent changes to the project which raise ethical issues not covered in the original application you should submit details to research.office@education.ox.ac.uk for consideration.

Good luck with your research study.

With kind regards,
Dr. Sandra Mathers
Senior Researcher
Department of Education
University of Oxford

Appendix F: Test of Instrument Reliability: Cronbach's Alpha

The Cronbach's alpha test was used to measure the reliability of the scales of each outcome measure in the questionnaire: (1) Areas of learning supported by VR technology; (2) VR benefits, (3) VR drawbacks; (4) implementation challenges; and (5) implementation enabling factors. The test results are shown below.

Table F1: Reliability statistics for all outcome measure

Outcome Measure	Cronbach's Alpha	Cronbach's Alpha Based on Standardized Items	Number of Items
Areas of Learning	.858	.859	6
Benefits	.931	.932	14
Drawbacks	.795	.797	7
Implementation Challenges	.666	.666	8
Implementation Enabling Factors	.873	.878	13

As seen in the table, the scales used to measure outcome measure *areas of learning*, *VR benefits*, and *implementation enabling factors* achieved high internal consistency, where Cronbach's $\alpha > .8$. On the other hand, the scales used to measure outcome measure *VR drawbacks* and *implementation challenges* achieved comparatively lower reliability scores, Cronbach's $\alpha = .795$ and .666 respectively (< .8).

Appendix G: Online Survey Questionnaire

Online survey can be found here: link

Virtual Reality (VR-based) programs in language learning refer to the use of VR tools, which allow users to immerse in real-life scenarios to practice their language production. VR tools include a wide range of modern technologies, including computers, phones, tablets, headsets, 3D glasses to simulate the real environments, thus potentially bringing learners the most authentic experiences.

*ImmerseMe is a renowned VR learning program that offers students **engagement with native speakers** in real-life scenarios without real-time stress. For each lesson, students have opportunities to develop a conversation with native speakers according to the lesson knowledge and the levels of difficulty. There are four modes of practice: **Pronunciation – Writing – Translation – Immersion**, allowing learners to practice repeatedly to improve certain areas of learning. With the **voice recognition feature**, learners can only proceed the conversation by producing correct pronunciation. At the beginning and pre-intermediate levels, suggested utterances are provided (**watch a short video of ImmerseMe**). At higher levels of the Immersion mode, learners need to come up with their own answers using given word cues. At the moment, ImmerseMe offers VR-based content for **9 languages**.

