A Mixed-method Study of the Impact of High-immersion Virtual Reality on Foreign Language Speaking Anxiety

Miaomiao Ding

A Thesis Submitted in Fulfilment of the Requirements for the Degree of Doctor of Philosophy (PhD)

School of Education
College of Social Science
University of Glasgow
June 2023

Abstract

Previous research shows that anxiety about speaking foreign languages is common among foreign language learners. Foreign Language Speaking Anxiety (FLSA) has been studied extensively, but insights into approaches to alleviate it remain limited. Some studies found that immersion in the target-language speaking environment may help to alleviate FLSA, but not every foreign language learner can visit a foreign country or live there for an extended period because of the cost in money and time. High-immersion Virtual Reality (HiVR) appears to offer an alternative to this. Wearing a head-mounted display allows a first-person perspective in a virtual environment while using HiVR. In addition, HiVR provides direct interaction with the environment through gloves or controllers. These features of HiVR technology seem to contribute to a great sense of presence in the virtual environment, which means users may feel as though they are in the real world, and to a high degree of agency in the control of the learning process.

The investigation of HiVR in foreign language learning has increased in the last five years, primarily due to decreased cost and technological development of HiVR. However, most studies are on the effects of HiVR on acquiring linguistic knowledge, such as words and expressions, and the impact of HiVR on FLSA remains underexplored. In the field of psychology, HiVR is often used to treat psychological disorders, such as social anxiety disorders, fear of flying and fear of spiders. Therefore, in my study, it was assumed that HiVR may be used to help foreign language learners cope with FLSA.

Through a mixed-method experimental design, my study quantitatively examines the change in FLSA levels among four groups of Chinese English as a Foreign Language (EFL) learners (140 students in total, 35 students in each group) before and after an intervention. A nine-session intervention was conducted for each group at a Chinese university, and the learning sessions were designed with a different combination of two learning environments (HiVR or classroom) and two learning approaches (situated learning or teacher-centred learning). The FLSA levels were measured from two perspectives: general FLSA via a Likert scale and specific FLSA via a self-rating of anxiety levels during a two-way role-play speaking test. The statistical results indicated that students' general FLSA did not change after the intervention in each group, but FLSA levels in the role-play speaking tests decreased significantly in each group. According to students' responses in the open-ended questionnaires, the main reason for the decrease in FLSA levels in the speaking tests appears

to be the fact that they were given more practice in role-play tasks during the intervention rather than the use of HiVR *per se*.

According to the quantitative results, neither HiVR nor situated learning resulted in statistically significant decreases in FLSA levels compared to traditional classroom or teacher-centred learning. Even so, the qualitative results showed that most participants felt positive about HiVR and its potential to alleviate FLSA. The positive perceptions of using HiVR to cope with FLSA included that HiVR provided an authentic and low-anxiety environment, increased learning engagement, and offered more language practice opportunities. Similar opinions were noted regarding the advantages of HiVR in oral English learning, although some disadvantages related to hardware and learning efficiency were reported. According to the qualitative analysis in my study, task repetition and pleasant experiences played significant roles in alleviating FLSA among Chinese EFL learners. When using HiVR to address FLSA, it is worth considering the learning approaches underlying the learning design in HiVR as well as the use of the main learning affordances of HiVR (such as presence and agency) to enhance positive emotions and engagement in the learning process.

Table of Contents

List of Figures

List of Tables

Acknowledgement

My appreciation goes out to many people who helped me and supported me throughout this research.

In the first place, I would like to extend my gratitude to my family. My late grandma, who loved me more than anything in the world, understood me, listened to me, and encouraged me to complete my PhD. My mother's remote accompaniment and my father's financial support over the past four and a half years are also much appreciated. I could not have made it this far without them.

It is my good fortune to be supervised by Professor Ellen Boeren, Dr Gabriella Rodolico, and Dr Ide Haghi. Without their excellent supervision, guidance, care, and support, I would not be able to cope with the disruption to my PhD programme in the first year. In addition, I would like to extend my great appreciation to my principal supervisor, Professor Ellen Boeren, for her professional guidance, encouragement, and assistance throughout the entire process. Besides being my academic mentor, she is also my role model as a human being.

I also would like to thank Linyi Jiang for providing technical support for my research. It is impossible for me to realise all my learning designs in Virtual Reality without him. Furthermore, I would like to extend my sincere gratitude to all 140 participants in my study, as well as to Mr. Zhang, the head of the Foreign Language College at the site, where I conducted my fieldwork.

As well, I would like to express my thanks to all my close friends back home in China, in the UK, in the US, and in Spain. Thanks for supporting me mentally throughout the journey. My heart is filled with memories of every kind word received from my close friends.

Author's Declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Printed Name: Miaomiao Ding

Signature:

Chapter 1. INTRODUCTION

> "This is the OASIS. It is a place where the limits of reality are your own imagination. You can do anything, go anywhere. Like the vacation planet. Surf a 50-foot monster wave in Hawaii. You can ski down the pyramids. You can climb Mount Everest with Batman. Check out this place".
>
> --Wade Watts talking about the OASIS (Ontologically Anthropocentric Sensory Immersive Simulation) in the film of Ready Player One (Spielberg, 2018).

1.1 Research Background

High-immersion Virtual Reality (HiVR) is a technology that immerses users in a virtual environment through a head-mounted display device, and interaction with the environment is achieved by using controllers or gloves with sensors (Peixoto *et al.*, 2021). HiVR provides users with a sense of presence, which is a feeling of being in the virtual environment, and agency, the freedom to control the environment (Petersen, Petkakis, & Makransky, 2022). Using HiVR technology as a tool to stimulate learning has gained the attention of scholars because of its decreased cost and technological development. However, there is a lack of research on HiVR in the field of second/foreign language learning (Peixoto *et al.*, 2019), especially the affective and emotional dimensions of second/foreign language learning.

In the second language acquisition field, language affect is "the feelings or emotional reactions about the language, about the people who speak that language, about the experience of learning and/or being taught a language, or about the culture where that language is spoken" (Gass, Behney, & Plonsky, 2020, p. 530). Affect and emotion in second language acquisition were not at the forefront of discussions in the last century; the primary focus was on the development of knowledge and use of the new language (Garrett & Young, 2009). However, MacIntyre (2002, p. 63) pointed out that "a better understanding of emotion has the capacity to explain cases where students endorse orientations but might not be energised to take action, and also cases where action is prevented by emotional arousal, either present or anticipated". Thus, understanding affect in second language acquisition may help with understanding the process of language learning and use.

The study of affect in language learning can be traced back to the1980s, when Krashen (1982) proposed the Affective Filter Hypothesis. The Affective Filter Hypothesis (Krashen, 1982) claims that negative internal feelings or emotions would hinder language acquisition,

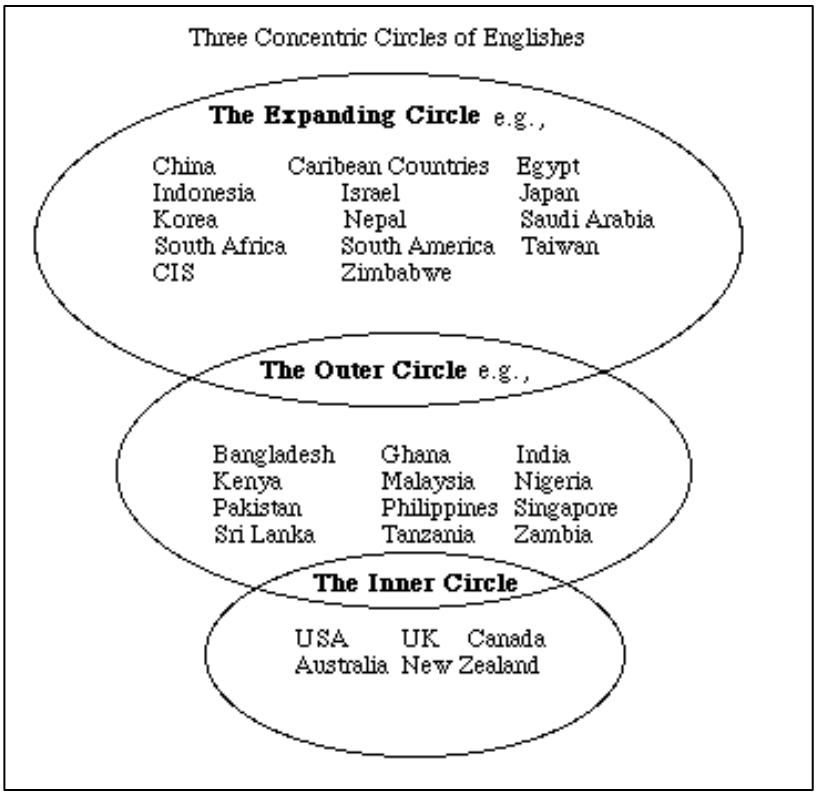
because they create filter that impede the comprehensible input pass to the language acquisition device. Krashen (1985) emphasised the importance of avoiding negative affective reactions and finding a way to establish a positive climate for language learning. This mindset underpins the research of this thesis. Positive language affect that can contribute to successful second language acquisition, from Krashen's perspective, includes "high motivation, strong self-confidence, and low anxiety" (Benesch, 2013, p. 6). Indeed, anxiety is one of the most common emotions felt by language learners, and it is often deemed a negative emotion that impairs language learning (Gass, Behney, & Plonsky, 2020).

Language anxiety is a multifaceted emotion that many foreign language learners experience. It can be influenced by "internal physiological processes, cognitive and emotional states along with the demands of situation and the presence of other people, among other things, considered over different timescales" (MacIntyre, 2017, p. 53). Evidence showed that levels of language anxiety among Asian learners were different from learners in the rest of the world, and Chinese EFL (English as a Foreign Language) learners were more anxious than EFL learners from other countries (Dewaele & MacIntyre, 2014). This may be attributed to the exam-oriented Chinese education system, which requires students to follow models and standards and has low tolerance for mistakes (Jiang & Dewaele, 2019; Kirkpatrick & Zhichang, 2002).

According to the Model of World Englishes[1] (Kachru, 1996), China belongs to the expanding circle (see Figure 1-1), in which English is used as a foreign language. The English used in countries or regions from the expanding circle is considered norm-dependent; it relies on standards provided by native speakers in the inner circle (Tajeddin & Pakzadian, 2020). Although China has the largest population of English learners in the world, less than one per cent of people in China speak English conversationally (Smith, 2017).

[1] World Englishes refers to different varieties of English used in various contexts across the world (Bhatt, 2001).

Figure 1- 1 *The concentric circle model of the global profile of English (Kachru, 1996, p.25)*



Spoken English is undervalued in most regions in China, because students have to focus on reading and grammatical skills in order to pass the National College Entrance Examination (Gaokao[2]). Spoken English is not required in the national English exams (e.g. CET4, CET6[3]) for non-English majors at universities. Even for English majors, the speaking tests involved in the national English exams (e.g. TEM-4, TEM-8[4]) do not require two-way communication but a one-way interpretation or monologic presentation (Wang, 2014). Several studies (He, 2011; Liu, 2007; Shao, Yu, & Ji, 2013; Tien, 2018; Timina, 2015) showed that most Chinese EFL students at universities experience moderate or high English-speaking anxiety levels. Helping Chinese EFL learners cope with speaking anxiety is important, because language anxiety inhibits language learning and use (Subekti, 2018), has a negative influence on language performance (Liu & Jackson, 2008), and hides the willingness to participate in tasks in foreign language classes (Gregersen & Horwitz, 2002). The causes and effects of

[2] The Gaokao is an academic examination held annually as a prerequisite for entrance into almost all higher education institutions at undergraduate level. English tests for Gaokao usually include listening, reading, and writing; speaking is only included in some big cities as additional tests, but the level of difficulty is very basic.
[3] College English Test (CET) at the fourth and sixth level in China. These tests are for non-English major students.
[4] Test for English Major (TEM) at the fourth and eighth level in China. These tests are for English major students.

Foreign Language Speaking Anxiety (FLSA) will be further unpacked in one of the Literature Review chapters (Chapter 2).

1.2 Research Rationale

On a personal level, this study was inspired by my experience of learning and teaching English and of studying abroad. As a Chinese EFL learner myself, the first time I felt anxious about speaking English was when taking the International English Language Testing System test in order to apply to study in the UK. I felt very nervous while waiting for my test and noticed that my heart rate was faster. When I entered the room and faced the examiner, who was a native speaker of English, I tried to calm down, but I spoke very fast during the test even though I did not always understand the meaning of the questions. When I walked out of the test centre after finishing the test, my mind was totally blank. However, after getting used to an English-speaking environment and making some English-speaking friends when studying in the UK, I became more at ease while speaking English. When I went back to China to teach English after graduating, I was aware that most of my students have similar problems with speaking English. All these experiences motivated me to investigate the topic of anxiety when speaking a foreign language.

Reviewing empirical studies on language anxiety, it was found that the majority of studies were on the investigation of factors or sources for foreign language anxiety in the classrooms (Horwitz, 2010; Horwitz, Horwitz, & Cope, 1986; Ohata, 2005; Onwuegbuzie, Bailey, & Daley, 1999; Tanveer, 2007), and on the relationship between foreign language anxiety or foreign language speaking anxiety and language performance or proficiency (Balemir, 2009; Craig & Kim, 2010; Hewitt & Stephenson, 2012; Phillips, 1992). A few studies (Chou, 2018; El-Sakka, 2016; He, 2017; Rafieyan & Yamanashi, 2016; Sun, 2022) explored approaches to dealing with foreign language speaking anxiety. There is limited research on the effectiveness of different approaches to alleviating foreign language speaking anxiety, including a negotiated syllabus (Pakdaman, Alibakhshi, & Baradaran, 2022), the flipped classroom approach (Parvaneh, Zoghi, & Asadi, 2022), the use of paralinguistic cues (Uştuk & Aydın, 2018), self-talk (Toyama & Yamazaki, 2021) and cooperative learning (Mon, 2019).

Onwuegbuzie, Bailey, and Daley (1999) found a significant negative correlation between visiting foreign countries and foreign language anxiety. The more time a person spent in a foreign country speaking the target language, the less anxious they felt. Baker and MacIntyre

(2000) noted that students in an immersion programme were less anxious when communicating in a foreign language. In the immersion learning programme applied in the study of Baker and MacIntyre (2000), students have more opportunities to speak the target language than in the non-immersion programme. Matsuda and Gobel (2004) reported that the self-confidence of students with overseas experience was higher, so they were less anxious when speaking English. According to Dewaele (2004), this is due to an increase in the amount of authentic communication in the target language. Authentic communication means that foreign language learners use the target language to communicate a message to others who want or need it for some engaging and/or meaningful purpose (Ozawa, 2006). It also includes understanding oral or written messages from others and engaging in conversations that build up ideas (Ozawa, 2006).

It is not possible for all EFL learners to travel to or study in an English-speaking country because of cost and time constraints, but HiVR technology can offer foreign language learners a virtual environment in which to use the target language authentically without travelling (Peixoto *et al.*, 2021). This made me wonder whether practising oral English in HiVR simulating real-life scenarios, could help to reduce the level of FLSA among Chinese EFL learners.

HiVR as a cutting-edge technology is garnering increasing attention among researchers in the field of second language acquisition, but there seemed to be no studies investigating its effects on FLSA when I started my PhD project in 2018. It is claimed that the head-mounted display used in HiVR provides users with a sense of presence without the distraction of the outside world, which offers autistic people a sense of calm (Newbutt, Bradley, & Conley, 2019) and helps reduce social stress in interactions (Hartanto *et al.*, 2014). HiVR also has a positive influence on other aspects of anxiety, such as social anxiety disorders (for example, anxiety about holding conversations, meeting new people, or public speaking) and test anxiety (Alsina-Jurnet, Carvallo-Beciu, & Gutiérrez-Maldonado, 2007; Harris, Kemmerling, & North, 2002). In terms of second language acquisition, the benefits of HiVR are often seen in linguistic knowledge learning, such as learning words (Vázquez *et al.*, 2018), cultural knowledge (Cheng, Yang, & Andersen, 2017), and expressions (Sakamoto & Sakata, 2018). The effects of HiVR on language anxiety are underexplored; and examining it would broaden understanding of the emerging technology and its use with anxious students in foreign language learning.

1.3 Research Aims and Questions

At the time I designed my study, no research effort had been invested in investigating whether and to what extent HiVR can affect FLSA levels. Little attention has been devoted to understanding the use of HiVR in oral English learning from an affective or an emotional perspective. Based on OECD (2015)[5], the effectiveness of technology in facilitating learning depends on how the technology is exploited rather than the technology *per se*. My study therefore aims to compare the potential effect on FLSA levels of interventions designed with a different combination of two learning environments (HiVR and classroom) and two learning approaches (situated learning and teacher-centred learning). Both quantitative and qualitative data of participants' FLSA change before and after the intervention were collected in order to gain insights into the potential impact of HiVR on FLSA through different lenses. My research also aimed to understand participants' experience of using HiVR in oral English learning. Specifically, my study sought to investigate the following research questions:

RQ 1: To what extent (if any) do the FLSA levels of students who used HiVR change after the intervention?

RQ 2: To what extent (if any) do two learning environments (HiVR and classroom) combined with two learning approaches (situated learning and teacher-centred learning) have different degrees of influence on students' FLSA levels?

RQ 3: How do students' perceptions of their FLSA levels change before and after the intervention support, explain or refute the quantitative results?

RQ 4: What are students' perceptions of the impact of HiVR on FLSA?

RQ 5: What are students' experiences of using HiVR in oral English learning?

1.4 Research Design and Methodology

A mixed-method experimental design (Creswell, 2021) was employed to illuminate the above research questions. Details about the research design and research methods can be found in Chapter 5. An experiment with 2*2 factorial design (Fisher, 1936) was conducted to collect quantitative data. The specific group categories were shown in Table 1-1, which also related to the intervention in each group.

[5] Organization for Economic Cooperation and Development.

Table 1- 1 *Group categories and corresponding learning approaches and environments in my study*

Group name	Learning Approaches	Learning Environment
SVR	situated learning	HiVR
TVR	teacher-centred learning	HiVR
SC	situated learning	classroom
TC	teacher-centred learning	classroom

FLSA levels were obtained from four groups of participants (140 Chinese university students in total, 35 students in each group) before and after interventions. Multiple instruments were employed to obtain the FLSA levels, including a Likert scale, English speaking tests, and self-ratings of FLSA during specific tasks. The quantitative data was analysed by comparing the FLSA levels before and after the intervention among and within groups. An open-ended questionnaire was used after the experiment to gather students' perceptions and experiences of FLSA and HiVR. The qualitative data was integrated with the quantitative results for a complete understanding of the impact of HiVR on Chinese EFL learners' FLSA levels.