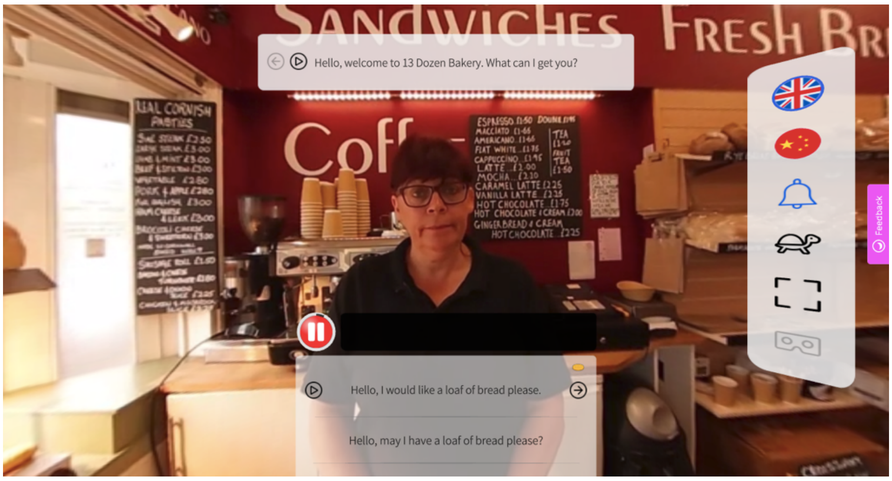


Figure 1: Screenshot of ImmerseMe content at beginner level

Questionnaire

	Categories	Measure	Question
1	**PI: Age**	Insert number (>18)	How old are you?
2	**PI: Gender**	Multiple Choice Question (MCQ)	Are you: Male, Female or Other?
3	**PI: Years of experience**	Insert number	How many years of experience do you have for language teaching overall?
4	**PI: Taught Level**	MCQ	What levels are you currently teaching?
5	**PI: Countries**	MCQ	In which country are you currently residing?
6	**PI: Experience with Technology**	**Likert Scale** 1 – Not at all useful 2 – Slightly useful 3 – Moderately useful 4 – Really useful 5 – Extremely useful	How useful are computer-based programs to your language teaching? ***Digital technologies** are electronic tools, systems, devices, and resources that generate, store, or process data. Examples include social media, online games, multimedia, and mobile phones.
7	**PI: Technology Use Frequency**	**Likert Scale** 1 – Never 2 – Occasionally (once/two weeks) 3 – Often (once/week) 4 – Frequently (more than 3 times/week) 5 – Everyday	How often do you use computer-based programs for language teaching?
8		**Yes/No**	Have you ever used a Virtual Reality (VR-based) program such as ImmerseMe?
9	**Experience with VR**	**Likert Scale** 1 – Not at all useful 2 – Slightly useful 3 – Moderately useful 4 – Very useful 5 – Extremely useful	How useful are Virtual Reality programs (e.g., ImmerseMe) for language teaching?
	VR Benefits	**Likert Scale** 1 – Strongly agree 2 – Agree 3 – Neutral	

		4 – Disagree 5 – Strongly disagree	
10-13	Set 1: In terms of learning approaches		- It promotes task-based[1] learning which helps students with problem-solving skills - It facilitates experiential and contextualised[2] learning - It helps students meet different needs in language learning - It empowers autonomous and active learning
14-17	Set 2: In terms of pedagogy and teaching approaches		- The VR-supported pedagogical approach allows the integration of different language activities - It allows teachers to implement constructivist* teaching approaches - It allows learner-centred mode of instruction rather than conventional teacher - centred orientation. - It promotes rich and natural interaction environments for language output by providing exposure to authentic culture
18-23	Set 3: In terms of linguistic acquisition		- It improves learners' speaking skills - It helps students improve general language skills and knowledge - It gives learners opportunities to improve their language confidence - It generally improves students' language performance and grades - It helps learners with self-correcting skills - It helps students with long-term memory retention of learnt knowledge
24		Optional short paragraph	Are there any other benefits of VR-based programs that you want to share with us?
	VR Drawbacks	**Likert Scale** 1 – Strongly agree 2 – Agree 3 – Neutral 4 – Disagree 5 – Strongly disagree	

25-26	Set 1: In terms of educational value		- The gamification and competition features create more recreational rather than educational appeal to learners - It distracts learners by the visual appeal and interactivity of the stimulated environments
27-28	Set 2: In terms of learning autonomy		- Students over-rely on technology and avoid face-to-face conversations - Students tend to follow the script from the scenarios and struggle to create sentences themselves
29-30	Set 3: In terms of lesson planning		- It is time-consuming for lesson design and setup - There is insufficient practice time when integrating VR tools
31	Set 4: In terms of classroom environments		- It is harder to control the class while students are using the VR-based programs
32		Optional short paragraph	Are there any other drawbacks of VR-based programs that you want to share with us?
	Implementation Challenges	**Likert Scale** 1 – Strongly agree 2 – Agree 3 – Neutral 4 – Disagree 5 – Strongly disagree	
33-36	Set 1: In terms of technical-related issues		- I believe I cannot cope with technical issues related to VR-based tools - Technical failures frustrate students, resulting in ineffective teaching and learning - There are not enough facilities (well-functioning computers/hardware) to run VR-based programs in class - Before this survey, I never knew of any good examples of using VR-based in language learning
37-38	Set 2: In terms of lesson design		- I do not know how to integrate a VR program into my curriculum