1.5 Significance of the Study

The significance of this study encompasses several areas. First, it broadens understanding of emerging technology (HiVR) in foreign language learning. As stated in Section 1.2, HiVR technology is often investigated in mental health research (for example social anxiety or public speaking anxiety) and in linguistic knowledge learning, including vocabulary and expressions. My study aims to fill the research gap by examining the impact of HiVR on the affective dimensions of foreign language learning and use. Employing a mixed methods approach, my research aims to provide a comprehensive understanding of this topic from different perspectives. It also takes learning approaches into account when investigating the use of HiVR technology in foreign language learning. This is pioneering in research into HiVR and FLSA.

I revolutionised the means to measure FLSA by adapting the previous retrospective method (Likert scale) and involving a self-rating approach for FLSA in specific tasks. Both exploratory and confirmatory factor analyses were conducted with the adapted Likert scale in order to further understand the construct of FLSA among Chinese EFL learners. The self-

rating of FLSA serves as an alternative for examining FLSA in specific tasks, and it gives new insights into the ontological discussion of FLSA.

1.6 Outline of the Thesis

The thesis is divided into seven chapters. Table 1-2 summarises the goal of each chapter. Chapter 1, the current chapter, introduces the background as well as the personal motivation and rationale for this research. The research aims and research questions are briefly presented, followed by the research design and methodology, and the significance of the current research.

Chapter 2 conceptualises FLSA by comparing it with similar concepts. The measures of FLSA are reviewed and challenged according to my research context. The causes and effects of FLSA are explained in order to provide an understanding of FLSA. A discussion of various approaches researched in the past with the intention of alleviating FLSA is presented at the end of the chapter.

Chapter 3 provides a thorough review of HiVR technology and its affordances in learning. It explores the potentials and the challenges of HiVR in foreign language learning, especially in terms of FLSA alleviation, in order to identify the research gap. The existing empirical studies of HiVR and FLSA are examined at the end of this chapter to present substantive findings on this topic.

Based on the literature review chapters (Chapter 2 and Chapter 3), a theoretical framework is developed in Chapter 4 in order to provide the theoretical basis for the research assumptions. In addition, it derives the research design of the current study.

Chapter 5 presents the methods and methodology of the study. It provides the rationale for the mixed methods experimental design and introduces the quantitative and qualitative data collection methods and procedure. It demonstrates the instruments used in my study, including questionnaires, speaking tests and open-ended questionnaires. The data analysis methods and the integration of quantitative and qualitative data are explained along with validity and reliability in relation to this study. Finally, the ethical considerations of this research are presented.

Table 1- 2 *Chapter summary*

Chapter title	Goal
Chapter 1. Introduction	To introduce the background, rationale, aims, design, and significance of my research
Chapter 2. Literature Review: Foreign Language Speaking Anxiety	To illustrate the concept that drives my study: FLSA
Chapter 3. Literature Review: High-immersion Virtual Reality and Foreign Language Speaking Anxiety	To introduce HiVR technology, to understand the use of HiVR in current foreign language learning literature, and to identify research gaps
Chapter 4. Theoretical Framework	To illustrate the theoretical foundations of my research assumptions and research design
Chapter 5. Research Methodology and Methods	To demonstrate the research design and the fieldwork, to explain the data analysis methods
Chapter 6. Results	To report the results of the data analysis
Chapter 7. Discussion and Conclusion	To reflect on my results and link them to previous studies, to reflect on the limitations of my study, and to provide theoretical and practical implications and future research directions

Chapter 6 technically displays both the quantitative and qualitative results of this study, following the order of Research Questions (RQs). Statistical results are examined in order to identify the change and differences in FLSA before and after the intervention, which answers RQ 1 and RQ 2. The quantitative results are integrated with analysis of participants' responses to the open-ended questionnaire in order to answer RQ 3. Students' perceptions of HiVR and FLSA and their experiences of using HiVR for oral English learning are illustrated in detail at the end of this chapter. These reflections answer RQ 4 and RQ 5.

Finally, Chapter 7 synthesises the quantitative and qualitative findings of the study and links them to previous literature. It also provides theoretical and pedagogical implications for second language acquisition researchers, Chinese EFL practitioners and learners. Limitations of the current research and suggestions for future research are presented at the end of the thesis along with the conclusion.

1.7 Glossary

Agency (in a virtual environment): One's perception of generating and controlling their actions in a virtual environment (Johnson-Glenberg, 2019).

Exposure-based therapy: A psychology method of exposing an individual to fear-provoking stimuli in the absence of aversive feedback (Craske et al., 2014).

Foreign language anxiety: "A distinct complex of self-perceptions, beliefs, feelings, and behaviours related to classroom language learning arising from the uniqueness of the language learning process" (Horwitz, Horwitz, & Cope, 1986, p.128).

Foreign language speaking anxiety: "An individual's fear or nervousness associated with either real or anticipated oral communication in a foreign language with another person or persons" (He, 2018, p.4).

High-immersion Virtual Reality: A simulation of a three-dimensional virtual environment that is generated by a computer and in which a person wears a helmet equipped with a built-in screen to move around, observe, and interact with virtual objects (Peixoto et al., 2021).

Low-immersion Virtual Reality: A type of Virtual Reality that the virtual 3D environment is viewed on a two-dimensional computer monitor (Kaplan-Rakowski & Gruber, 2019).

Presence (in a virtual environment): One's psychological state or subjective experience that overlooking the technology at some levels while using the technology (Schuemie et al., 2001). It is a feeling of being there, even when one is not physically there (Riva & Ijsselsteijn, 2003).

Situated learning: A learning theory views learning as a social, cultural, and contextual activity that should take place in the situation in which it will be applied in the future (Lave and Wenger, 1991). This theory is used as a model for learning design in my thesis. It guides the creation of the learning environment in HiVR with several characteristics, including providing an authentic context and authentic activities, providing access to expert performance and multiple roles, supporting collaborative construction of knowledge, and providing coaching and scaffolding at critical times.

Teacher-centred learning: A teaching method in which teachers are the authority of the class and control the teaching process (Jin & Cortazzi, 2006). For teacher-centred language teaching, teachers spoon-feed linguistic knowledge (e.g. grammar rules, textbook vocabulary) to the students (Su, 2019).

Virtual Reality exposure-based therapy: An application of Virtual Reality technology to achieve exposure-based therapy that exposes individuals to virtual fear-provoking stimuli without receiving aversive feedback.

Chapter 2. LITERATURE REVIEW: FOREIGN LANGUAGE SPEAKING ANXIETY

This chapter introduces several aspects of Foreign Language Speaking Anxiety (FLSA), in order to provide an understanding of this concept and highlight the importance of investigating approaches to alleviate it. It presents an overview of FLSA, including its definition, measures, causes, effects, and some commonly used approaches to address it in the field of foreign language learning. The conceptualisation of FLSA is first explained in relation to other types of anxiety or related concepts. Two methods for measuring FLSA are reviewed, including retrospective (such as a Likert scale) and dynamic approaches (such as moment-by-moment rating), in order to guide the FLSA measures adopted in my study. It illustrates two main causes of FLSA: individual factors and situational factors and summarises specific factors for FLSA among Chinese English learners. The effects of FLSA are discussed from three aspects (academic, psychological, and social) to underscore the importance of alleviating it. The final section reviews conventional approaches (that do not use technology) proposed for alleviating FLSA and empirical studies on the effectiveness of various approaches in alleviating FLSA.

2.1 Definition of FLSA

To understand the concept of FLSA as it is used in my research, this section illustrates the differences between it and some relevant concepts. The concept of anxiety is introduced because it helps to define the specific type of anxiety that FLSA sits in. Foreign language anxiety is often confused with the concept of FLSA in literature, so untangling these two concepts is necessary. Communication apprehension, public speaking anxiety, reticence, and unwillingness to communicate are concepts often discussed along with FLSA in the literature, because they are associated with the speaking situations. Distinguishing these concepts is also important.

2.1.1 Anxiety

Spielberger (1971) defined anxiety as a negative emotional state encompassing feelings of tension, apprehension, and nervousness, with intense activity of the autonomic nervous system, such as an increase in heart rate, blood pressure, and galvanic skin response. The concept of anxiety can be described as a mood state concerned with a future situation that

one is not ready or prepared to deal with (Barlow, 2000). In the psychological field, anxiety is often subdivided into trait anxiety, state anxiety and situation-specific anxiety. Trait anxiety is a relatively stable personality characteristic in which individuals are anxious in a wide range of situations (Spielberger, 1971). In other words, some people are more likely to experience anxiety than others (Crozier, 2013). By contrast, state anxiety is a transient emotional state triggered by a threatening or frightening situation (Spielberger, 1971). The experience of state anxiety can vary in intensity, frequency and duration depending on the situation (Wiedemann, 2001). Situation-specific anxiety bridges the gap between trait anxiety and state anxiety, but it takes into consideration the particular type of situation within which the anxiety occurs (MacIntyre, 2017). Situation-specific anxiety can be seen as "trait anxiety measures limited to a given context" (MacIntyre & Gardner, 1991c, p. 90).

FLSA is often considered a situation-specific anxiety (Akkakoson, 2016). According to the definition of situation-specific anxiety by MacIntyre and Gardner (1991c), FLSA can be seen as trait anxiety associated with foreign language speaking situations. However, there are various foreign language speaking situations, such as giving a presentation in public, answering the teacher's questions in class, or communicating with native speakers. Foreign language learners' speaking anxiety levels may vary in different speaking situations. The measurement of FLSA may need to consider the different speaking situations. These guide the definition and measures of FLSA in my study, which will be further unpacked in the following sections.

2.1.2 Foreign Language Anxiety

FLSA is sometimes confused with foreign language anxiety. Foreign language anxiety is also known as foreign language classroom anxiety, which in turn is sometimes referred to as language anxiety in the literature. Foreign language anxiety is seen as a situation-specific anxiety related to language learning and using context (Horwitz, 2017). Horwitz, Horwitz, and Cope (1986, p. 128) defined foreign language anxiety as "a distinct complex of self-perceptions, beliefs, feelings, and behaviours related to classroom language learning arising from the uniqueness of the language learning process". The reason a classroom language learning setting was involved in this definition may be due to the measure of foreign language anxiety developed in their study, Foreign Language Classroom Anxiety Scale, was based on participants' perceptions of foreign language learning and use at school. However,

language learning can also happen outside the classroom. From this perspective, FLSA can occur both in and outside the classroom and in any circumstances involving the oral use of the target language (Woodrow, 2006).

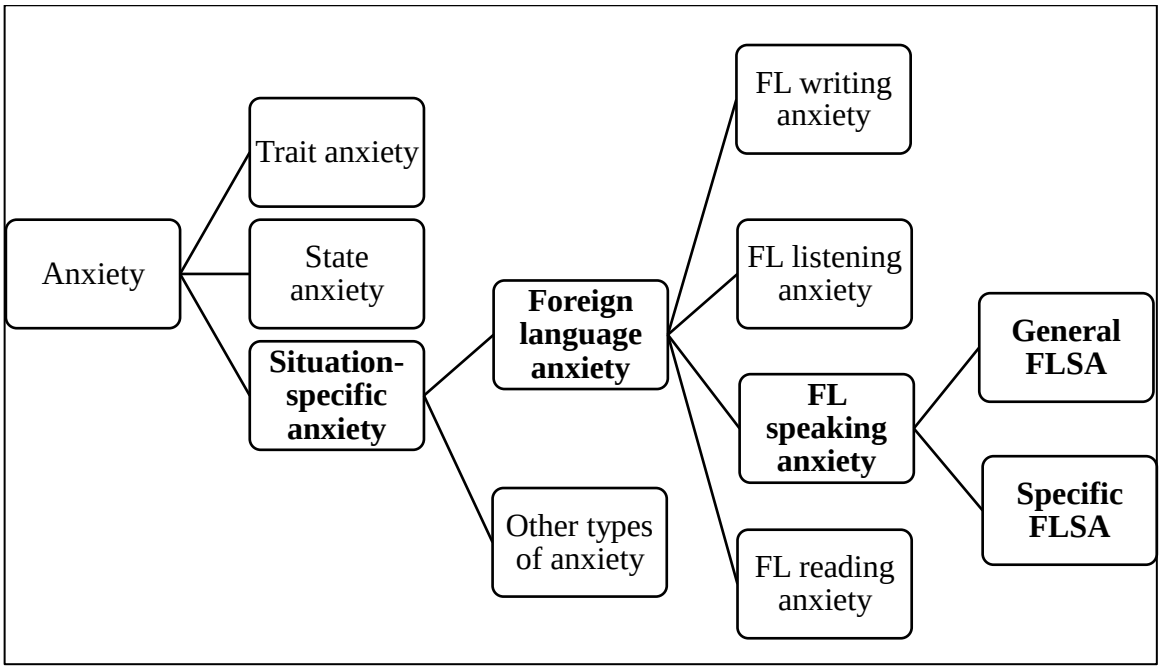
Horwitz, Horwitz, and Cope (1986) have drawn three analogies of foreign language anxiety, which are 1) communication apprehension, 2) test anxiety, and 3) fear of negative evaluation. *Communication apprehension* is a type of fear or anxiety associated with real or anticipated communication with others, which is a subconstruct of reticence or unwillingness to communicate (McCroskey, 1977). *Test anxiety* stems from a fear of failure in a test-taking situation. *Fear of negative evaluation* is broader in scope than test anxiety, since it can occur in a variety of social or evaluative situations (Horwitz, Horwitz, & Cope, 1986).

Many researchers (Al-Shuaibi, Hamdan-Mansour, & Azzeghaiby, 2014; Liu & Huang, 2011) view foreign language anxiety as a combination of the above three concepts, whereas Horwitz (2017) contended that this was a misunderstanding of the nature of foreign language anxiety. The identification of communication apprehension, test anxiety and fear of negative evaluation should not simply be seen as three-factor model of foreign language anxiety (Horwitz, 2017; Park, 2014). In addition, Horwitz (2016) strongly recommended examining the components of foreign language anxiety in different learner populations and learning contexts. This is because foreign language anxiety may vary with respect to the culture of foreign language learners, and it is possible to find different factor structures of foreign language anxiety in different learning contexts. I take the same position in relation to the concept of FLSA and have examined the factors for Chinese EFL learners throughout the literature review (see Section 2.3.3). A measurement for FLSA among Chinese EFL leaners (see Section 4.2.4) was adapted from previous scales, and the factor analysis to it was conducted to clarify the construct of FLSA among Chinese EFL learners (see Section 5.2.1.2). Qualitative data was used to further identify factors for FLSA among Chinese EFL learners (see Section 5.3.1).

Foreign language anxiety has been extended to four skills of language learning: foreign language speaking anxiety (Gregersen & Horwitz, 2002), foreign language reading anxiety (Saito, Garza, & Horwitz, 1999), foreign language writing anxiety (Cheng, 2004), and foreign language listening anxiety (Elkhafaifi, 2005). However, as discussed at the end of Section 2.1.1, it is not clear whether foreign language learners are anxious in some foreign

language speaking situations but not anxious in others. Therefore, I argue that FLSA can be divided into the concept of general FLSA (generally anxious in speaking the foreign language) and specific FLSA (being anxious in specific speaking situations). As Figure 2-1 shows, based on my understanding, foreign language anxiety is a type of situation-specific anxiety, and FLSA is a distinct type of foreign language anxiety in terms of speaking the foreign language (FL), but may differ in general and specific speaking situations.

Figure 2- 1 *The relationship of anxiety, foreign language anxiety and FLSA (my own creation)*



Foreign language anxiety and FLSA share another similarity in that they are unlikely to be invariable. Students who do not initially suffer any anxiety in foreign language learning may suffer it after repeated negative experiences, and the level of foreign language anxiety may increase after the poor experiences (Gkonou, 2017). In the same way, students may encounter FLSA after conversing with native speakers or advanced foreign language learners, and finding that their oral proficiency is not good enough to complete the conversation (Woodrow, 2006). As Gkonou (2017) states, language anxiety is not merely a matter of whether students have it or not; it is a phenomenon that is primarily caused by specific situations. This also applies to the concept of FLSA as defined in my study, in that general FLSA may differ from FLSA in specific speaking situations. For instance, a foreign language learner may not be anxious about speaking the target language in general but becomes anxious in a conversation about a complicated topic with a native speaker of the target language. This aligns with my distinction between general and specific FLSA in this thesis.

2.1.3 Other Relevant Concepts

Communication apprehension, public speaking anxiety, reticence, and unwillingness to communicate are negative emotions associated with speaking situations, which have often been mixed with the investigation of language anxiety in previous studies (Liu & Jackson, 2008, 2011). Understanding the differences and relationship of these concepts with FLSA helps to identify the concept of FLSA for my study.

Communication apprehension refers to one's fear or anxiety related to real or anticipated oral communication situations, which could be trait (feeling anxious about any communication act), context-based (a particular type of communication), audience-based (communicating with particular individuals), and situation-based (communication under a context with particular individuals) (McCroskey, 1984). Communication apprehension can be related to any languages, but FLSA discussed in my thesis is focused on speaking a foreign language. Moreover, communication apprehension can be trait, which occurs in any oral communication, but it does not apply to FLSA. FLSA is a situation-specific anxiety that only happens when an individual is speaking a foreign language.

In addition, McCroskey (1977) found that apprehensive communicators were unlikely to communicate with others, and also had low communication competence. It is not the same in terms of FLSA, where the communication competence of language learners with FLSA may not be low. Saito and Samimy (1996) observed that some advanced foreign language learners may experience high level of FLSA.

To the best of my knowledge, public speaking anxiety is a type of communication apprehension that relates to state anxiety during an oral presentation in public (Shi, Brinthaupt, & McCree, 2015). It is also possible to define public speaking anxiety as a type of social anxiety or phobia, but this definition needs to be further investigated if the distinction is to be made (Blöte *et al.*, 2009). Public speaking anxiety is different from the concept of FLSA discussed in my study, which focuses on oral interaction rather than oral presentation. Public speaking anxiety is therefore excluded from further discussion in this thesis, although a large number of studies have been conducted on the effects of High-immersion Virtual Reality (HiVR) on it.

Social anxiety, also known as social phobia, is a fear of social situations, such as meeting strangers, having conversations, and joining groups (Schlenker & Leary, 1982). Social anxiety is a domain-general psychopathological phenomenon that can happen in various context, but FLSA is domain-specific that only appears when one communicates in a foreign language with others (Zhang & Zhang, 2022). However, due to the fact that social anxiety and FLSA as discussed in my study both relate to interpersonal oral interaction, studies on HiVR and social anxiety disorders are reviewed in Section 3.4.2.