			- The scenarios might not match my lesson design and so I need to re-plan my lessons
39-40	Set 3: In terms of teacher training		- There are not enough training materials for the implementation of VR-based tools - There is a lack of teacher training in effective use of VR-based tools
41		Optional short paragraph	Are there any other barriers to effective use of VR-based programs that you want to share with us?
	Areas of Learning Supported by VR Technology	**Likert Scale** 1 – Strongly agree 2 – Agree 3 – Neutral 4 – Disagree 5 – Strongly disagree	
42-44	Set 1: Attention to linguistic forms (grammatical and lexical related)		- Students produce more complex and grammatically accurate language in task performance - Students develop self-correcting grammatical and lexical skills - Students use vocabulary more flexibly in different contexts
45-46	Set 2: Productive skills		- Students' written performance is better after using the VR-based program - Their speaking abilities get better after practices
47	Set 3: Receptive skills		- Students' listening skill is improved significantly
48		Optional short paragraph	Are there any other areas of learning that might be benefited from the use of VR-based programs that you want to share with us?
	Factors to Effective Use of VR	**Likert Scale** 1 – Not at all important 2 – Slightly important 3 – Moderately important 4 – Very important 5 – Extremely important	

49-51	Set 1: In terms of technology		- Teacher and student's digital literacy[3] - Teacher and student's comfort with the use of Virtual Reality - Good functioning facilities (computers and hardware)
52-56	Set 2: In terms of curriculum		- Adaptive curricula - Flexible lesson structure - Long practice time - Consistent practice - Class management
57-61	Set 3: In terms of student's experience		- Student's level of engagement - Student's learning autonomy - Student's concentration - Student's strong theoretical understanding of grammar - Student's pronunciation of learnt vocabulary
62		Optional short paragraph	Are there any other factors which are important to the effective use of VR-based tools in foreign language classroom that you want to share with us?

[1]: Task-based learning refers to the ways in which students solve a task that involves authentic use of language, rather than completing simple language questions about grammar or vocabulary.

[2]: Contextualised learning focuses on the practicality of knowledge, referring to how students engage in active learning to make meaning out of information they are obtaining

[3]: Digital literacy refers to individual's ability to use digital-related platforms and programs

Appendix H: Descriptive Statistics of Questionnaire Items

Outcome Measure	Sub Outcome Measure	Item	*M*	*SD*	Skewness	Kurtosis
Areas of Learning Supported by VR Technology	**Linguistic forms**	42 (more complex and grammatically accurate language in task performances)	3.52	.819	-.168	.74
		43 (develop self-correcting grammatical and lexical skills)	3.53	.733	-.026	-.242
		44 (use vocabulary more flexibly)	3.64	.807	-.251	.243
		45 (better written performances)	3.32	.834	-.164	.707

	Productive skills	46 (better speaking abilities)	4.04	.675	-.234	-.140
	Receptive skills	47 (improve listening skill)	4.03	.683	-.364	-.713
Benefits of VR Technology	**Learning approaches**	10 (promotes task-based learning and problem-solving)	4.06	.715	-.56	.54
		11 (facilitates experiential and contextualised learning)	4.27	.578	-.098	-.484
		12 (meet students' different needs)	4.00	.756	-.539	.248
		13 (empower autonomous annd active learning)	4.18	.644	-.401	.329
	Teaching approaches	14 (VR-supported pedagogical approach allows the integration of different language activities)	4.08	.700	-.445	.236
		15 (allow teachers to implement constructivist teaching approaches)	4.01	.683	-.377	.314
		16 (allow learner-centred mode of instruction)	4.12	.700	-.343	-.299
		17 (promotes rich and natural interaction environments for language output by providing exposure to authentic culture)	4.18	.659	-.411	.164
	Language Acquisition	18 (improve speaking skills)	4.02	.743	-.456	.067
		19 (improve general language skills and knowledge)	4.05	.681	-.427	.418
		20 (improve language confidence)	4.15	.659	-.170	-.694
		21 (improve language performance and grades)	3.79	.836	-.088	-.234
		22 (help self-correcting skills)	3.84	.732	-.334	.058
		23 (help long-term memory retention of learnt knowledge	3.84	.678	.021	-.421
Drawbacks of VR Technology	**Educational value**	25 (the gamification and competition features create more recreational rather than educational appeal to learners)	2.77	1.017	.471	-.310
		26 (distracting by the visual appeal)	2.46	.938	.499	-.464
	Learning Autonomy	27 (students over-rely on technology and avoid face-to-face conversations)	2.75	1.113	.395	-.651
		28 (students follow script and struggle to create sentences themselves)	3.20	.909	-.019	-.795