Burgoon (1976) noted that reticence was more than just fear of public speaking situations and found it in a variety of interpersonal communication situations. Unwillingness to communicate can be attributed to communication reticence, which is "a chronic tendency to avoid and/or devalue oral communication" along with a view of the communication process as relatively unrewarding in comparison with other situations (Burgoon, 1976, p. 60). In my understanding, unwillingness to communicate can be both a cause and an effect of FLSA. That is to say, foreign language learners with FLSA may avoid or devalue speaking the foreign language, and the unwillingness to communicate may also cause their anxiety about speaking the foreign language. This will be discussed in Sections 2.3 and 2.4.

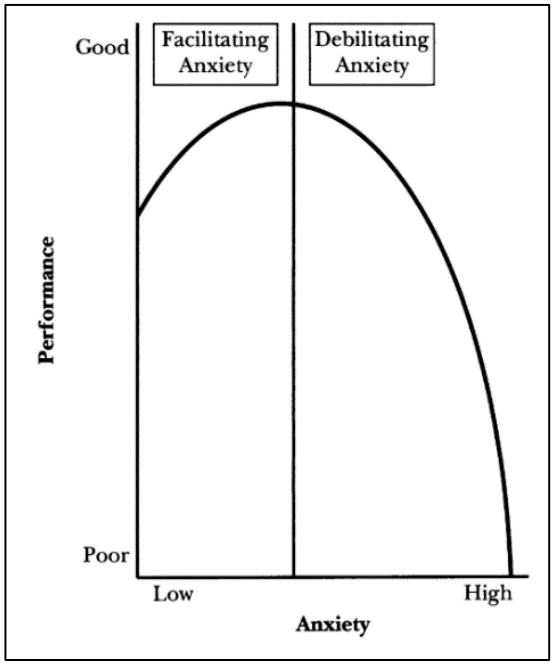
2.1.4 FLSA in this Thesis

FLSA is, above all, a distinct anxiety that arises from the use of a foreign language in a speaking situation either inside or outside the classroom. According to He (2018), FLSA is "an individual's fear or nervousness associated with either real or anticipated oral communication in a foreign language with another person or persons" (p. 4). My study takes this definition of FLSA into account but divides it into general FLSA related to any foreign language speaking situations and specific FLSA in particular speaking situations or with specific interlocutors.

Although facilitating effects of anxiety on learning has been argued by some researchers (MacIntyre, 1995), the FLSA discussed in my study is primarily responsible for the debilitating effects of anxiety on foreign language learning and use. Facilitating anxiety is seen as an appropriate level of anxiety that would motivate learners to deal with learning tasks. Figure 2-2 displays the inverted U relationship between anxiety and performance proposed by MacIntyre (1995). It demonstrates that increasing anxiety first enhances performance (facilitating anxiety), and then impairs it after the optimal level (debilitating

anxiety). However, as Horwitz (2017) argued, foreign language learning is inherently stressful for some foreign language learners, so searching for facilitating anxiety is a step backwards. Thus, the debilitating effects of FLSA, which would result in students' avoidance of speaking the foreign language, is the focus of my research.

Figure 2- 2 *Inverted U relationship between anxiety and performance (MacIntyre, 1995, p. 92)*



2.2 Measures of FLSA

There is no unified measurement of FLSA; rather, it is often mixed with what used to measure foreign language anxiety. As a result, discussion of FLSA measures cannot be separated from those concerning foreign language anxiety. Foreign language anxiety and FLSA are mainly measured in two ways: retrospectively and dynamically. The retrospective approach often deems foreign language anxiety or FLSA as an invariable concept and measures it based on memories of previous events via a scale (MacIntyre, 2017). According to the dynamic perspective, foreign language anxiety or FLSA fluctuates throughout a conversation or a speech and should be measured moment-by-moment (MacIntyre, 2017). The approaches to measuring foreign language anxiety and FLSA in previous studies are illustrated in this section in order to address the adaptation of measurement of FLSA in my study.

2.2.1 Retrospective Approach

Scales for foreign language anxiety in classrooms

In general, the most commonly used method of assessing foreign language anxiety or FLSA is the retrospective approach, for example a self-reported Likert scale (Botes, Dewaele, & Greiff, 2020). Table 2-1 presents a summary of Likert scales employed in previous studies to measure foreign language anxiety.

Table 2- 1 *Scales for foreign language anxiety developed in previous studies (my own creation)*

Year	Authors	Scale name or purpose
1975	Gardner and Smythe	French Class Anxiety Scale
1977	Clément, Gardner and Smythe	English language use anxiety in the classroom
1985	Gardner	French language use anxiety in the classroom
1986	Ely	Language Class Risk-taking, Sociability and Discomfort Scales
1986	Horwitz, Horwitz, and Cope	Foreign Language Classroom Anxiety Scale

The first self-reported measure of anxiety surrounding second language learning was the French Class Anxiety Scale (Gardner & Smythe, 1975). The survey consists of 11 items related to uncomfortable feelings or anxiety while speaking French or participating in a French class (Gardner & Smythe, 1975, p. 280). Clément, Gardner, and Smythe (1977) proposed a five-item scale focused on English language use anxiety in the classroom, and it was adapted by Gardner (1985) to examine anxiety in the French classroom. Language learning anxiety seems to be only a subset of these scales, rather than the main focus, and the scales seem too simplistic with just five items. Ely (1986) developed three scales for language class risk-taking, sociability and discomfort with some other short scales, in order to understand their relationships with each other. Two items of the Language Classroom Discomfort Scale in the study of Ely (1986) were adapted for the measurement of FLSA in my study (see Section 5.3.4.1).

For measuring specific foreign language classroom anxiety, Horwitz, Horwitz, and Cope (1986) developed a Foreign Language Classroom Anxiety Scale (FLCAS), which has

become the most widely used approach to examining foreign language anxiety. The FLCAS is a five-point Likert scale of 33 items, based on the findings of 30 university students' potential sources of anxiety in foreign language classrooms. It is reflective of communication apprehension, test anxiety, and fear of negative evaluation (Horwitz, Horwitz, & Cope, 1986). Discussions on these three concepts can be found in Section 2.1.2. Two items of the FLCAS were borrowed for the scale used to assess the FLSA in my study (see Section 5.3.4.1).

In addition, researchers (Aida, 1994; Park, 2014) have conducted factor analysis to determine the components of the FLCAS in order to clarify the nature of foreign language anxiety and identify specific strategies to help students who suffer from foreign language anxiety. I will apply factor analysis to the scale used to measure FLSA in my study in order to understand Chinese EFL learners' speaking anxiety (see Section 6.2.1.2). Factor analysis is a method of reducing the number of variables and classifying them (He, 2011). A pioneering study of factor analysis on the FLCAS was conducted by Aida (1994), who identified four factors influencing foreign language anxiety levels: 1) speech anxiety and fear of negative evaluation, 2) fear of failing the class, 3) comfortableness in speaking with native speakers, and 4) negative attitudes towards the language class. Several studies have attempted to determine the components of foreign language anxiety since Aida's work. Cheng, Horwitz, and Schallert (1999) identified two components of the FLCAS, namely low self-confidence in speaking English and general English classroom performance anxiety. Matsuda and Gobel (2004) considered these two factors to be meaningful elements influencing foreign language anxiety.

Nevertheless, most of the studies explored foreign language anxiety dimensions based only on their data rather than confirming it with the FLCAS framework (Park, 2014). Park (2014) used an exploratory factor analysis, which is used to reduce data to a smaller set of summary variables, followed by a confirmatory factor analysis, which is used to determine whether components generated by exploratory factor analysis adequately fit the data to clarify the factor structure of FLCAS. Park (2014) found that 1) communication apprehension and understanding, and 2) communication apprehension and confidence accounted for the most variance among his Korean students of English. Park (2014) emphasised the results of previous studies may be affected by different translations of FLCAS. Furthermore, Horwitz (2016a) suggested that differences in population and learning situations would influence

foreign language anxiety. Attention should be paid to cultural background and learning context in empirical studies of the components of foreign language anxiety (Horwitz, 2016). Therefore, I developed an English Speaking Anxiety Scale (see Appendix 1) for Chinese EFL learners in my study. Both exploratory factor analysis and confirmatory factor analysis were conducted to identify the components of FLSA investigated in my study (see Section 5.2.1.2), in order to understand FLSA among Chinese EFL learners.

Specific scales for FLSA

In recognition of the fact that scales (e.g. FLCAS) used to measure general foreign language anxiety did not reflect the exact language use situation, several researchers (Teimouri, Goetze, & Plonsky, 2019) developed language skill-specific anxiety scales. Scales were created for measuring foreign language writing anxiety (Cheng, 2004), foreign language reading anxiety (Saito, Garza, & Horwitz, 1999), and foreign language listening anxiety (Elkhafaifi, 2005). Despite this, measurement of FLSA did not see a significant improvement, but FLCAS is often used for measuring FLSA in previous studies. This is because FLCAS is viewed as focusing primarily on the oral aspects of language use, including speaking and listening (Phillips, 1992). However, FLCAS is mainly focused on classroom settings and does not measure anxiety related to specific oral tasks (Phillips, 1992).

Due to the limitation of FLCAS in measuring FLSA outside classrooms or during specific tasks, some researchers adapted it for their specific research purposes. Table 2-2 presents the main scales adapted from FLCAS for FLSA measurement in existing literature.

Table 2- 2 *Scales developed for FLSA measurement in previous literature (my own creation)*

Year	Authors	Scale name or purpose
2003	Saltan	Speaking anxiety questionnaire
2006	Woodrow	Second Language Speaking Anxiety Scale
2009	Balemir	Foreign Language Speaking Anxiety Scale
2012	Mahmoodzadeh	Speaking anxiety questionnaire
2013	Melouah	Foreign Language Speaking Anxiety Scale
2016	Li	Oral English Anxiety Scale for Chinese Learners of English

In the study by Mahmoodzadeh (2012) as an example, the FLCAS was modified by dividing it into three subsets based on interlanguage phonology, interlanguage grammar and interlanguage meaning system. Through this approach, Mahmoodzadeh (2012) developed an 18-item 5-point Likert scale to measure FLSA among EFL learners. Seven items from his speaking anxiety questionnaire were adapted for my study use (see Section 5.3.4.1). Similarly, Melouah (2013) constructed a Foreign Language Speaking Anxiety Scale (FLSAS) by selecting and modifying 23 items from the FLCAS, in order to examine the nature and causes of EFL speaking anxiety in students. Six items from the FLSAS (Melouah, 2013) were selected for my research (see Section 5.3.4.1).

In order to create an FLSA questionnaire, Saltan (2003) selected 18 items concerning speaking anxiety from the FLCAS along with some items from a questionnaire by Young (1990). This scale, which Saltan (2003) developed for his Master's dissertation, was employed by other researchers (Çağatay, 2015; Karatas *et al.*, 2016; Öztürk & Gürbüz, 2014) to collect FLSA data. Since Saltan's (2003) questionnaire has some overlap with the scales developed by Mahmoodzadeh (2012) and Melouah (2013), it was not taken into consideration for my measurement of FLSA.

The FLCAS has sometimes been used in conjunction with the Unwillingness to Communicate Scale in order to examine students' FLSA (Yalçın & İnceçay, 2014). The Unwillingness to Communicate Scale is a 26-item scale developed by Burgoon (1976) and can be applied to assess students' attitudes to communication, as well as their communication behaviours with family, peers, teachers and administrators. One item in the Unwillingness to Communicate Scale was borrowed for the scale used in my study (see Section 5.3.4.1).

In addition, some researchers have created instruments for their own research purposes in order to examine FLSA. Woodrow (2006) introduced the Second Language Speaking Anxiety Scale, which is a five-point Likert-scale that takes into account the communication situation the participants encountered, including communicative setting, interlocutor variables (e.g. the number of speakers, the status of the speakers, and whether the speakers were native or non-native speakers) and the nature of their communication. As the Second

Language Speaking Anxiety Scale (Woodrow, 2006) is accounted for a second language[6] rather than a foreign language, it was not adapted for my research.

Balemir (2009) developed a 28-item 5-point Likert scale to acquire specific information related to the sources of speaking anxiety by adapting items from Foreign Language Speaking Anxiety Scale (Huang, 2004). One item was borrowed from the scale of Balemir (2009) for the measurement of FLSA in my study (see Section 5.3.4.1), without overlapping with items adapted from other scales.

Li (2016) developed an Oral English Anxiety Scale for Chinese Learners of English (OECASCL) to obtain participants' attitudes to and experience of speaking English in class. OECASCL is a lengthy scale created by adapting and combining items from the FLCAS (Horwitz, Horwitz, & Cope, 1986), the Unwillingness to Communicate Scale (Burgoon, 1976), the State-Trait Anxiety Inventory (Spielberger, 1971), and the Language Class Risk-taking Scale and Language Class Sociability Scale (Ely, 1986). The focus of OECASCL is on speaking English in the classroom, so only two items related to general speaking situations were selected for my research use (see Section 5.3.4.1).

2.2.2 Dynamic Approach

Since retrospective measurement cannot capture the dynamics of in-the-moment anxiety on short timescales (Gregersen, MacIntyre & Olson, 2017), alternative measures for the fluctuations of anxiety in real time have been developed. This is based on the dynamic perspective of foreign language anxiety, which emphasises changes in foreign language anxiety within a different time span and the coexistence of contradictory feelings (e.g. foreign language anxiety and foreign language enjoyment) (Gregersen, 2020). In the dynamic perspective, anxiety can be seen as an emotion fluctuating over time, and the fluctuation can be revealed in a short-term (e.g. in a short communication) or a long-term (e.g. weeks or months or years of classes) period (Gregersen, MacIntyre, & Olsen, 2017). Following the dynamic approach, individual reactions to the event are analysed rather than the group-level or individual-level traits (Boudreau, MacIntyre, & Dewaele, 2018). In

[6] Second language refers to a language that leaners are naturally exposed to rather than their mother tongue (Smelser & Baltes, 2021).

general, FLSA measurement underlying dynamic approach could be divided into three categories: self-rating, observer rating, and physiological measure.

Self-rating

Dynamic self-ratings of FLSA levels have been explored for three decades by a few scholars. A self-rating scale, the *anxometer*, was developed by MacIntyre and Gardner (1991a) by adapting the fear thermometer (Walk, 1956) to measure language anxiety in production tasks. *Anxometers* are "single item, visual analogue scales with a range values from one to ten printed along with the side of a thermometer figure" (MacIntyre & Gardner, 1991a, p. 298). Participants were asked to draw a line on the *anxometer* to indicate their anxiety levels during each oral task. Gardner *et al.* (2004) developed a measure for state anxiety in class, which was a 7-point Likert scale ranging from -3 (low anxiety) to +3 (high anxiety). A similar printed grid with a 7-point Likert scale (1 is the lowest anxiety level, and 7 is the highest anxiety level) with 18 blank items (each item for an interval of 15 minutes) was used by Mahmoodzadeh (2015) to measure the intensity of students' in-class anxiety. However, the comparison in these studies were all made at a group level rather than an individual reaction level. For example, Mahmoodzadeh (2015) analysed males' and females' differences in in-class foreign language anxiety. This is not strictly a dynamic approach, which only considers the individual FLSA rather than any group levels.

An idiodynamic approach was proposed by MacIntyre and Legatto (2011) to study moment-to-moment fluctuations in Willingness to Communicate using a software application (https://petermacintyre.weebly.com/idiodynamic-software.html). By reviewing the recordings of a communication test using this software, participants can rate their levels of willingness to communicate from -5 to +5 in each second based on their performance. A graph showing the fluctuations in willingness to communicate levels was then generated by the software. This idiodynamic method was further developed to investigate anxiety levels in oral presentations (Gregersen, Meza, & MacIntyre, 2014) and both language anxiety and enjoyment levels in oral tasks (Boudreau, MacIntyre, & Dewaele, 2018).

The self-rating method used at the early stage (Gardner *et al.*, 2004; MacIntyre & Gardner, 1991a) was adapted to measure participants' FLSA levels in the speaking tasks in my study (see Section 5.3.6), and the analysis was made at a group level rather than individual level (see Section 6.2.2). This is because the purpose of my study was to examine the impact of

High-immersion Virtual Reality on FLSA among a group of EFL learners rather than identify the changes of FLSA levels in each individual. I therefore named it self-ratings of FLSA in specific tasks rather than a dynamic approach (see Section 5.3.6).

Observer rating

The idiodynamic method mentioned above has been used to measure FLSA by participants themselves for immediate rating as well as by external observers (e.g. peers, teachers, and groups) for later rating (Gregersen, MacIntyre, & Olsen, 2017). The early stages of FLSA measurement excluded observer rating, because it was impossible and/or extremely difficult to determine observable behaviours related to communication (McCroskey, 1970). This may be due to the limitations of technology in that era and no anxiety patterns were established at that time.

With the development of technology such as cameras and computers, a study (Gregersen, 2005) of video recordings of an oral examination identified nonverbal behaviours that could help detect FLSA. The findings revealed that learners who had limited facial expressions (e.g. eyebrow movement and smiling), less eye contact, and more hand use (e.g. self-touching and object manipulation) were more likely to be anxious. Gregersen (2007) confirmed that nonverbal awareness training would enhance teachers' accuracy in decoding FLSA. However, these studies (Gregersen, 2005, 2007) were undertaken with a small sample size (N = 6), which provides insufficient evidence to demonstrate a relationship between individual nonverbal behaviours and FLSA. Take the less eyebrow movement, which Gregersen (2005) identified as an indicator for FLSA as an example; it may be caused not only by FLSA, but also by concentration, hesitation, and searching for words (Kamisznikow, 2003). Moreover, Gregersen (2005, 2007) detected facial expressions separately as signs of FLSA. This is not convincing to me because facial expressions are complex dynamic signals and should be represented as action units. Action units are the fundamental actions of individual muscles or groups of muscles (Yao *et al.*, 2021). Hence, a single facial signal may not represent an internal emotion, but a complex combination of several facial expressions may.

There is evidence to suggest that both visual and auditory data could be used to determine FLSA more effectively than visual-only or auditory-only data (Gregersen, 2009). Nevertheless, smiling and pausing are the most problematic cues in detecting FLSA due to

the difficulty in distinguishing fake smiles from genuine smiles, as well as in identifying a pause due to lack of language proficiency or a pause caused by negative emotions. In their study, Gregersen, MacIntyre, and Olsen (2017) compared the accuracy of observer ratings by peers and expert practitioners with students' own ratings. They found that it was easier to identify indicators of rising anxiety than declining anxiety, and that experts provided more accurate decoding of FLSA than peer observers did. Considering the insufficient evidence in detecting nonverbal behaviours for the occurrence of FLSA, observer rating was excluded for my thesis.

Physiological measures

Other than dynamic subjective measures (e.g. self-rating or observer-rating) as discussed above, a few objective approaches were utilised to measure FLSA dynamically. Some researchers (Gregersen, Meza, & MacIntyre, 2014) monitored participants' heart rate during presentations or speaking the foreign language in class. They concluded that changes in heart rate could be attributed to other body movements rather than affective arousal (e.g. a rise in anxiety levels), so heart rate may not be a reliable measure of FLSA.