	Lesson Planning	29 (time-consuming for lesson design and setup)	3.32	1.065	-.387	.477
		30 (insufficient practice)	2.99	.951	.154	-.729
	Classroom environment	31 (uncontrollable class)	2.76	.991	.492	-.318
Implementation Challenges	**Technical-related issues**	33 (unable to cope technical issues)	2.57	.966	.326	.496
		34 (frustrate students and ineffective teaching and learning)	3.31	1.008	-.207	-.656
		35 (insufficient facilities)	3.49	1.044	-.179	-.969
		36 (no access to VR tools before)	3.50	1.062	-.389	-.644
	Lesson Design	37 (unable to integrate VR into curriculum)	2.58	1.013	.465	-.348
		38 (replanning the lesson design)	3.08	1.066	-.009	-.777
	Teacher Training	39 (lack of training materials)	3.08	1.021	-.099	-.844
		40 (lack of teacher training)	3.42	1.004	-.344	-.713
Implementation Enabling Factors	**Technology**	49 (teacher and student's digital literacy)	4.17	.765	-.697	-.208
		50 (teacher and student's comfort with the use of VR)	3.98	.759	-.371	-.189
		51 (well-functioning facilities)	4.52	.695	-1.490	2.127
	Curriculum	52 (adaptive curricula)	3.96	.784	-.177	-.761
		53 (flexible lesson structure)	4.01	.803	-.359	-.543
		54 (long practice time)	3.64	.856	.121	-.748
		55 (consistent practice)	4.00	.797	-.233	-.812
		56 (class management)	3.90	.894	-.799	-.713
	Student's contribution	57 (student's level of engagement)	4.26	.721	-1.216	3.355
		58 (student's learning autonomy)	4.18	.782	-.946	1.596
		59 (student's concentration)	4.20	.642	-.428	.386
		60 (student's strong theoretical understanding of grammar)	3.14	.975	.025	-.073
		61 (student's pronunciation of learnt vocabulary)	3.71	.895	-.546	.025

Appendix I: Paired T-Test Results for Outcome Measure Areas of Learning (Table 2)

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q42 – Q43	-.009	.640	.062	-.133	.114	-.152	105	.880
Q42 – Q44	-.123	.752	.073	-.268	.022	-1.679	105	.096
Q42 – Q45	.198	.821	.080	.040	.356	2.483	105	.015
Q42 – Q46	-.519	.875	.085	-.067	-.035	-6.103	105	.000*
Q42 – Q47	-.509	.784	.076	-.660	-.359	-6.692	105	.000*
Q43 – Q44	-.113	.772	.075	-.262	.036	-1.509	105	.134
Q43 – Q45	.208	.700	.068	.073	.342	3.053	105	.000*
Q43 – Q46	-.509	.796	.077	.073	-.356	-6.591	105	.000*
Q43 – Q47	-.500	.694	.067	-.634	-.366	-7.423	105	.000*
Q44 – Q45	.321	.763	.074	.174	.468	4.329	105	.000*
Q44 – Q46	- .396	.801	.078	-.055	-.242	-5.093	105	.000*
Q44 – Q47	- .387	.738	.072	-.529	-.245	-5.398	105	.000*
Q45 – Q46	-.717	-848	.082	-.880	-.554	-8.705	105	.000*
Q45 – Q47	-.708	.084	.078	-.862	-.553	-9.056	105	.000*
Q46 – Q47	.009	.543	.053	-.095	.114	.179	105	.858

Appendix J: Paired T-Test Results for Each Subscale of Benefits of VR (Table 4)