In a recent study by Thrasher (2022), salivary cortisol levels were used to determine FLSA levels. It has been found that salivary cortisol levels are moderately correlated with FLSA levels measured by a foreign language anxiety questionnaire adapted from the Foreign Language Classroom Anxiety Scale (Horwitz, Horwitz, & Cope, 1986) and Second Language Speaking Anxiety Scale (Woodrow, 2006). These two scales are discussed in Section 2.2.1. The measurement of salivary cortisol levels may provide an objective means of examining FLSA levels, although further research is required. Due to the fact that there was as yet no convincing evidence of accurate physiological measures of FLSA, this method was not included in my research.

2.2.3 FLSA Measures in China

As mentioned in Section 2.1.2, target population and learning context should be considered when measuring language anxiety (Horwitz, 2016). For foreign language learners in China, especially for English learners, the development of measurement for FLSA is in its infancy. Some researchers applied Foreign Language Classroom Anxiety Scale (FLCAS) with minor changes to measure FLSA. For example, Liu (2007) replaced *foreign language* with *English*

in FLCAS to measure FLSA in Chinese English learners. Liu and Jackson (2008) added three further items to it to reflect the situation in English classrooms in China. These were "I get tense and nervous when talking to a person whose sex is opposite to me", "I get tense and nervous when I have to discuss things unfamiliar to me in English", and "I feel overwhelmed by the number of words I have to learn to speak English".

He (2011) proposed a differentiated version of Foreign Language Speaking Anxiety Scale (FLSAS) to examine Chinese EFL learners' speaking anxiety. The FLSAS assesses both debilitating and facilitating FLSA, self-perception, oral communication anxiety, oral test anxiety and fear of negative evaluation. He (2011, 2018) used a combination of FLCAS and FLSAS to obtain both general and specific pictures of students' foreign language anxiety and FLSA levels. Six items in the FLSAS (He, 2011, 2018) have been adapted to the Likert scale used to measure FLSA in my study (see Section 5.3.4).

2.3 Causes of FLSA

The majority of previous research focused on factors contributed to language anxiety in classrooms rather than FLSA as defined in my study, which can also be related to situations outside the classroom. According to MacIntyre (2017), language anxiety is described as a feeling of worry and negative emotion that arises while learning or using a second/foreign language, and this anxiety has both internal and social dimensions. This means that students' internal physiological process, cognition, and emotional states, as well as the demands of the situation and the presence of other people, are contributing factors to language anxiety (MacIntyre, 2017). There are other ways to categorise the causes of foreign language anxiety, such as learner-internal (learner self-related) and learner-external (teacher- and peer-related) variables (Jiang & Dewaele, 2019). Amorati and Venturin (2021) divided foreign language anxiety factors into four categories: 1) instructional practices and classroom procedures; 2) counterproductive beliefs about language learning; 3) fear of errors and errors correction; and 4) underestimation of one's own abilities.

The separate investigation of factors related to FLSA was primarily initiated within the past decade. However, there is some overlap between the causes of foreign language anxiety and FLSA. Thus, the discussion of factors relating to FLSA in this section covers both foreign language anxiety and FLSA and is based on the individual and situational categories

proposed by MacIntyre (2017). A framework for the causes of Chinese EFL learners' FLSA is given at the end of this section.

2.3.1 Individual Factors

The term *individual factors* refers to a student's personal and internal elements that relate to language learning. As a result of reviewing previous studies (Balemir, 2009; Dewaele & Al-Saraj, 2015; Gregersen & Horwitz, 2002; Liu, 2006; Omaggio, 1992; Young, 1992), foreign language anxiety or FLSA may be caused by individual factors such as low self-esteem and personality characteristics (such as shyness, reticence or perfectionism).

Low self-esteem

An individual's self-esteem is a subjective evaluation of their abilities and limitations (Orth & Robins, 2014). Researchers (Brown, 2000; Cohen & Norst, 1989) found that low self-esteem or a loss of self-esteem is a distinct obstacle in second/foreign language acquisition. A terror management theory proposed by Greenberg *et al.* (1992) asserts that self-esteem provides protection against anxiety, leading to a lower level of anxiety in people with high self-esteem. The learner's self-esteem in regard to language learning (e.g. perceived competence and subjective proficiency) is more important than external reality (e.g. objectively measured achievement and proficiency) (Bailey, 1983). According to Mak and White (1997), a lack of self-esteem is the main source of communication apprehension (see Section 2.1.3 for a discussion of communication apprehension). A low level of self-esteem among students causes great concern about the correctness of their language and the quality of their accent (Horwitz, Horwitz, & Cope, 1986). Clément, Dörnyei, and Noels (1994) and Onwuegbuzie, Bailey, and Daley (1999) all confirmed that subjective proficiency and perceived competence are related to foreign language anxiety. If students perceive their proficiency level to be higher and consider themselves competent, they are more inclined to take risks, feel secure, and put in effort into foreign language learning (De Andrés & Arnold, 2009).

In terms of FLSA, anxious learners often underestimate their speaking proficiency. The main concern pertains to pronunciation, since students strive to achieve native-like pronunciation and accent (Alnahidh & Altalhab, 2020; Ansari, 2015). The underestimation of linguistic skills or speaking abilities results in over-focus on deficiencies, thereby increasing students'

FLSA levels (Kitano, 2001). For example, Kurdish university-level EFL learners were not satisfied with their levels of speaking proficiency and were concerned about failing in English classes. This resulted in the students experiencing a high level of foreign language anxiety (Ahmed, 2016). Akkakoson (2016) found that 38.09% of Thai EFL learners identified limited vocabulary as the primary cause of FLSA in his study. A lack of grammatical knowledge was revealed to be another main cause of FLSA among Saudi EFL learners (Alnahidh & Altalhab, 2020), Iranian EFL learners (Kasbi & Shirvan, 2017), and Indian EFL/ESL (English as a second language) learners (Ansari, 2015). In addition, speakers may feel anxious if they are unable to understand other speakers since the act of speaking is a reciprocal process in which interlocutors take turns being speakers and listeners, and listeners must first understand the other person's utterance in order to respond (Feng, 2007). Therefore, improving one's perceived language proficiency has the potential to reduce FLSA levels.

Personality

Personality traits such as shyness, quietness or reticence are also seen as basic causes of FLSA (Balemir, 2009; Dalkiliç, 2001). Personality traits are consistent patterns in the way individuals behave, feel and think (Cervone & Pervin, 2015). Dewaele and Al-Saraj (2015) found that students who are willing to take active social interaction are less likely to suffer from FLSA. However, students with perfectionism are likely to experience FLSA. According to Gregersen and Horwitz (2002), perfectionism and language anxiety share similar manifestations, resulting in fear of making mistakes and eagerness to produce error-free utterances. The fear of being laughed at, embarrassed and making a fool of oneself also can be associated with perfectionism (Liu, 2006). Students' low self-esteem is also correlated with their perfectionism. People with low self-esteem often worry about what others think and attempt to please others, which can lead to a considerable amount of anxiety (Omaggio, 1992).

In addition, students with trait competitiveness are more likely to experience foreign language anxiety or FLSA. Trait competitiveness refers to the desire to win and be better than others (Spence, 1984). However, the competitiveness linking to FLSA is negative comparisons with other learners, which would result in students' low self-concept (Bailey, 1983; Rubio, 2014). Self-concept refers to "the individual's belief about himself or herself,

including the person's attributes and who and what the self is" (Baumeister, 1999, p. 247). Learners with competitive attitudes may have unrealistic expectations, or may be perfectionists (Ohata, 2005). Competitiveness can also be regarded as a situational factor aroused by the external environment. It is because there should be peers who the foreign language learners compared themselves with in the social or situational environment. For example, the foreign language instructor may increase competitiveness among students in class by asking questions that students must answer (Price, 1988; Saito & Samimy, 1996). Certain educational systems, particularly Chinese exam-oriented education system, may also contribute to the competitiveness, because students are often taught that they have to perform better than others in tests. This also links to the cultural influences discussed in Section 2.3.2.3.

Changing a personality trait is difficult, but educators or instructors may make efforts to reduce the negative aspects of competitiveness among foreign language learners in order to alleviate their FLSA.

2.3.2 Situational Factors

Situational factors refer to the social and external environment involved in the student's language learning process. Several situational factors contribute to foreign language anxiety or FLSA, including poor prior experience in foreign language learning and use (Gkonou, 2017), lack of experience of language use (Ansari, 2015), and cultural influences (Ohata, 2005).

Poor experience

In terms of foreign language anxiety, poor experience in foreign language learning is considered an important contributing factor (Gkonou, 2017). Poor prior experience may include misunderstanding the words, providing wrong answers in front of classmates, looking or sounding stupid as well as instructors correcting students' errors harshly (Young, 1991). Moreover, when students over-study without improving their grades (Horwitz, Horwitz, & Cope, 1986), fail to communicate effectively or make errors in pronunciation (Price, 1988), they tend to become anxious. The specific language-skill anxiety (e.g. FLSA) may be formed during a poor experience of the foreign language learning process (Cheng, Horwitz, & Schallert, 1999). In addition, Cheng, Horwitz, and Schallert (1999) claimed that

failing to perform one's language skills (e.g. speaking) as expected might result in negative emotions and attitudes about that specific language skill. Thus, it is important to reduce poor experiences and increase pleasant experiences in learning and speaking the foreign language in order to alleviate FLSA.

Lack of experience

Another important source of FLSA is lack of experience in or low frequency of language use. Ansari (2015) interpreted this factor as a lack of exposure and practice platforms. Akkakoson (2016) found that 15.08% of Thai EFL learners attributed their FLSA to not having opportunities to speak English in their daily lives. Baker and MacIntyre (2000) suggested that immersion in the target language reduces foreign language anxiety. According to their study, the more opportunities students have to speak the target language, the better their ability to predict and confirm expectations during the conversation, resulting in a higher sense of perceived competence or self-esteem. Dewaele (2008) further confirmed that when students practiced authentic communication in the target language, their self-confidence increased, and their foreign language anxiety decreased with the higher frequency of language use. Authentic communication means the foreign language is used for communication purpose to convey meaning rather than for accomplishing quizzes focusing on the norms of the language (Ozawa, 2006).

A prior visit to foreign countries can predict FLSA, which may also be attributed to the increase in authentic communication. Onwuegbuzie, Bailey, and Daley (1999) found a significant negative correlation between students visiting foreign countries and FLSA. The more a person spends in a foreign country speaking the target language, the less FLSA they experience. Matsuda and Gobel (2004) reported that foreign language learners with overseas experience in the target-language-speaking country had higher self-confidence and were less anxious while speaking the foreign language. Mede and Karaırmak (2017) found that having a foreign friend and having prior foreign language experience in or outside the learning context contributed to the reduction of FLSA.

Nonetheless, it is not possible for every foreign language learner to visit the foreign countries in real life. High-immersion Virtual Reality (HiVR) may help with this because it can immerse the users in the virtual environment via using a head-mounted display and provide opportunities for authentic communication through simulating real-life scenarios. This is the

main reason why my study hypothesises HiVR may have potential for alleviating FLSA. This will be further illustrated in Section 3.2.

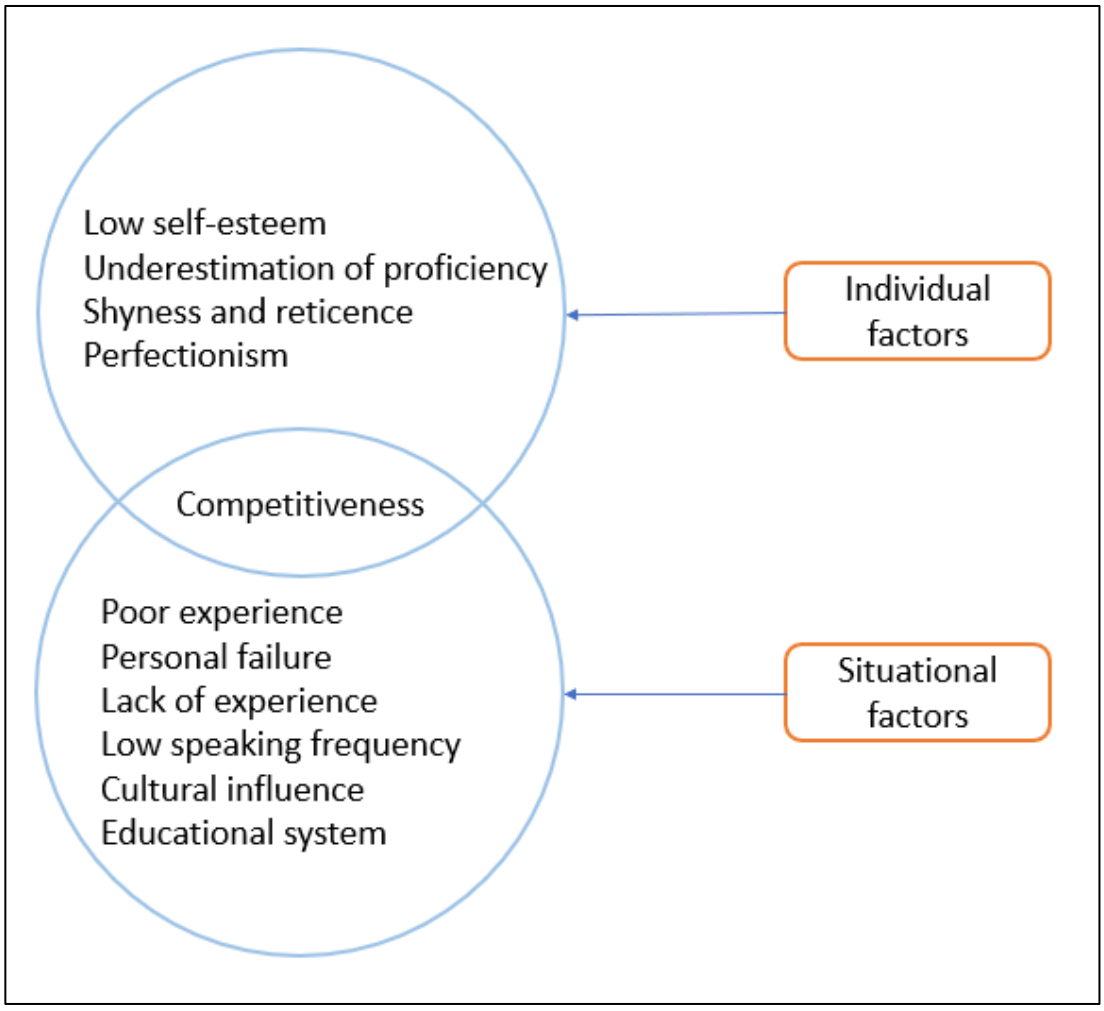
Cultural influences

Culture also plays a significant role in the emergence of anxiety in foreign language learning and use. Young (1992) stated that anxiety might arise unless one feels very proficient and comfortable in the foreign culture. It was noted in the study of Ohata (2005) that Japanese English learners had difficulty adapting to Western culture, which caused them to feel frustrated by the way they spoke and behaved in the new environment. A sense of identity loss is accompanied by anxiety because the motivation to be part of the community of target-language speakers would force students to align their utterance with those of the native speakers (Omaggio, 1992).

In addition, the educational system plays a critical role in the influence of a culture. Suleimenova (2013) suggested that students may experience anxiety when speaking a foreign language if the educational system is not designed to increase the communicative competence of students. In Iran, for example, the university entrance exam focuses on grammar and vocabulary rather than speaking (Kasbi & Shirvan, 2017). China also heavily emphasises written English instead of speaking (Jiang & Dewaele, 2019). The specific influence of Chinese culture on English learners' anxiety is demonstrated in the following section.

The factors relating to FLSA based on previous literature are summarised in Figure 2-3. In terms of individual factors, low self-esteem and underestimation of proficiency are the main predictors of FLSA. Personality traits such as shyness, reticence, perfectionism, and competitiveness may also result in FLSA. However, competitiveness can also be seen as a situational factor caused by the external environment. With regard to situational factors, poor experience including personal failure, lack of experience and low speaking frequency, as well as cultural influences such as the educational system, may contribute to the occurrence of FLSA.

Figure 2- 3 *My own summary of factors for FLSA based on previous literature*



2.3.3 Factors for FLSA Among Chinese EFL Learners

It is necessary to investigate the uniqueness of Chinese EFL learners' emotions because of the specific cultural and educational context (Jiang & Dewaele, 2019). The causes of FLSA specifically among Chinese EFL learners were investigated by several researchers in recent years (He, 2011, 2013; Liu, 2006, 2007; Mak, 2011; Tien, 2018; Timina, 2015). Among the causes, Liu (2007) emphasised that Chinese EFL learners lack experience in speaking English both in and out of the classroom. This is due to the fact that of the three concentric circles of World Englishes (Kachru, 1996), China belongs to the expanding circle (See Chapter 1). In class, few opportunities exist for practising English orally because of the exam-oriented learning (Jiang & Dewaele, 2019). The importance of spoken English is undervalued, since students have to develop reading and grammatical skills in order to pass the National College Entrance Examination (Gaokao, see explanation in Section 1.1) in China (Su, 2019). Out of class, Chinese people rarely speak English in their daily lives (Gass,

Behney, & Plonsky, 2020). As a result of the lack of experience, students lack familiarity with the conversation topics, the partners, and the speaking environment, which in turn increases their anxiety levels while speaking the foreign language.

The teacher-centred instruction style and hostile classroom atmosphere also contribute to the increase in anxiety levels (Suleimenova, 2013; Woodrow, 2006). Teacher-centred instruction is widely used in teaching English in China. In this method, teachers spoon-feed linguistic knowledge of English (e.g. grammar rules, textbook vocabulary) to the students (Su, 2019). In teacher-centred instruction, teachers may force students to participate and answer questions (Alnahidh & Altalhab, 2020), which leads to students becoming afraid of receiving a negative evaluation (Ahmed, 2016) or being judged by others (Melouah, 2013). In Chinese culture, students are also afraid of making mistakes due to their desire to save face (Liu, 2006). The term 'face' refers to how one is perceived by others (Jiang & Dewaele, 2019). Avoiding losing face, or saving face, means avoiding losing others' respect or avoiding embarrassment and is extremely important in Chinese culture (Lindridge & Wang, 2008).

In addition, according to Chinese tradition, students are not encouraged to express themselves or ask questions but rather to listen to the teachers (Hu, 2002). The teacher is the authority in the classroom, and often points out students' mistakes harshly and directly (Hu, 2002). Due to the harsh correction in class, Chinese EFL learners are more concerned with the fact that their accent and pronunciation are not native-norm, and they are eager to speak fluently (Yan & Horwitz, 2008). The majority of Chinese students feel that their proficiency in English is not good enough for communicating with others (Liu & Jackson, 2008). Typically, they are worried about their vocabulary range, grammatical knowledge and listening comprehension ability (Yan & Horwitz, 2008). Moreover, because Confucianism states that silence has a positive connotation, Chinese people are generally shy and dislike speaking in public (Woodrow, 2006).

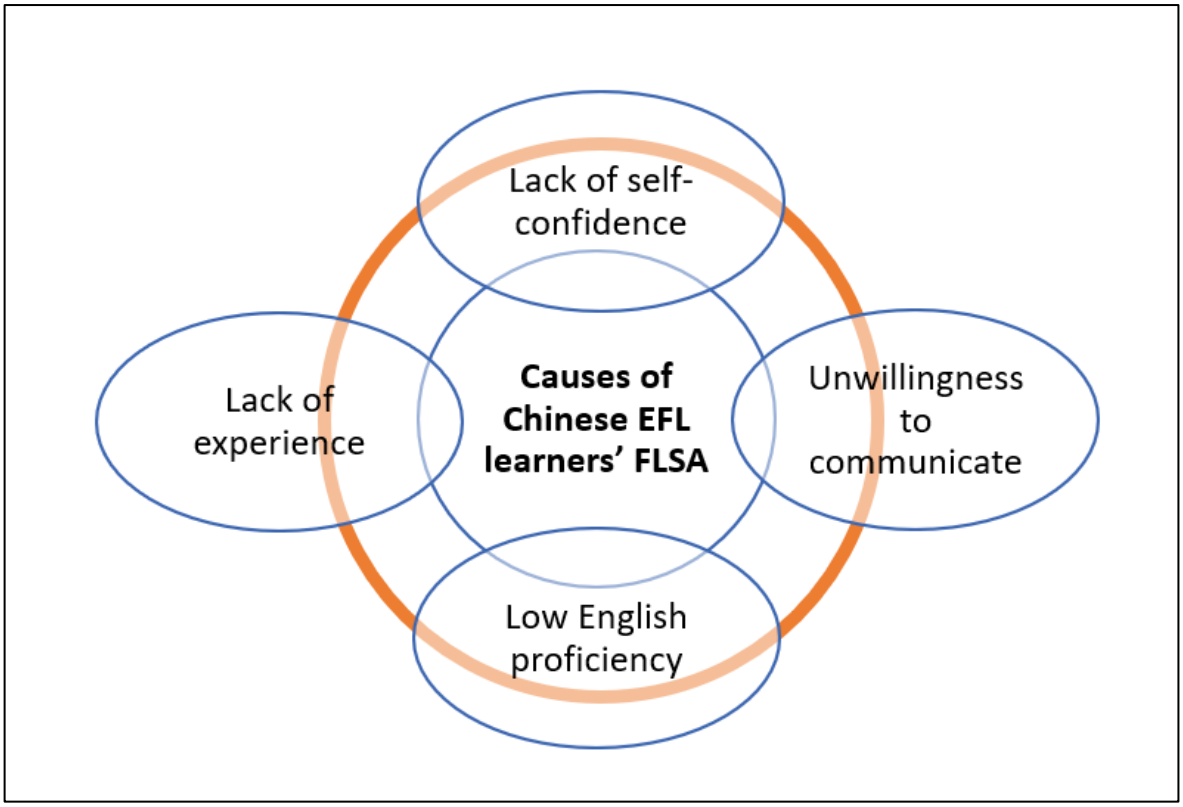
The literature (He, 2011, 2013; Liu, 2006, 2007; Mak, 2011; Tien, 2018; Timina, 2015) identifies a number of factors relating to FLSA among Chinese EFL learners. I have grouped these factors into four categories: lack of experience, lack of self-confidence, low English proficiency, and unwillingness to communicate (see Table 2-3).

Table 2- 3 *My categorisation of causes of Chinese EFL learners' FLSA based on previous literature*

Categories	Factors
Lack of experience	Lack of familiarity with the topic; lack of familiarity with speaking partners; lack of familiarity with the environment; lack of preparation; lack of practice in speaking English; teachers do not encourage students to speak English
Lack of self-confidence	Concern about pronunciation; fear of being unable to follow and understand others or task; fear of making mistakes; fear of being laughed at/embarrassed; fear of being the focus of attention; fear of losing face; eagerness to speak fluently
Low English proficiency	Lack of vocabulary; incomplete understanding of other's utterances; lack of linguistic knowledge to express ideas; inadequate grammatical knowledge
Unwillingness to communicate	Personality (shyness, reticence, and reluctance to speak in public); difficulty of the task; time pressure; lack of preparation; personal failure while speaking English; harsh error correction and low tolerance of errors

These four categories are not separated but interweave with each other. Low English proficiency can lead to unwillingness to communicate and deteriorate the lack of experience in speaking English. A lack of self-confidence occurs when students lack experience, and this can lead to unwillingness to communicate. Thus, it is not any single factor causes FLSA, but all four categories and the factors within each of them (see Figure 2-4) cause FLSA among Chinese EFL learners.

Figure 2- 4 *My own interpretation of relationship of four categories of causes of Chinese EFL learners' FLSA*



The four groups of factors shown in Figure 2-4 may not cover all the sources of FLSA, nor do they summarise all the reasons for FLSA among Chinese EFL learners. However, it serves as a framework of causes of FLSA guiding the design of FLSA measurement (see Section 4.2.4) in my study.

2.4 Effects of FLSA

In order to understand the importance of alleviating FLSA, it is essential to reveal the adverse effects it produces. This section discusses the academic, psychological, and social effects of FLSA.

2.4.1 Academic Effects

The effects of foreign language anxiety on foreign language performance have been studied since the 1960s, and the results were inconsistent at the early stage. The inconsistency of results was largely resolved once researchers used more specific methods to measure foreign language anxiety. This allowed the negative correlation between foreign language anxiety

and foreign language achievement to be observed (Teimouri, Goetze, & Plonsky, 2019). A higher level of foreign language anxiety is associated with lower course grades (Ganschow & Sparks, 1996; MacIntyre *et al.*, 1998), more impaired performance (Gardner, 1987), lower ratings of achievement by teachers (Trylong, 1988), and lower student self-rating of their language proficiency (Gardner & MacIntyre, 1993). Since the focus of my thesis is on FLSA rather than foreign language anxiety, these general academic effects of foreign language anxiety are not examined in detail; the effects of FLSA on specific oral performance will be illustrated, however.

It was generally revealed that FLSA seems to have detrimental effects on students' oral performance, although various measurements of FLSA and different oral tasks were utilised in previous studies. Phillips (1992) conducted pioneering research on the effects of language anxiety on specific oral performance variables in French learners. The relationship was examined between eight variables (percentage of total words in communication units[7] (CUs), average length of CUs, percentage of error-free CUs, percentage of words in error-free CUs, percentage of total words in mazes[8], average length of mazes, number of target structures[9] and number of dependent clauses) on the oral exam and language anxiety measured by Foreign Language Classroom Anxiety Scale (introduced in Section 2.2.1.1). It was found that non-anxious students produced more dependent clauses and longer CUs than anxious students.

As part of a replication of the study by Phillips (1992), Hewitt and Stephenson (2012) confirmed certain aspects of Phillips' findings but observed some contradictory findings. For example, they found that higher anxious student produced longer communication units but of poorer quality. Meanwhile, FLSA measured by the Foreign Language Classroom Anxiety Scale and oral exam scores in certain aspects of grammar, vocabulary, pronunciation, and fluency were moderately negatively correlated.

Following the findings of the above two studies, Pérez Castillejo (2019) investigated the specific relationship between FLSA and utterance fluency among Spanish learners.

[7] Communication unit refers to an utterance that cannot be divided without losing its essential meaning or a dependent clause that is part of an independent predication.
[8] A maze is a word, or several words, or a fragment of a word that is extraneous, incorrect, or in the learner's native language. It refers to any filled pause, false start, repetition, or reformulation and is marked by enclosing that part of the utterance in parentheses.
[9] Target structure refers to a usable piece of language, which could be a single word or a string of words.

Utterance fluency refers to the acoustic characteristics of performance in language speaking, such as speed and pauses, which reflect the operation of underlying cognitive processes (Segalowitz, 2010). Segalowitz (2010) additionally applied a proficiency assessment, Spanish EIT, which was useful for the focus on utterance and cognitive fluency. Pérez Castillejo (2019) found that foreign language anxiety was negatively correlated with the following aspects of utterance fluency in a foreign language:

1. Ratio of mid-AS pauses: pause frequency within AS-units[10];
2. Length of end-AS pauses: pause length between AS-units;
3. Mean length of run: the number of meaningful syllables produced between pauses;
4. Phonation-time ratio: percentage of times spent speaking, calculated as the total speaking time (excluding pauses) divided by the total duration of a participant's performance.

Other researchers (Salem & Al Dyiar, 2014; Sanaei, Zafarghandi, & Sabet, 2015) also have affirmed that oral fluency may predict speaking anxiety, though the details of the variables were not specified. Despite that some fluency variables are found to be correlated to speaking anxiety, it is not equivalent to the measurement of FLSA. Therefore, utterance fluency was not investigated in my study.

In addition, FLSA levels were found to be negatively correlated with oral performance that was evaluated subjectively rather than objectively. In the study of Woodrow (2006), the Chinese EFL learners' oral performance was assessed by the examiners from the levels of A to F, and their oral performance was found to be negatively correlated with both in-class anxiety (e.g. role-play in front of the class) and out-of-class anxiety (e.g. speaking with native speakers). Liu (2018) also found a negative relationship between FLSA and oral performance. The oral performance was subjectively rated from 1 to 20 by teachers in the study of Liu (2018). Based on the results obtained by Hamamorad (2020), there was a negative statistically significant correlation between communicative performance and

[10] As a type of Analysis of Speech unit, an AS-unit is mainly a syntactic unit consisting of an independent clause or sub-clause, with any subordinate (dependent) clause(s) attached to either of them (Foster, Tonkyn, & Wigglesworth, 2000).

anxiety as measured by Foreign Language Classroom Anxiety Scale. In the study of Hamamorad (2020), communicative competence was marked on a scale of 1 to 5.

Last but not least, some researchers questioned the existence of foreign language anxiety independent of language achievement, arguing that foreign language anxiety stemmed from poor language learning achievement (Sparks & Ganschow, 1991). MacIntyre (2017) asserted that anxiety was both a cause and a consequence of poor performance. As an example, anxious students may have English learning problems and attain low grades; meanwhile, they may have difficulties in classroom learning and exhibit poor developmental skills (Chen & Chang, 2004). In my view, FLSA is both a consequence and a cause of poor oral performance; students who are anxious about speaking a foreign language may performance poorly on oral tasks, and this poor oral performance may increase their anxiety about speaking. This means it is vital to find a way to improve students' oral proficiency and reduce their anxiety about speaking.

2.4.2 Cognitive Effects

Some researchers revealed a relationship between language anxiety and cognitive process involving language learning. The concept of 'cognitive processes' refers to the tasks the brain conducts during the process of receiving information, such as acquiring, storing, interpreting, manipulating, transforming, and using knowledge ("Cognitive Process," n.d.). From a psychological perspective, Eysenck (1979) proposed that anxiety reduces processing effectiveness, so worry and other task-irrelevant cognitive activities associated with anxiety would compete with task-relevant information for processing space.

From an educational psychology perspective, Tobias (1979) proposed a model with three classic information-processing components in the instructional process: input, processing and output. Input refers to the stage of presenting instructional materials to students; processing consists of the registration, recording, organisation, storage, and retrieval of information; and output denotes students' production of their own learning materials. This model was adapted to discuss foreign language anxiety and foreign language learning by MacIntyre and Gardner (1991b). They found that anxiety might prevent the initial reaction to a given stimulus in memory at the input stage. In other words, anxious students are more likely to miss the incoming information or remember less information (MacIntyre & Gardner, 1994). During the processing stage, anxiety impairs the speed and accuracy of organising,

storing and assimilating information (MacIntyre & Gardner, 1994). The output stage might be disrupted by the retrieval of information, which then affects the quality of second language communication, such as improper use of grammar rules, ineffective retrieval of vocabulary, or even inability to respond at all (MacIntyre & Gardner, 1994). Indeed, anxious learners tend to freeze up in oral activities and forget what they have learnt (Ely, 1986; Steinberg & Horwitz, 1986).

Additionally, anxious learners may not have the ability to respond effectively in response to their errors but tend to think negatively about their substandard performance (Gregersen, 2003). Gregersen and Horwitz (2002) stated that foreign language anxiety and perfectionism manifest similarly in anxious learners that they set higher standards for their performance. Moreover, when students receive feedback, less anxious learners tend to be more effectively in learning as being more aware of differences between their utterances and target-like forms, or to be more successful in perceiving recasts and metalinguistic feedback than anxious students (Rassaei, 2015). Recast feedback refers to "the teacher repeats the students' utterance in the correct form without pointing out the students' error" (Martin & Valdivia, 2017, p. 4); this is similar to modelling discussed above. Metalinguistic feedback means "the teacher gives a hint or a clue without specifically pointing out the mistake" (Martin & Valdivia, 2017, p. 4). Thus, from a cognitive perspective, it is important to alleviate FLSA in order to enhance foreign language learning.

2.4.3 Social Effects

From a social perspective, high-level anxious students are not willing to communicate with others (Price, 1988). Horwitz (2010) noted that language anxiety causes students to avoid communicating in English. It is generally believed that this type of unwillingness is mainly related to the student's self-concept in the foreign language. Learners' self-perceptions of foreign language competence or proficiency is often seen as a predictor for foreign language anxiety (Cheng, Horwitz, & Schallert, 1999; MacIntyre, 2007). This is in accordance with cognitive self-evaluation theory, which states that people become anxious in social activities not because they lack related skills, but because they believe they lack them (Leary & Kowalski, 1997). The discussion of self-esteem/self-concept as a factor related to FLSA can be found in Section 2.3.1.1.

Evidence for FLSA as a cause of unwillingness to communicate can be found in some previous empirical studies. As Bailey (1983) reported, anxious students perceived their language skills as weaker than others. Similarly, Horwitz (2010) contended that foreign language learners were fearful of not meeting the expectations of their peers or significant others, which led to unwillingness to speak the foreign language. The anxious students were relatively silent when they had to communicate with others or simply speak the foreign language (Liu & Jackson, 2008). Further, Jee (2018) found a moderate correlation between foreign language anxiety and unwillingness to communicate among Korean students.

In addition, in foreign language classes, anxious students were less likely to take risks and socialise with others (Liu & Jackson, 2008; Pyun et al., 2014; Saito & Samimy, 1996). Risk-taking is an individual's tendency to assume risks when using a foreign language (Dittmann, 2018). The chances of the anxious students voluntarily answering questions or participating in oral classroom activities were lower (Ely, 1986). In speaking the foreign language, anxious students may be more reluctant to use linguistic forms they are not so confident with (MacIntyre & Gardner, 1991a), or practise new grammatical structures and use improvisation (Young, 1991). The negative correlation between risk-taking and foreign language anxiety was found to be statistically significant in some previous studies (Falkoni-Mjehović, Tolj, & Baranac, 2022; Jee, 2018). When it comes to foreign language speaking, risk-takers are more active and productive, they seize every opportunity to use the foreign language (Lin & Lin, 2020).

In conclusion, FLSA may negatively impact oral performance, inhibit cognition processes, and discourage willingness to communicate and risk-taking. It is therefore necessary to find ways to reduce FLSA levels in order to improve foreign language learning and use.

2.5 Approaches for Alleviating FLSA

The adverse effects of FLSA outlined above have led to several suggestions of approaches to help foreign language learners cope with them. However, only a few of these suggestions have been empirically tested. This section introduces some of the approaches previous researchers have suggested for reducing FLSA levels and presents an up-to-date review of relevant empirical studies on them.

2.5.1 Proposed Approaches

Scholars have proposed methods for coping with language anxiety among language learners based on their reflections on factors associated with foreign language anxiety in the early stage. This can be divided into two categories: the student's level and the instructor's level. Although these approaches are mainly intended to deal with anxiety in foreign language learning, they have potential for FLSA alleviation because most of the discussion of foreign language anxiety in the existing literature related to speaking the foreign language.

From the students' perspective, there are a number of ways to reduce foreign language anxiety levels, such as having realistic expectations and beliefs about language learning (Crookall & Oxford, 1991), and realising that many variables (e.g. attitude, motivation, anxiety, learning style, tolerance for ambiguity and risk-taking) are associated with language learning rather than aptitude (Phillips, 1999). It is important to help students change their beliefs about foreign language learning in order to alleviate foreign language anxiety as well as FLSA. Foreign language learners should be aware that language learning can take a considerable amount of time (Crookall & Oxford, 1991), and that making mistakes is not a catastrophic problem in language use (Gregersen, 2003).

The above student approaches to reducing FLSA are highly correlated with instructors. Without the assistance of teachers, students may not be able to overcome their anxiety. If teachers create a supportive and constructive classroom atmosphere, it may decrease students' foreign language anxiety levels. Palacios (1999) argued that a less competitive, warm, and easy-going classroom climate may help alleviate students' anxiety. It is also stated that students feel less anxious when their teachers help them, communicate openly with them with trust, and show interest in their ideas (Horwitz, Horwitz, & Cope, 1986; Trickett & Moos, 1995). Through low-risk self-disclosure activities, students can learn information about each other. Teachers may also become more accepting of their students, which in turn may result in a reduction in anxiety levels in class (Dörnyei & Murphey, 2003). Accepting students refers to non-judgemental and positive attitudes to students (King & Smith, 2017).

Additionally, if instructors adopt non-harsh, friendly and relaxing error correction techniques, this may decrease students' anxiety levels (Young, 1990). It has been revealed that correcting every error and overcorrecting may demotivate students and result in them focusing on form rather than meaning in their language learning (Gregersen, 2003).

Although this may be rare in current Western education systems, it is still common in examination-oriented Chinese EFL classrooms (Su, 2019). As Walz (1982) suggested, when the goal of using language is communication, correcting errors that cause misunderstanding should take priority over correcting other errors, and students should correct their own oral errors first. This is difficult to achieve in China, where teachers focus on correcting students' grammatical errors and misuse of vocabulary (Su, 2019).

Phillips (1999) recommended that teachers should continually remind students making mistakes is natural in the foreign language learning process. However, mistakes are not tolerated in Chinese EFL classrooms, especially in grammar (Su, 2019). It is advised by Phillips (1999) that corrections of form should be limited to mechanical practice; if corrections are used for authentic communication, it may be entirely inappropriate. Phillips (1999) proposed a technique named modelling, in which teachers respond to students' statements, while at the same time modelling the correct form or vocabulary. The following is an example of modelling, in which an indirect correction is used.