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q10 – Q11	-.217	.647	.063	-.342	-.092	-3.451	105	.001*
Q10 – Q12	.057	.701	.068	-.078	.192	.831	105	.408
Q10 – Q13	-.123	.658	.064	-.249	.004	-1.920	105	.058
Q11 – Q12	.274	.655	.064	.147	.400	4.299	105	.000*
Q11 – Q13	.094	.594	.058	-.020	.209	1.635	105	.105
Q12 – Q13	-.179	.701	.058	-.314	-.044	-2.633	105	.010
Q14 – Q15	.066	.590	.057	-.048	.180	1.153	105	.252
Q14 – Q16	-.047	.709	.069	-.184	.089	-.685	105	.495
Q14 – Q17	-.104	.631	.061	-.225	.018	-1.692	105	.094
Q15 – Q16	-.113	.558	.054	-.221	-.006	-2.090	105	.039
Q15 – Q17	-.170	.683	.066	-.301	-.038	-2.561	105	.012
Q16 – Q17	-.057	.674	.065	-.186	.073	-.865	105	.389
Q18 – Q19	-.28	.654	.064	-.154	.098	-.446	105	.657
Q18 – Q20	-.132	.618	.060	-.251	-.013	-2.199	105	.030
Q18 – Q21	.226	.637	.062	.104	.349	3.661	105	.000*
Q18 – Q22	.179	.659	.064	.052	.306	2.801	105	.006
Q18 – Q23	.179	.714	.069	.042	.317	2.583	105	.011
Q19 – Q20	-.104	.551	.054	-.210	.002	-1.940	105	.055
Q19 – Q21	.255	.691	.067	.122	.388	3.796	105	.000*
Q19 – Q22	.208	.672	.065	.078	.337	3.179	105	.002*
Q19 – Q23	.208	.658	.064	.081	.334	3.248	105	.002*
Q20 – Q21	.358	.733	.071	.217	.500	5.037	105	.000*
Q20 – Q22	.311	.667	.065	.183	.440	4.805	105	.000*
Q20 – Q23	.311	.638	.062	.188	.434	5.025	105	.000*
Q21 – Q22	-.047	.681	.066	-.178	.084	-.713	105	.478
Q21 – Q23	-.047	.709	.069	-.184	.089	-.685	105	.495
Q22 – Q23	.000	.756	.073	-.146	.146	.000	105	1.000

Appendix K: Paired T-Test Results for Outcome Measure VR Drawbacks (Table 8)

Item Pair	**Mean**	**Std. Deviation**	**Std. Error Mean Difference**	**95% Confidence Interval**		***t***	***df***	***p***
				Lower	Upper			
Q25 – Q26	.311	1.045	.102	.110	.513	3.066	105	.003
Q25 – Q27	.028	1.142	.111	-.192	.248	.255	105	.799
Q25 – Q28	-.425	1.242	.121	-.664	-.185	-3.520	105	.001*
Q25 – Q29	-.547	1.374	.133	-.812	-.283	-4.100	105	.000*
Q25 – Q30	-.217	1.272	.124	-.462	.028	-1.756	105	.082
Q25 – Q31	.009	1.207	.117	-.223	.242	.080	105	.936
Q26 – Q27	-.283	1.012	.098	-.478	-.088	-2.880	105	.005
Q26 – Q28	-.736	.989	.096	-.926	-.545	-7.664	105	.000*
Q26 – Q29	-.858	1.099	.107	-1.070	-.647	-8.041	105	.000*
Q26 – Q30	-.528	1.053	.102	-.731	-.326	-5.166	105	.000*
Q26 – Q31	-.302	.968	.094	-.488	-.116	-3.212	105	.002*
Q27 – Q28	-.453	1.243	.121	-.692	-.213	-3.750	105	.000*
Q27 – Q29	-.575	1.265	.123	-.819	-.332	-4.685	105	.000*
Q27 – Q30	-.245	1.271	.123	-.490	-.001	-1.987	105	.050
Q27 – Q31	-.019	1.138	.111	-.238	.200	-.171	105	.865
Q28 – Q29	-.123	1.084	.105	-.331	.086	-1.165	105	.247
Q28 – Q30	.208	1.039	.101	.007	.408	2.056	105	.042
Q28 – Q31	.434	1.163	.113	.210	.658	3.841	105	.000*
Q29 – Q30	.330	.983	.095	.141	.519	3.459	105	.001*
Q29 – Q31	.557	1.079	.105	.349	.764	5.313	105	.000*
Q30 – Q31	.226	1.035	.101	.027	.426	2.251	105	.026

Appendix L: Paired T-Test Results for VR Implementation Challenges (Table 10)

Item Pair	Mean	Std. Deviation	Std. Error Mean Difference	95% Confidence Interval		*t*	*df*	*p*
				Lower	Upper			
Q33 – Q34	-.745	1.096	.106	-.956	-.534	-7.001	105	.000*
Q33 – Q35	-.925	1.165	.142	-1.027	-.642	-6.497	105	.000*
Q33 – Q36	-.934	1.354	.132	-.1.195	-.673	-7.101	105	.000*
Q33 – Q37	-.019	1.219	.118	-.254	.216	-.159	105	.874
Q33 – Q38	-.509	1.340	.130	-.767	-.251	-3.915	105	.000*
Q33 – Q39	-.509	1.396	.136	-.778	-.241	-3.758	105	.000*
Q33 – Q40	-.858	1.390	.135	-1.126	-.591	-6.359	105	.000*
Q34 – Q35	-.179	1.440	.140	-.456	.098	-1.282	105	.203
Q34 – Q36	-.189	1.346	.131	-.448	.071	-1.443	105	.152
Q34 – Q37	.726	1.444	.140	.448	1.005	5.178	105	.000*
Q34 – Q38	.236	1.313	.128	-.017	.489	1.849	105	.067
Q34 – Q39	.236	1.231	.120	-.001	.473	1.973	105	.051
Q34 – Q40	-.113	1.245	.121	-.353	.126	-.936	105	.351
Q35 – Q36	-.009	1.483	.144	-.295	.276	-.065	105	.948
Q35 – Q37	.906	1.299	.126	.656	1.156	7.181	105	.000*
Q35 – Q38	.415	1.241	.121	.176	.654	3.443	105	.001*
Q35 – Q39	.415	.994	.097	.224	.607	4.300	105	.000*
Q35 – Q40	.066	.998	.097	-.126	.258	.681	105	.497
Q36 – Q37	.915	1.448	.141	.636	1.194	6.505	105	.000*
Q36 – Q38	.425	1.414	.137	.152	.697	3.091	105	.003
Q36 – Q39	.425	1.447	.141	.146	.703	3.020	105	.003
Q36 – Q40	.075	1.478	.144	-.209	.360	.526	105	.600
Q37 – Q38	-.491	1.189	.116	-.720	-.262	-4.247	105	.000*
Q37 – Q39	-.491	1.221	.119	-.726	-.255	-4.137	105	.000*
Q37 – Q40	-.840	1.266	.123	-1.083	-.596	-6.828	105	.000*
Q38– Q39	.000	1.155	.112	-.222	.222	.000	105	1.000
Q38 – Q40	-.349	1.219	.118	-.584	-.114	-2.948	105	.004
Q39 – Q40	-.349	-.873	.085	-.517	-.181	-4.115	105	.000*

Appendix M: Paired T-Test Results for Implementation Enabling Factors (Table 13)

Item Pair	Mean	Std. Deviation	Std. Error Mean Difference	95% Confidence Interval		*t*	*df*	*p*
				Lower	Upper			
Q60 – Q54	-.495	1.066	.104	-.702	-.289	-4.759	105	.000*
Q60 – Q61	-.571	1.082	.106	-.781	-.362	-5.413	105	.000*
Q60 – Q56	-.762	1.140	.111	-.982	-.541	-6.851	105	.000*
Q60 – Q52	-.819	1.054	.103	-1.023	-.615	-7.962	105	.000*
Q54 – Q52	-.324	.791	.077	-.477	-.171	-4.197	105	.000*
Q61 – Q49	-.457	1.047	.102	-.660	-.254	-4.473	105	.000*
Q56 – Q51	-.619	1.023	.100	-.817	-.421	-6.203	105	.000*
Q52 – Q51	-.562	.876	.086	-.731	-.392	-6.571	105	.000*
Q50 – Q51	-.543	.797	.078	-.697	-.389	-6.980	105	.000*
Q55 – Q51	-.524	.962	.094	-.710	-.338	-5.581	105	.000*
Q53 – Q51	-.514	.856	.084	-.680	-.349	-6.155	105	.000*
Q49 – Q51	-.352	.784	.077	-.504	-.201	-4.604	105	.001*
Q58 – Q51	-.343	.853	.083	-.508	-.178	-4.119	105	.000*
Q59 – Q51	-.324	.753	.073	-.470	-.178	-4.406	105	.000*
Q57 – Q51	-.267	.737	.072	-.409	-.124	-3.706	105	.000*