> "Teacher: What did you do last night, Susan?
> Susan: I goed to the movie *Captain Hook*.
> Teacher: Oh, you went to the movie. I went to see *Captain Hook*, too. What a coincidence! We went to the same movie." (Phillips, 1999, p. 129).

In an online foreign language learning environment, Martin and Valdivia (2017) found that anxious students prefer recast and metalinguistic feedback (see Section 2.4.2 for explanation) from teachers, and that error corrections were necessary. Therefore, the techniques to correct students' errors should be considered in order to reduce students' FLSA levels.

In addition, learning tasks may have an influence on students' foreign language anxiety levels. In terms of foreign language learning, a task refers to an activity focused on meaning that students need to complete with their linguistic knowledge (Ellis, 2003). In dealing with foreign language anxiety, pair-work and personalised discussions (Koch & Terrell, 1991), as well as small-group speaking activities (Young, 1990), were found to be more acceptable than other tasks (e.g. presenting in front of the whole class). By creating a community of learners who support each other and work collaboratively to produce meaningful utterances, it might be possible to improve authentic communication and help with reducing foreign language anxiety levels (Larrivee, 2005; Phillips, 1999).

Phillips (1999) recommended some tasks for groups and pairs to alleviate anxiety in language learning and use, including *recognition activity, cued response (pair work), information gap activities, interviews, and surveys (individual), cartoon stories and role play*. Considering the affordances of High-immersion Virtual Reality (HiVR) technology (see Section 3.2), *information gap activities* and *role-play* were utilised in my learning design. In the information gap task, each person receives a part of the information needed to solve a puzzle, and students solve the puzzle by exchanging information. For the role-play task, a situation is set in which students taking their characters of the task. The detailed learning activity design is presented in Section 5.3.3.

2.5.2 Empirical Studies on Approaches to Alleviating FLSA

In recent years, researchers have begun to investigate the effectiveness of different learning activities on reducing students' FLSA levels. Cooperative learning seems to be beneficial for FLSA alleviation. The reason for this may be that students are less concerned with errors in linguistic errors or about failing the task because anxiety levels are distributed at the group level rather than at the individual level (Tsiplakides & Keramida, 2009).

An intervention involving pair-work and small-group work were employed in the study of Nagahashi (2007) among 38 EFL learners to investigate FLSA. The cooperative learning activities included comparing homework answers, quizzing each other on new vocabulary, dictating model paragraphs, reading, and commenting on each other's essays, and reading texts aloud and so on. The majority of participants (more than 80%) in the study of Nagahashi (2007) had a positive perception of the cooperative learning activities.

A group work (*project work*) combining building *a supportive classroom atmosphere* was designed by Tsiplakides and Keramida (2009) to investigate their effects on reducing FLSA. The supportive atmosphere was established by promoting teacher student relations, providing indirect corrections, protecting students' self-worth, teacher's positive verbal and non-verbal immediacy, and non-verbal praise. Despite the absence of statistical comparisons of FLSA levels before and after the intervention, Tsiplakides and Keramida (2009) found that students were more willing to participate in speaking tasks after the interventions.

Yalçın and İnceçay (2014) found the same positive results in their study of spontaneous collaborative speaking activities (games, role-plays, and debates) and FLSA levels in 12

students from the English teaching department. It was stressed during the intervention that the importance of being unprepared and feeling that you were successful were important in alleviating FLSA. This may be because that without preparation of the tasks can arouse anxiety, and the successful experience of completing the tasks contribute to the build of self-confidence. This is in line with the point of view stated in Section 2.3.2.1 that poor experiences can result in FLSA, but pleasant or successful experiences may help to reduce it.

The benefit of cooperative speaking activities on alleviating FLSA was also illuminated by Mon (2019). The learning activities included *finding someone who* (finding a person with a certain characteristic), *a cup of conversation* (pick questions written on a slip of paper contained in a cup to answer), *doctor's appointment* (role-play), *charade* (guessing the word acting by others), and *debate* (argue for different points on a topic). It was further found that having a humorous, knowledgeable, patient and encouraging teacher reduced students' anxiety and prevented them from worrying about making mistakes while speaking English (Mon, 2019).

Additionally, fluency-based activities may contribute to lowering FLSA levels, since students' errors are tolerated until the communication is complete and may be implicitly corrected (Hammad & Ghali, 2015). It is consistent with the finding by Ansari (2015) that it is necessary to convince anxious students to continue speaking even though mistakes are made. Moreover, Ansari (2015) argued that the activities should reflect students' interests and be appropriate for their proficiency levels. Therefore, the learning design in my study took these suggestions into consideration; the priority of the intervention was to provide students an interesting and relaxing environment to complete the tasks via communicating with each other, rather than correcting their mistakes. Details of the learning design can be found in Section 5.3.3.

A positive psychology approach was examined by Jin, Dewaele, and MacIntyre (2021) using an experimental design. They found that reminiscing about language achievement could significantly reduce foreign language anxiety. Positive psychology approach involves boosting positive emotions in foreign language learning in order to reduce negative emotions. Several researchers (Dewaele *et al.*, 2019; Dewaele & Li, 2020) advocate positive psychology when dealing with foreign language anxiety, since it can enhance students'

ability to notice in class, strengthen students' awareness of language input, and reduce the negative emotional effects of learning foreign language, such as anxiety (Gregersen, 2013). It is important to note that the concept of foreign language enjoyment (Dewaele *et al.*, 2016; Dewaele *et al.*, 2018) is presented in a sequential manner based on a positive psychology approach. However, the discussion of foreign language enjoyment is excluded from this thesis, because it is a completely different topic that would require major investigation. Future studies could investigate whether and how the use of HiVR technology influences foreign language enjoyment. Although positive psychology was not directly applied to my thesis, the purpose of my learning design in HiVR was to maximise the creation of positive emotions during the foreign language speaking process.

2.5.3 Approaches to Alleviating FLSA Among Chinese EFL Learners

Research on coping with FLSA among Chinese EFL learners seems to be in the theoretical stage rather than the empirical stage. The 12 solutions suggested by participants in the study of He (2018) provided directions for reducing Chinese EFL learners' foreign language anxiety, which can be categorised as follows:

1. Teachers being humorous, patient, encouraging;
2. Speaking English in a friendly environment with indirect error correction and not focusing on accuracy;
3. Being aware that making mistakes is common;
4. Joining small-group activities;
5. Playing English games and doing relaxation exercises;
6. Talking about anxiety about speaking English.

Other pedagogical suggestions were made by Jiang and Dewaele (2019) and Sun (2022) based on their investigations of Chinese EFL learners in universities. They emphasised the need for teachers to use humour and encouragement to create a relaxing and positive classroom atmosphere. Teachers should learn more about EFL learning anxiety and pay attention to students' mental states. Curriculum and syllabus should be more creative so that teachers can use innovative teaching methods rather than follow strict guidelines. Both teachers and students may find these suggestions challenging, since classroom atmosphere in China is often serious, not relaxed or open, with students sitting and listening to their teachers (Lim, 2007).

Considering that China has a very large number of English learners but few high-quality teachers as well as a strict examination-oriented English learning curriculum, creating a relaxed classroom atmosphere, implementing group/pair work, and applying appropriate ways to correct errors will not be easy. It is therefore worthwhile to investigate alternative strategies for reducing FLSA among Chinese EFL learners. The integration of technology into the foreign language learning process seems to be a potential solution. However, the impact of technology on FLSA has received little attention, and the results of existing studies are mixed (Aydın, 2018). My study aims to fill this research gap by investigating the effects of HiVR technology on FLSA. The theoretical foundations and empirical evidence for conducting this research are illustrated in the following chapters.

2.6 Summary

This chapter examined the concept which drives my thesis – FLSA – including its definition (Section 2.1), measures (Section 2.2), causes (Section 2.3) and effects (Section 2.4). While some traditional solutions (e.g. collaborative tasks) have been found to be effective in alleviating FLSA, there have been obstacles to implementing them in Chinese EFL classrooms (see Section 2.5). In the following chapter, I will discuss the potential effects of High-immersion Virtual Reality technology on FLSA in an effort to gain a better understanding of my study's rationale and identify potential research gaps.

Chapter 3. LITERATURE REVIEW: HIGH-IMMERSION VIRTUAL REALITY AND FOREIGN LANGUAGE SPEAKING ANXIETY

As stated at the end of Chapter 2, involving technology may be a possible way to cope with Foreign Language Speaking Anxiety (FLSA) among Chinese English-as-a-Foreign-Language (EFL) learners. This chapter opens with an introduction to High-immersion Virtual Reality (HiVR) and provides an overview of its main affordances in learning. The potential and challenges of HiVR in foreign language learning are discussed in order to identify the research gap. The evaluation of existing empirical studies on HiVR and FLSA are presented at the end of this chapter in order to gain a substantive understanding of this topic.

3.1 Introduction to VR

3.1.1 Definitions and Classifications of VR

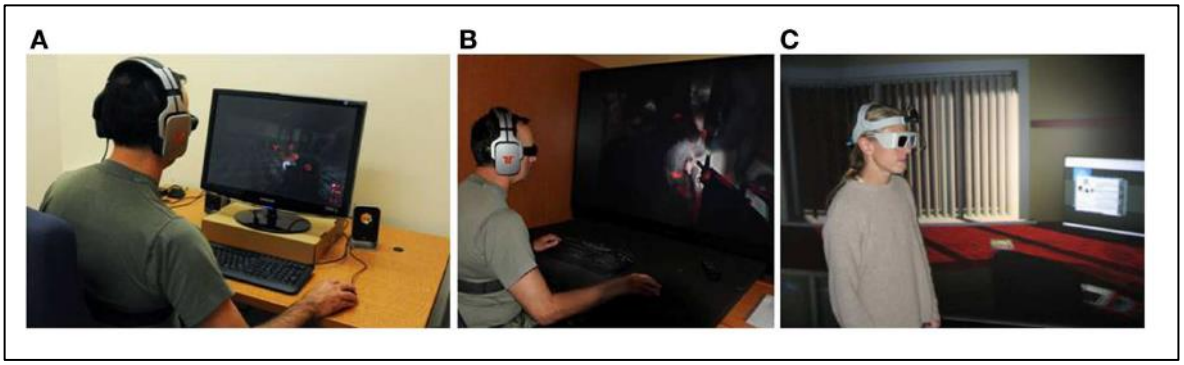
Since the term Virtual Reality (VR) coined by Jaron Lanier in 1989, its definition has constantly evolved. In the past, the definitions of VR were usually complex (Kaplan-Rakowski & Gruber, 2019). According to Coates (1992), VR is "electronic simulations of environments experienced via head-mounted eye goggles and wired clothing enabling the end user to interact in realistic three-dimensional situations" (cited in Steuer, 1992, p. 74). Greenbaum (1992, p. 58) also mentioned the use of goggles and data gloves in his definition as follows:

> "Virtual Reality is an alternate world filled with computer-generated images that respond to human movements. These simulated environments are usually visited with the aid of an expensive data suite which features stereophonic video goggles and fibre-optic data gloves".

These early VR definitions relate more to immersive VR or High-immersion VR, which is the focus of my study. Recent definitions of VR have become simpler, excluding the use of goggles or gloves but expanding to all simulations presented through the use of a display. Serrano *et al.* (2013, p. 648) defined VR as "a computer-based technology that creates a synthetic reality using three-dimensional graphics". Similarly, Girvan (2018, p. 1098) stated that VR "is a technical system through which a user or multiple users can experience a simulated environment".

A detailed division of VR categories is necessary because the late VR definitions expand to simulations presented in various displays. VR technology can generally be divided into three categories based on the degree to which it displaces physical reality: non-immersive VR, semi-immersive VR, and immersive VR (Carrier, Damerow, & Bailey, 2017). Examples of the three types of VR can be found in the paper of Baus and Bouchard (2010, p. 3) as shown in Figure 3-1.

Figure 3- 1 *Examples of VR: (A) non-immersive VR, (B) semi-immersive VR, and (C) immersive VR (Baus & Bouchard, 2010, p. 3)*



Non-immersive VR (see A in Figure 3-1), often refers to Virtual Worlds, a desktop-based three-dimensional (3D) virtual environment that can be shared by multiple users simultaneously online (Dickey, 2005). A non-immersive VR environment is displayed on the computer monitor, the users are represented by avatars[11], and interaction with the virtual environment is carried out using a keyboard, mouse, joystick or touch screen (Lee & Wong, 2014). In semi-immersive VR (see B in Figure 3-1), a large concave screen, a projection system and a computer are often required (Carrier, Damerow, & Bailey, 2017). The Cave Automatic Virtual Environment is a common type of semi-immersive VR where projectors are directed to between three and six walls of a room-sized cube, and users' physical movement can be captured. However, real-world surroundings are still visible in semi-immersive VR.

Through immersive VR (see C in Figure 3-1), users are able to experience a lifelike virtual environment by immersing themselves fully within the virtual environment through a head-mounted display and interacting with it directly through gloves or motion controllers (Makransky & Lilleholt, 2018). A head-mounted display is the primary component of VR

[11] Avatars are virtual characters used to present the users in a virtual environment instead of their real appearance.

headsets, which consists of an embedded liquid crystal display that provides 360-degree video images from slightly different angles for each eye (Hamilton *et al.*, 2021). In immersive VR, motion controllers or gloves replaces the mouse, making interaction with the virtual environment more realistic and authentic (Peixoto *et al.*, 2021). Immersive VR is more closely related to early definitions of VR proposed by Coates (1992) and Greenbaum (1992).

The classification of VR in the 2020s differs, particularly in the field of foreign language learning (Kaplan-Rakowski & Gruber, 2019). The concept of immersion level was used to distinguish the types of VR rather than the accessibility to physical reality. Based on this, scholars (Kaplan-Rakowski & Gruber, 2019; Makransky, Terkildsen, & Mayer, 2019) divided VR into two dimensions: Low-immersion VR and High-immersion VR. Low-immersion VR (LiVR) is referred to as desktop VR, in which a virtual 3D environment is viewed on a two-dimensional computer monitor (Kaplan-Rakowski & Gruber, 2019). In High-immersion VR (HiVR), users are typically required to wear a head-mounted display in order to perceive a 3D virtual environment as spatially realistic, a result of disconnecting themselves from the real world (Kaplan-Rakowski & Gruber, 2019). In a recent article by Peixoto *et al.* (2021), a definition of HiVR was provided. This definition stated that HiVR involves a simulation of a three-dimensional virtual environment that is generated by a computer and in which a person wears a helmet equipped with a built-in screen to move around, observe, and interact with virtual objects.

In the past two decades, a number of researchers (Arnold, 2007; Hammick & Lee, 2014; Kruk, 2016, 2019; Reinders & Wattana, 2015; Wehner, Gump, & Downey, 2011) have demonstrated the benefits of LiVR in alleviating FLSA. However, the purpose of my study is to fill the research gap regarding the impact of HiVR rather than LiVR on FLSA. In the following two sections, a brief history of HiVR technology and the types of current HiVR devices are provided.

3.1.2 Brief History of HiVR

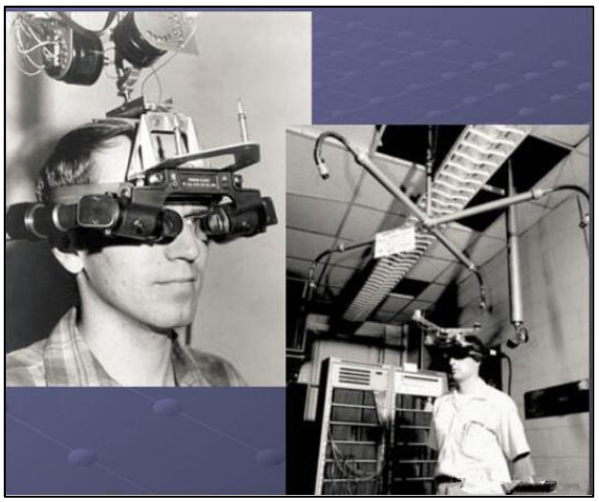
HiVR technology can be traced back to the development of 360-degree art through panoramic murals appearing in the 19th century, while the first immersive and multi-sensory mechanical device was created by Morton Heilig in 1962, named *Sensorama* (see A in Figure 3-2). *Sensorama* simulated a motorcycle ride with colourful display of streets through the

screen, fan-generated wind, sounds and smells of the city, as well as motion from the chair (Freina & Ott, 2015). As early as 1968, Ivan Sutherland and his students at MIT's Lincoln Laboratory developed *Sword of Damocles* (see B in Figure 3-2), the first VR head-mounted display. Although this VR head-mounted display was so heavy that it had to be suspended from the ceiling, headtracking became possible within it.

Figure 3- 2 *Previous VR devices: (A) Sensorama[12], (B) Sword of Damocles[13], (C) VPL VR Products[14], (D) Cyberface[15]*

A

B



[12] Retrieved from http://uschefnerarchive.com/morton-heilig-inventor-vr/
[13] Adapted from Sutherland (1968)
[14] Retrieved from https://commons.wikimedia.org/wiki/File:VPL_DataSuit_1.jpg
[15] Retrieved from http://www.leepvr.com/cyberface1.php

VR devices developed in the 1980s were lighter and easier to wear. Among the first companies to develop and sell VR products was Visual Programming Lab[16], founded by Jaron Lanier in 1984 (Berg & Vance, 2017). They invented the DataGlove, EyePhone, and DataSuit for VR systems (see C in Figure 3-2). As the first head-mounted display to be designed for commercial use, *Cyberface* (see D in Figure 3-2) was based on wide-angle LEEP viewing lenses. Since the 1990s, HiVR technology has transformed step by step that enabling it to become smaller, lighter, and cheaper in systems such as Sega VR and the Matrix. In 2016, HiVR technology was accelerated by the development of hardware and software by many high-tech companies, including Oculus, HTC, Google, and Microsoft (Hilfert & König, 2016). Since HiVR devices had become more easily accessible, and the application of HiVR to foreign language learning was underexplored, I chose to investigate HiVR rather than LiVR when I began my PhD in 2018.

3.1.3 Existing Types of HiVR

According to Radianti *et al.* (2020), HiVR technology can be divided into three areas: mobile HiVR (used with a mobile phone), high-end HiVR (connected to a computer), and enhanced HiVR (a combination of head-mounted displays and data gloves or bodysuits). However, I have adopted the categories proposed by Hamilton (2021) in my study, as they appear to be more clearly defined. Hamilton (2021) distinguishes three types of VR headset: tethered (e.g. Oculus Rift, HTC Vive), stand-alone (e.g. Oculus Quest, HTC Vive Focus), and mobile (e.g. Google Cardboard, Samsung Gear). The stand-alone VR headset bridges the gap between mobile and tethered VR headsets through the use of a headset that includes a built-in screen and battery as well as one or two motion controllers (Carruth, 2017). Examples of different types of HiVR devices are shown in Figure 3-3.