Appendix N: Focus Group Discussion Procedure and Questions

<u>Before the Discussion</u>

1. **Link** to join the meeting should be sent 3 days in advance to the participants.
2. Participants are required to confirm whether they have received the email
3. **Timeline** of the discussion will be sent out to all participants

Timeline (British Summer Time)	Main Activity
09:00 – 09:01	*Welcome: Introduction to the session (~30s)*
09:01 – 09:02	*The purpose of the discussion (~1 min)*
09:02 – 09:04	*Ground rules to discussion (~2 mins)*
09:04 – 09:29	*Key Questions (~25 mins)*
09:29 – 09:30	*Ending the discussion (~1 min)*

4. Researcher **pilots** the meeting to confirm: Teams Recording/Transcription functions
5. The moderator joins the meeting 15 minutes in advance, prepare:
    - Notepads/pens/pencils
    - Timer (clock)
    - Focus Group Procedure & Questions

<u>During the Discussion</u>

1. *Welcome: Introduction to the session (~30s)*

Good morning and welcome to our focus discussion group session. Thank you for taking the time to join us. My name is xxxxx and I am currently a Master student in Applied Linguistics at the University of Oxford.

2. *The purpose of the discussion (~1 min)*

The purpose of this group discussion is to discuss your opinions and experiences about the use of Virtual Reality (VR) technology such as ImmerseMe in foreign language classrooms. You were selected because you are a foreign teacher who might have or have not used ImmerseMe as a part of your teaching. As we've mentioned, we'd like to video-record the session because we don't want to miss any of your comments. We are on a first name basis, and we won't use any names in our reports. You may be assured of complete confidentiality.

3. *Ground rules to discussion (~2 mins)*

Before we start, I would like to briefly give you some guidelines before we start. Please turn on your camera all time and only turn on your mic when you speak. It is important that you feel comfortable during the discussion so please let me know if you have ANY issue at any point. If you experience a technical issue and would have to leave, please let me know. My role as moderator would be to ONLY guide the discussion, which means that I will not take part in the discussion. You all are free to comment and respond upon each other's opinions and please listen respectfully when others share their views. There are no right or wrong answers, only different points of view. Please feel free to share your point of view even if it differs from what others have said.

4. Key Questions (~25 mins)

Well, let's begin!

Question 1: Some of you agreed that VR-based programs benefit [subscale] (e.g., teaching approaches). Can each of you give a real-life example? For those who have never used VR-based tools, can you imagine how the program might benefit in terms of [teaching approaches] in real-life classroom settings? *(~5 mins)*

Thank you for your discussion. I would like to move on to the next question.

Question 2: Some of you concurred that VR-based programs have detrimental effects to the process of teaching in terms of [subscale] (e.g., lesson planning). Can you elaborate on your answer? For those who never used VR-based tools, how do you imagine the program might be detrimental to your [subscale]? *(~5 mins)*

Thank you for your discussion. I would like to move on to the next question.

Question 3: Apart from the implementation challenges mentioned in the survey (e.g., unpredictable IT-related issues, complicated lesson design, lack of teacher training), what other barriers do you find, or are expected to find when using VR tools in your classroom? *(~5 mins)*

Thank you for your discussion. I would like to move on to the next question.

Question 4: In your opinion, which areas of learning (e.g., speaking, listening, speaking, writing) are supported using VR technology? Please elaborate your answer. *(~5 mins)*

Question 5: Can you name three most important factors that have affected or will be affecting the effective use of VR-based tools in foreign language classes? (*~5 mins*)

5. Ending the discussion (~1 min)

This would be the end of the discussion. Please let me know if you have any question about the discussion today, or how to contact us about your confidentiality. Thank you so much for your contribution to our research. If you are happy, I would like to end the discussion now! Thank you.