Each type of HiVR device has advantages and disadvantages. Generally, tethered HiVR provides the highest level of immersion and comes with dedicated controllers for game playing, but it is often heavy and difficult to use (Newbutt, Bradley, & Conley, 2019). In spite of the fact that mobile HiVR is generally low-cost, its interaction with the environment is difficult to achieve, and its immersion level is low (Newbutt, Bradley, & Conley, 2019). A stand-alone HiVR device can, however, provide a better immersion experience than

[16] VPL filed for bankruptcy and its patents were bought by Sun Microsystems.

mobile HiVR using the head-mounted display, and is more convenient and easier to use than tethered VR headsets due to its simpler system and lighter headset.

Figure 3- 3 *Examples of HiVR devices: (A) tethered HiVR[17], (B) stand-alone HiVR[18], and (C) mobile HiVR[19]*

A B C

Similar use of HiVR devices can be found in Educational research (Dhimolea, Kaplan-Rakowski, & Lin, 2022; Peixoto *et al.*, 2021). Research on high-end HiVR (or tethered VR headsets) appears to be the most common (nearly 50% of studies), followed by mobile HiVR (around 25%), but stand-alone HiVR has caught the attention of few scholars. It is possible that stand-alone devices were developed later than the other two types of HiVR device (tethered and mobile), which have been available since 2019. When I was preparing for my fieldwork, stand-alone HiVR had just begun to gain popularity. Due to the limited empirical research of stand-alone HiVR in the field of foreign language learning as well as its convenience for educational use with smaller and lighter hardware, I chose the stand-alone HiVR device for my research. In Chapter 5, I provide a more detailed description of the HiVR device used in my research.

3.2 Main Affordances of HiVR in Learning

HiVR technology has distinct affordances that may be beneficial for foreign language learning as well as for alleviating FLSA. The concept of technology affordance refers to the property of a technology to provide potential for practice (e.g. foreign language learning and use in my study) (Bobsin, Petrini, & Pozzebon, 2019). In this section, a discussion is provided of the two most important affordances of HiVR in learning: presence and agency.

[17] HTC Vive Pro, retrieved from https://www.vive.com/eu/product/vive-pro/
[18] HTC Vive Focus Plus, retrieved from https://enterprise.vive.com/ca/product/focus-plus/
[19] Google Cardboard viewer 2nd generation, retrieved from https://en.wikipedia.org/wiki/Google_Cardboard

3.2.1 Presence

In terms of HiVR, presence often refers to telepresence, which can be roughly understood as one's perception of *being there*, even when one is not physically there (Riva, Davide, & IJsselsteijn, 2003). Presence is a psychological state or a subjective experience that an individual overlook the technology at some levels while using the technology (Schuemie *et al.*, 2001). Different individuals may therefore experience different amounts of presence, depending on how individual perceive the experience (Makransky & Petersen, 2021). An important determinant of presence is related to how realistically the virtual environment is displayed and how smooth the view changes in the environment are (Dalgarno & Lee, 2010). With the head tracking function built in the head-mounted displays of HiVR, users are able to view their digital surroundings by turning their heads as they do in the real world (Dhimolea, Kaplan-Rakowski, & Lin, 2022). This seems to contribute to a higher level of sense of presence in HiVR than in other virtual environments (Petersen, Petkakis, & Makransky, 2022). In HiVR, the sense of presence occurs when the brain and nervous system react to the virtual environment as if it were a real-world situation (Slater, 2003).

The immersive features of HiVR make a major contribution to the sense of presence. Immersion, from a technological perspective, can be defined as "the extent to which the computer displays are capable of delivering an inclusive, extensive, surrounding, and vivid illusion of reality" (Slater & Wilbur, 1997, pp. 604-605). In contrast to presence as a subjective phenomenon, immersion is an objective description of a system (Slater & Wilbur, 1997). Through the use of head-mounted displays within HiVR, the user is able to experience a high level of immersion in the virtual environment, allowing them to perceive the virtual environment from a first-person perspective and receive sounds or other stimuli without distraction from the outside world (Freina & Ott, 2015).

The sense of presence in HiVR may contribute to the alleviation of FLSA because HiVR can provide an immersive environment in which the foreign language learners can speak the target language. As stated in Section 2.3.2, the more foreign language learners practise authentic communication, the less anxious they may feel while speaking the foreign language (Dewaele, 2008; Mede & Karaırmak, 2017). Although the target-language speaking environment in HiVR is a virtual environment, it may have potential for alleviating

FLSA, because foreign language learners using HiVR can experience real-life scenarios from a first-person perspective, as though they were acting in the real world.

3.2.2 Agency

The concept of agency is another important learning affordance of HiVR. Agency refers to a person's perception of generating and controlling their actions (Moore & Fletcher, 2012). Specifically, the sense of agency is related to the freedom of controlling one's actions in the virtual environment (Johnson-Glenberg, 2019). When the predicted sensory consequences of a given movement are related to the actual sensory feedback signals arising from the movement, a sense of agency arises in the central nervous system (Farrer *et al.*, 2008).

According to Petersen, Petkakis, and Makransky (2022), interactivity in HiVR technology contributes to a sense of agency because users are able to determine their course of action through interacting with the virtual environment. Interactivity refers to "the extent to which users can participate in modifying the form and content of a mediated environment in real time" (Steuer, 1992, p. 84). In HiVR, the location, the looking point, and the time to perform each action are all controlled by the users themselves, which contributes to high agency (Taguchi, 2022). Compared to other types of VR, Velev and Zlateva (2017) stated that interactivity within HiVR is more realistic and natural because of the direct interaction with the virtual environment through the use of sensor controllers or gloves.

The sense of agency experienced while using HiVR may be pleasurable for foreign language learners because they are engaged in the virtual environment and can control it. This may help to alleviate FLSA because it enhances foreign language learners' motivation to learn the language and provide a pleasant experience of using the language. Empirical evidence of utilising HiVR in foreign language learning will be discussed in the following section.

3.3 HiVR and Foreign Language Learning

With the reduction in cost and developments in technology of HiVR, more researchers are investigating its application in the field of foreign language learning, but the area is still under-researched (Peixoto *et al.*, 2021). The potential of HiVR in foreign language learning, especially for alleviating FLSA, is examined in this section – as are the challenges.

3.3.1 Potential of Using HiVR in Foreign Language Learning

Potential to increase experience

The development of foreign language learning platforms in HiVR may help to increase foreign language learners' frequency of speaking the target language in an authentic scenario. This has the potential to alleviate FLSA, because learner's FLSA may decrease as their experience in the target language increases (Dewaele, 2008). It is claimed by Onwuegbuzie, Bailey, and Daley (1999) that exposure to the target-language environment is an ideal method for improving foreign language learners' confidence in speaking the foreign language. However, such practical experience is not possible for every foreign language learner due to the high cost and time commitment required. HiVR might provide a solution to this issue, because it can simulate the real world, thus allowing foreign language learners to sit in on authentic communications (Peixoto *et al.*, 2021). When using HiVR for foreign language learning, daily-life locations are often created, such as zoos, cafés, airports, cinemas and museums (Alfadil, 2020), hotels and restaurants (Yang *et al.*, 2020), as well as supermarkets, hospitals, and schools (Wang, Lian, *et al.*, 2021). Based on the affordance of presence (see Section 3.2.1), students using HiVR can experience the feeling of being in target-language scenarios without the expense of a trip to a foreign country. However, it is underexplored that if the increase of experience in speaking the target language in virtual environment while using HiVR can make a difference to the FLSA levels as the real-life immersion does.

Potential to enhance learning engagement

HiVR has the potential to enhance learning engagement. Learning engagement refers to students' voluntary participation in learning activities (Hu & Hui, 2012). A total of 91% of students in the study of Kaplan-Rakowski and Wojdynski (2018) and 70% of participants in the research of Sally Wu and Alan Hung (2022) reported that their engagement in learning activities was high while using HiVR. The positive effects of HiVR on learning engagement were revealed in studies with specific foci, such as vocabulary learning (Alfadil, 2020; Tai, Chen, & Todd, 2022), communication or interaction (Tseng & Geng, 2021; Yang *et al.*, 2020), and speaking practice (Enkin, 2022).

Enhancing learning engagement may help to alleviate FLSA because it may switch learners' focus from worries about their speaking competence to the learning tasks at hand. I assume that when concentrating on learning activities in HiVR, foreign language learners may feel more relaxed, which in turn enhances the pleasure of learning.

Potential to create a low anxiety environment

It is evident, both quantitatively and qualitatively, that HiVR may provide an environment that offers a low level of anxiety. Using HiVR may improve students' willingness to communicate and their motivation to learn a foreign language. In the study of Ebadi and Ebadijalal (2020), after the intervention of four presentations, willingness to communicate of EFL learners with levels ranging from A1 to C2[20] who used HiVR was statistically significantly higher than that of students who did not use HiVR. The same positive statistical results were reported by Enkin (2022), demonstrating that advanced Spanish-major students found speaking in HiVR to be more enjoyable than speaking face-to-face and felt less self-conscious while doing so.

Several qualitative studies (Awada & Diab, 2018; Chen *et al.*, 2020; Sally Wu & Alan Hung, 2022) indicated that students became more interested in learning English when offered HiVR. The relaxing and safe environment of HiVR boosted participants' confidence (Ebadi & Ebadijalal, 2020), and encouraged more risk-taking (Enkin, 2022). Specifically, Xie, Ryder, and Chen (2019) found that 12 students (10 English speakers, 2 Chinese heritage learners) who used HiVR for oral presentation in Chinese mentioned that they felt more relaxed when delivering the speech within HiVR. This is because the audience's attention was not solely focused on their presentation but could shift to other things in the virtual environment, and extra time and space to relax during the presentations was allowed while using HiVR.

Nonetheless, not all HiVR environment can claim to be low-anxiety. Participants in the study of Sally Wu and Alan Hung (2022) did not experience a positive effect of HiVR on their willingness-to-communicate levels. This may be attributed to the learning approach involved in researchers' learning design: students had to memorise words from scripts, and only used HiVR to learn words and dialogue. The learning affordances of HiVR had not been employed, which resulted in a very basic use of HiVR as a learning modality. It is, thus,

[20] Based on Common European Framework of Reference for language (CEFR).

important to investigate whether different learning approaches in HiVR can affect the degree of FLSA change. Specifically, I compared the influence of two learning approaches on FLSA: teacher-centred learning approach as applied to the study of Sally Wu and Alan Hung (2022) and a learning approach underpinning situated learning theory. Chapter 4 demonstrates the situated learning approach and details of the learning design can be found in Chapter 5.

Potential to improve oral proficiency

In terms of oral proficiency, inconsistent results are revealed with the use of HiVR. Many scholars have reported that vocabulary learning may be improved by using HiVR (Ebadi & Ebadijalal, 2020; Sally Wu & Alan Hung, 2022; Xie, Chen, & Ryder, 2021). According to Chen, Hung, and Yeh (2021), using HiVR with a problem-based learning approach improved students' English vocabulary learning compared to utilising problem-based learning without HiVR. It has to be noted that only studies using HiVR for oral tasks were reviewed here rather than those utilising HiVR for specific vocabulary learning purposes (e.g. Alfadil, 2020; Legault *et al.*, 2019; Tai, Chen, & Todd, 2022; Wilang & Soermphongsuwat, 2018). A case study (Xie, Ryder, & Chen, 2019) showed that while giving a presentation using HiVR, students used a more sophisticated lexicon than usual.

Nevertheless, the results of studies of the effects of HiVR on grammar and oral performance (e.g. fluency, accuracy, and pronunciation) were contradictory. No difference in grammatical knowledge between HiVR and non-HiVR groups was found by Xie, Chen, and Ryder (2021), but significant improvements in grammar ability were revealed in the HiVR group in the study by Sally Wu and Alan Hung (2022). With regard to pronunciation and fluency, no changes were found after using HiVR in previous studies (Ebadi & Ebadijalal, 2020; Sally Wu & Alan Hung, 2022). Furthermore, a slower speech rate in closed role-play tasks was found in HiVR compared to desktop VR by Taguchi (2022). In closed role-play tasks, students responded to a specific situation by reading it from the screen. This differs from the role-play tasks employed in my study that students playing different characters for the scenario to have a conversation (see Section 5.3.3).

It should be noted that a perceived enhancement in speaking skills was reported by 92.7% of participants (N = 124) in the study of Soto *et al.* (2020) using *ImmerseME*[21] for a speaking test, and by 58% of participants (N = 20) in the research of Liaw (2019) using *vTime*[22] for social activities. The perceived proficiency refers to students' own subjective judgement of proficiency, which is different from objectively measured proficiency. Moreover, an interesting result was found by Christoforou, Xerou, and Papadima-Sophocleous (2019): students performed poorly after the intervention using *Mondly*[23], and the post-test scores in the HiVR group were significantly lower than in the control group (learning through conventional lecture). The authors attributed the negative results to the small sample size (N = 18) in their study. However, it seems to be essential to investigate through a qualitative lens of reasons for poorer performance after using HiVR.

3.3.2 Challenges of Using HiVR in Foreign Language Learning

With regard to negative perceptions, feelings of cybersickness and dizziness while using the head-mounted display are mentioned the most. Several participants in the case study of Xie, Ryder, and Chen (2019) reported that they were not comfortable when using the HiVR device, experiencing physical dizziness. Similarly, Sally Wu and Alan Hung (2022) found that wearing the HiVR headset for a long time could cause participants' headache and dizziness. This is in line with findings in the study of Kaplan-Rakowski and Wojdynski (2018) that when the VR content was more dynamic (e.g. cycling), students experienced a higher level of discomfort. This may be due to the perceptual lag occurring with HiVR systems when the virtual environment is too graphically complex (Ryan *et al.*, 2019). The discomfort experience when using head-mounted display should be taken into consideration when using HiVR for foreign language learning and use.

Technical problems are another obstacle to the use of HiVR in foreign language learning. In some cases, the manipulation of the head-mounted display can be distracting to students while they are learning (Urueta & Ogi, 2019). It was difficult for participants in the study by Xie, Ryder, and Chen (2019) to understand and follow the presentation while using HiVR

[21] ImmerseMe is an HiVR platform simulating real-life scenarios for language learning. See https://immerseme.co/.
[22] vTime is a VR social networking site allowing socialisation in an immersive virtual environment, using not only HiVR but also using VR or desktop PC. See https://vtime.net/.
[23] Mondly is a language learning application involving interaction in different languages with avatars in VR. See https://www.mondly.com/.

because they did not know where to look, especially when the pronunciation of the partner was poor or unfamiliar vocabulary was used. A heavy reliance on network quality also caused problems with the use of HiVR devices in education (Urueta & Ogi, 2019; Xie, Ryder, & Chen, 2019). Some students were accustomed to traditional teaching styles (Soto *et al.*, 2020) and seemed to be resistant to novel technologies (Xie, Ryder, & Chen, 2019).

Above all, the challenges of using HiVR for foreign language learning may not be limited to the above two perspectives (discomfort while using a head-mounted display and technical problems). Thus, it is important to take into account foreign language learners' experience of using HiVR for foreign language learning in order to make the best use of it in alleviating FLSA. This requires a qualitative method to gain in-depth information.

3.4 Existing Studies on HiVR and FLSA

Even though the affordances of HiVR technology and the relevant literature suggest that this technology may assist with alleviating FLSA, empirical evidence is still needed. Some scholars (Gruber & Kaplan-Rakowski, 2020, 2022; Kaplan-Rakowski & Gruber, 2021; Thrasher, 2022; York *et al.*, 2021) have recently published empirical studies on the effects of HiVR on foreign language anxiety. Their research design and findings are noted here for an up-to-date understanding of HiVR and FLSA. They did not motivate the research design of my study, because they were published after my fieldwork.

3.4.1 Research Methods Employed in Previous Studies

Both qualitative and quantitative methods were employed in previous research on HiVR and foreign language anxiety. Quantitative research design was mainly applied by those researchers to compare the effectiveness of HiVR and other learning modalities on foreign language anxiety levels, such as in the classroom (Chen, Hung, & Yeh, 2021; Thrasher, 2022), videoconferencing (Gruber & Kaplan-Rakowski, 2022), desktop VR (Han, 2022), mobile phones (Jeong & Jeong, 2021), and two computer-mediated communications (voice and video) (York *et al.*, 2021). Semi-structured interviews (Gruber & Kaplan-Rakowski, 2020), reflective journal (Kaplan-Rakowski & Gruber, 2021), and open-ended questionnaires (Thrasher, 2022; York *et al.*, 2021) were used to collect qualitative data in the existing empirical studies on HiVR and FLSA.

In terms of measuring FLSA, in addition to self-reported FLSA levels, some physiological measures of anxiety levels were involved, such as heartrate and electrodermal activity (Gruber & Kaplan-Rakowski, 2022) – although the results are yet to be reported – and salivary cortisol levels (Thrasher, 2022). The HiVR devices employed in previous studies consisted of high-end HiVR, such as HTC Vive, Oculus Rift S (Gruber & Kaplan-Rakowski, 2020, 2022; Han, 2022; Jeong & Jeong, 2021; Thrasher, 2022; York *et al.*, 2021); mobile HiVR (Chen, Hung, & Yeh, 2021); and stand-alone HiVR, such as Oculus Go) (Kaplan-Rakowski & Gruber, 2021). This is consistent with the types of HiVR used in other educational research (see Section 3.1.3), which would suggest that stand-alone HiVR is underexplored.

3.4.2 Previous Quantitative Research Findings

Quantitative results in the studies on FLSA and HiVR are inconsistent in currently available studies. The FLSA levels in HiVR groups decreased statistically significantly in some research (Gruber & Kaplan-Rakowski, 2022; Jeong & Jeong, 2021; Thrasher, 2022; York *et al.*, 2021). When comparing HiVR with other technologies, findings on HiVR's effects on FLSA are contradictory. A lower level of FLSA in the HiVR group was found by comparing to a zoom group (Gruber & Kaplan-Rakowski, 2022), a mobile group (Jeong & Jeong, 2021), and a classroom group (Thrasher, 2022). By contrast, no statistically significant differences on FLSA levels were found among three modalities groups (voice, video, and HiVR) by York *et al.* (2021).

In my view, the inconsistency in FLSA measures, the duration of the intervention and the type of tasks used in the intervention might explain the contradictory quantitative results of the above studies. First, some problems can be identified with the measurement of FLSA in these studies. In the seven-item version of the Foreign Language Classroom Anxiety Scale used in the study of York *et al.* (2021), there were contradictory items. For example, item 2 ("I don't worry about making mistakes in language class") and item 6 ("I feel confident when I speak in foreign language class") were positive statements. The other five items were negative statements about FLSA, such as item 1 "I never feel quite sure of myself when I am speaking in my foreign language class", item 3 "I start to panic when I must speak without preparation in language class", and so forth.

The same problem occurs in the questionnaire used by Gruber and Kaplan-Rakowski (2022), where some statements (items 2, 7, 12, and 15) [24] have opposite meanings to the rest of the items reflecting anxiety about speaking English. This resulted in the FLSA levels being calculated through a problematic manner in their studies. In addition, Thrasher (2022) did not present the details of the 11-item Foreign Language Anxiety Questionnaire used to measure FLSA in her article. This made it difficult to confirm the validity of the instrument for obtaining FLSA scores in her study.

There are also some problems in the existing research in terms of research design and data analysis. York *et al.* (2021) collected FLSA levels on three occasions after students completed a spot-the-difference task using three different modalities (voice, video, and HiVR). The participants in each group were involved in three interventions of the same type of task, but with different order of using each modality. After each intervention, FLSA levels were collected from participants. There would then be a question to me regarding the use of which modality caused a decrease in participants' FLSA levels in the study of York *et al.* (2021), as there is no difference in the technologies each participant used but just the order they used them.

A similar issue can be identified in the study of Thrasher (2022), where both HiVR and classroom interventions are applied to each participant but at different times. It is difficult to determine which learning environment brought the reduction in FLSA levels. As Thrasher (2022) stated, the FLSA levels of all participants diminished as time went by. Thus, it might be deduced that it was the increase in practising speaking frequency that led to a lower FLSA level rather than the use of HiVR. Additionally, due to the small sample size (N = 25), Thrasher (2022) did not examine the statistical significance of the decrease in FLSA levels, but the effect sizes. The salivary cortisol levels collected halfway through each task were lower in HiVR sessions compared with classroom sessions, but the statistical significance of the difference was not examined by Thrasher (2022). Since salivary cortisol levels were collected from a small sample of participants (N = 6), the change of the salivary cortisol levels is not evident to be interpreted as the change of FLSA levels after the intervention.

[24] item 2: *I don't worry when I make mistakes in English*;
item 7: *I feel self-confident when I speak English*;
item 12: *when I know I am going to speak English, I feel very confident and relaxed*;
and item 15: *I would probably feel comfortable around native English speakers*.

Moreover, FLSA levels in HiVR groups were lower than in other groups (voice, video and classroom) as revealed in the studies of York *et al.* (2021) and Thrasher (2022). These studies just reflected the anxiety levels while using HiVR to speak rather than the effects of HiVR on students' FLSA levels transiting to real-world situations.

3.4.3 Previous Qualitative Research Findings

From a qualitative perspective, students' perceptions of HiVR and FLSA varied. Some studies revealed positive results, such as participants being willing to use HiVR to deal with their unease about speaking the foreign language (Gruber & Kaplan-Rakowski, 2020); HiVR group students feeling more confident in using English (Chen, Hung, & Yeh, 2021); and more opportunities for vocabulary learning, grammar practice and negotiation being provided in HiVR (Kaplan-Rakowski & Gruber, 2021). Participants in the study of York *et al.* (2021) also reported that HiVR enhanced their ease of expression and was perceived as the most effective modality for lowering FLSA levels compared to the other two computer-mediated communications (voice and video). Participants in the study by Enkin (2022) noted that speaking through avatars helped them feel less stress during conversations. The qualitative results in the study of Chen, Hung, and Yeh (2021) also indicated that student attitudes towards English were more positive in the HiVR group than in the classroom group.

Nevertheless, some negative perceptions of HiVR were reported. In addition to the challenges of using HiVR in foreign language learning discussed in Section 3.3.2, problems of HiVR are associated primarily with the distractions and realism of the virtual environment and avatars (Gruber & Kaplan-Rakowski, 2020; Kaplan-Rakowski & Gruber, 2021; York *et al.*, 2021), the overheating of VR headsets (Gruber & Kaplan-Rakowski, 2020), and the lack of body language, facial expressions and lip-reading (Kaplan-Rakowski & Gruber, 2021). In contrast to the results of Enkin (2022), anonymity was not seen as a benefit for FLSA alleviation by participants in York *et al.* (2021).

In conclusion, previous studies on HiVR and FLSA have used a variety of research methods and instruments and produced inconsistent quantitative and qualitative results. As stated above, all the studies reviewed in this section were published after my fieldwork was completed. Therefore, this section does not contribute to the research gap identification, but it provides a sustainable overview of my research topic.

3.5 Summary

This chapter introduced a technology that could mitigate FLSA – High-immersion Virtual Reality (HiVR). A brief introduction to HiVR technology was given in Section 3.1, followed by its main affordances in learning (see Section 3.2). The potential and challenges of HiVR in foreign language learning were illustrated in Section 3.3, mainly in terms of alleviating FLSA. Research gaps were identified indicating that insufficient research has been conducted to investigate the effects of HiVR on affective dimensions of foreign language learning (e.g. FLSA) and demonstrating that an integration of quantitative and qualitative data on this topic is essential. The existing empirical studies on HiVR and FLSA were evaluated at the end of this chapter to give an up-to-date overview of the current research topic. The following chapter will demonstrate the theoretical foundations for the research design of my study.

Chapter 4. THEORETICAL FRAMEWORK

From the discussion in the last two chapters (Chapter 2 and Chapter 3), it is clear that HiVR technology may be beneficial for alleviating FLSA, but the way the technology is exploited is crucial. Bearing in mind the causes of FLSA among Chinese EFL learners (see Section 2.3.3), the potential of HiVR to reduce FLSA should be considered from both individual/internal and situational/external perspectives. Internally, positive emotions should be encouraged while using the foreign language in order to improve learners' self-confidence and willingness to communicate. Externally, an environment needs to be created that will increase students' experience of using the target language authentically as well as improve their oral proficiency.

This chapter develops a theoretical framework to illustrate the basis for my research assumptions and design. I first introduce the theoretical foundations and explain their connections to my study. I then examine Virtual Reality (VR) exposure-based therapy and its application in treating social anxiety disorders; this is the main research rationale. Next, I illuminate situated learning theory and its role in guiding the learning design within HiVR. Finally, I outline the assumptions for this research. These assumptions guide the research design.

4.1 HiVR and Exposure-based Therapy

When I designed my research in 2019, to the best of my knowledge, no specific empirical studies had been published on HiVR and FLSA. My assumption for this thesis – that HiVR technology may help to reduce FLSA – was mainly predicted on the psychological application of HiVR in the treatment of mental disorders. VR exposure-based therapy is the common method used for treating mental disorders through HiVR technology. Its effectiveness has been confirmed in treating acrophobia (Coelho *et al.*, 2006), fear of flying (Maltby *et al.*, 2002), post-traumatic stress disorder (Botella *et al.*, 2015), panic disorder and agoraphobia (Botella *et al.*, 2007), social phobia (Gebara *et al.*, 2015), and public speaking anxiety (Safir, Wallach, & Bar-Zvi, 2012). The theoretical foundation and practical evidence for applying VR exposure-based therapy in treating FLSA are presented in this section.

4.1.1 Exposure-based Therapy

In order to understand VR exposure-based therapy, it is necessary to understand the theoretical fundamentals of exposure-based therapy. Exposure-based therapy (or exposure therapy) refers to exposing an individual into fear-provoking stimuli in the absence of aversive feedback (Craske *et al.*, 2014). For my study, the use of exposure-based therapy is to expose foreign language learners who are anxious about speaking the target language to the situation of speaking but receiving no negative feedback on their performance. This is because poor experience, such as harsh corrections from teachers, may increase students' anxiety about speaking the target language (see Section 2.3.2). Chinese EFL instructors are often strict about the use of grammar and vocabulary, and they correct students' each mistake, which results in students' fear of negative evaluation (Yan & Horwitz, 2008). In addition, due to the fact that China belongs to the expanding circle of World Englishes, Chinese EFL learners have limited access to a spontaneous English-speaking environment (see Chapter 1). Exposing students to English-speaking situations may also increase their experience of using the target language, which may contribute to the reduction of FLSA levels (see Section 2.3.3).

Exposure-based therapy is claimed to be derived from extinction learning, which means the response to a conditioned stimulus decreases gradually when the stimulus is presented without reinforcement (Hartley & Phelps, 2012). However, the theoretical mechanisms of exposure are a matter of debate. Emotional processing theory and inhibitory learning theory are both used to explain exposure therapy (Craske *et al.*, 2014; Deacon *et al.*, 2013; Foa & Kozak, 1986).

The concept of emotional processing originates in Lang (1977)'s work on fear-relevant imagery, which is a behaviour therapy for fear reduction. Lang (1977, 1984) proposed that a fear image is a cognitive structure containing stimulus, response, and meaning information that leads to avoidance behaviour. This is the basis for the concept of fear structure, which were divided into normal and pathological fear structures by Foa and Kozak (1986). A normal fear structure refers to fear generated in a situation that really needs the avoidance behaviour. For example, if a car hurtles toward me, my heart rate will accelerate and I will scan the road then move to safety (Foa, Huppert, & Cahill, 2006). In contrast, a pathological fear structure involves an excessive response to a situation that does not need such a degree of fear or avoidance behaviour. For instance, the situation is safe, but one avoids it (Foa, Huppert, & Cahill, 2006). Therefore, Foa and Kozak (1986) take the position that emotional

processing is the process of weakening erroneous pathological associations with stimuli, responses, and meaning, and replacing them with nonpathological associations.

Foa and Kozak (1986) proposed two conditions for emotional processing: the activation of a fear structure, and the incorporation of new information that is incompatible with the pathological elements of the fear structure. The incompatible information comes from two sources: within-session habituation and between-session habituation. Habituation means a decrease in response to a frequently repeated stimulus, not the disappearance of fear. Within-session habituation occurs when fear decreases as exposure is conducted during a therapy session, and between-session habituation occurs when fear decreases between or across therapy sessions. Within-session habituation is seen as the pre-requisite for between-session habituation, and between-session habituation is claimed to form the basis for long-term learning (Craske *et al.*, 2008). Learning, here, refers to "the capacity of a living organism to alter its behaviour as a result of experience" (Eelen, Hermans, & Baeyens, 2001, p. 250).

Successful emotional processing in exposure-based therapy includes 1) the initial fear activation; 2) within-session habituation; 3) between-session habituation (Foa, Huppert, & Cahill, 2006). The initial fear activation can be seen as the peak response during the first exposure trial, in which peak response is the difference between the baseline for the session and the maximum level recorded during the first part of the exposure (Kozak, Foa, & Steketee, 1988). In my study, being involved in an authentic communication in the target language can be seen as the initial fear activation for anxious foreign language learners. However, it is claimed that over-reliance on habituation as an indicator of learning may have negative consequences. This is because this idea emphasises that anxiety is a bad thing, and treatment is successful only when anxiety is completely eliminated (Jacoby & Abramowitz, 2016). Some research (Lang & Craske, 2000; Tsao & Craske, 2000) even supports improvement in anxiety in the absence of habituation. Acknowledging these criticisms, Lang, Craske, and Bjork (1999) proposed an alternative model to explain the mechanism for exposure-based therapy for fear and anxiety: inhibitory learning theory.

Inhibitory learning theory critiques the replacement of the new non-fear structure over fear structure but focuses on dealing with the fear-based association between a conditioned stimulus and an unconditioned stimulus (Arch & Abramowitz, 2015). In other words, inhibitory learning theory focuses on the diminishment of access to fear-based associations rather than the association itself. Linking to my study, this refers to anxiety about speaking

the foreign language still exists, but the feeling of anxiety is undermined after exposure. This can be connected to a Pavlovian conditioning model; a conditioned stimulus (e.g. a neutral stimulus) is followed by an unconditioned stimulus (e.g. an aversive stimulus), then the conditioned stimulus will elicit a conditional response (e.g. anticipatory fear reactions). The conditional response will become a predictor of the unconditioned stimulus, which means that presentation of the conditioned stimulus will activate the memory of the unconditioned stimulus. Therefore, fear or anxiety is evoked by thinking about the aversive unconditioned stimulus (Craske *et al.*, 2014), which is called fear conditioning.

Inhibitory learning theory emphasises extinction rather than habituation. Whereas habituation (in emotional processing theory) refers to a reduction in fearful responses to repeated presentations of a stimulus, extinction is a method of associative learning in which conditional stimulus is repeatedly presented in the absence of the unconditional stimulus (Craske *et al.*, 2008). When the absence of the unconditional stimulus is presented with repeated conditional stimulus, the unconditional stimulus will no longer be aroused by the conditional response (Craske *et al.*, 2014). Extinction is expected to result in a change of expectation about the feared consequence (e.g. no longer expecting the unconditional stimulus to follow the conditional stimulus) and the corresponding behaviour (e.g. finding approaches to solving the problem instead of avoidance) (Jacoby & Abramowitz, 2016). In the current study, a change in expectation about feared consequences (no longer expecting anxiety while speaking the foreign language) and the corresponding behaviour (finding ways to speak the foreign language instead of avoiding it) can be expected after applying exposure-based therapy.

Inhibitory learning theory also emphasises that the original fear-based associations could return after exposure-based therapy sessions (Jacoby & Abramowitz, 2016). Thus, if non-threatening learning is not continually implemented, the conditional stimulus – no unconditional stimulus association will be less and less accessible as time goes by. The newly created conditional stimulus – no unconditional stimulus associations should be remained and retrieved over time in the exposure (Tolin, 2019). That is to say, from the inhibitory learning perspective, exposure should be repeated over a long period in order to maintain its effect (Arch & Abramowitz, 2015). As a result, in order to maintain the potential effects of HiVR on FLSA, my study involved a regular nine-session intervention rather than

one or two trials. The intervention was originally intended to last for 12 sessions but had to be shortened due to the pandemic (see Section 5.3.3 for detailed explanation).

4.1.2 VR Exposure-based Therapy and Anxiety in Interpersonal Interactions

VR exposure-based therapy is an application of HiVR technology to achieve exposure-based therapy. The similarity of the HiVR environment to the real world provides the illusion that users are immersed and interacting with objects in the real world, and therapists can more easily control and manipulate exposure elements in HiVR than in *in vivo* exposure (Morina *et al.*, 2021). Through exposure to the designed virtual experience, people can change their behaviours, thoughts, and emotions (Riva, 2005). Moreover, VR exposure-based therapy is claimed to have real-life impact, which means an intervention conducted in HiVR could influence patients' behaviour in the real world (Opriş *et al.*, 2012).

Since the 1990s VR exposure-based therapy has been used to treat anxiety and related disorders (Oing & Prescott, 2018). In spite of the broad and promising research findings on VR exposure-based therapy and public speaking anxiety (Anderson *et al.*, 2005; Blöte *et al.*, 2009; Safir, Wallach, & Bar-Zvi, 2012; Yuen *et al.*, 2019), the effect of VR exposure-based therapy in interpersonal situations or interactions is not sufficiently addressed. In contrast to other systematic reviews of VR and anxiety disorders (Emmelkamp, Meyerbröker, & Morina, 2020; Oing & Prescott, 2018), the empirical studies I review here are focused on the use of VR exposure-based therapy in real-time interpersonal interactions. Moreover, only studies that employed High-immersion VR (HiVR, with a head-mounted display) as the technology for VR exposure-based therapy are discussed rather than those using Low-immersion VR (LiVR, such as desktop VR). The distinction between HiVR and LiVR is discussed in Section 3.1. In particular, studies on social anxiety disorders and HiVR are reviewed in this thesis, since social anxiety shares the same characteristics as FLSA that they are both related to spontaneous interpersonal interactions.

According to Stein and Stein (2008), social anxiety disorders (also known as social phobias) refer to the fear and avoidance of interpersonal encounters or endurance of such situations with intense discomfort due to the fear of negative evaluation. Individuals who suffer from social anxiety disorders are usually quiet in groups and avoid speaking in public or interacting with unfamiliar people. Research into social anxiety disorders and the use of VR

exposure-based therapy for social interaction is scarce, but promising results have been observed. Morina *et al.* (2015) provided initial evidence for the effectiveness of applying verbal interactions in HiVR interventions and illustrated its long-term benefits on social anxiety and self-efficacy. Specifically, participants in the high anxiety group in the study of Morina *et al.* (2015) witnessed a significant reduction in social anxiety levels and an increase in self-efficacy after three months of HiVR treatment. The free dialogues conducted in their study covered topics such as buying clothes (e.g. a bra) in a shop, attending a job interview, dining in a restaurant (e.g. with a blind date or a friend), talking to strangers, being interviewed by a journalist, and giving a presentation to an audience.

Kampmann *et al.* (2016) conducted a randomised controlled trial on the virtual scenarios developed by Morina *et al.* (2015) in order to compare their effectiveness in treating social anxiety disorders with *in vivo* exposure. It was one of the first studies to examine the efficacy of pure VR exposure-based therapy without any cognitive therapies. A randomised controlled trial is an experimental design that assigns participants to groups randomly for receiving the intervention (experimental group) or an alternative treatment (control group) in order to assess the effectiveness of the intervention by comparing the groups (Kendall, 2003). This is similar to the experiment design applied in my study (see Section 5.3). Kampmann *et al.* (2016) found a significant decrease in social anxiety levels after the intervention with HiVR. Moreover, Kampmann *et al.* (2016) claimed that the effects of VR exposure-based therapy on participants could be generalised to real-life social situations. They suggested that a larger number of more diverse and flexible conversations should be involved in VR exposure-based therapy for a more realistic and unpredictable social interaction experience. This was taken into account in the learning task design in my study (see Section 5.3.3).

Cognitive behavioural therapy is often combined with VR exposure-based therapy in treating social anxiety disorders. Cognitive behavioural therapy is a talking therapy that aims to change the way the client thinks and behaves to deal with mental health problems (Hofmann *et al.*, 2012). It is claimed to be the most effective treatment for social anxiety disorders (Mayo-Wilson *et al.*, 2014). The application of cognitive behavioural therapy in VR protects client's privacy and provides exposure situations (Geraets *et al.*, 2019). A randomised controlled trial conducted by Bouchard *et al.* (2017) showed that conducting cognitive behavioural therapy in VR exposure was more practical and effective than *in vivo* exposure,

and the effects were maintained after six months. The social interactions used in their study included speaking in front of an audience in a meeting room (two scenarios); having a job interview (two scenarios); introducing oneself and having a talk with supposed relatives in an apartment; acting under the scrutiny of strangers on a coffee shop patio; and facing criticism or insistence in two situations (meeting unfriendly neighbours, refusing to buy goods from a persistent seller in a store). In addition, Geraets *et al.* (2019) reported positive results regarding the effects of VR exposure-based therapy with cognitive behavioural therapy on social interaction anxiety, indicating that exposure to feared situations in HiVR reduced anxiety.

To summarise, when treating social anxiety disorders, VR exposure-based therapy seems to be beneficial, especially when implementing with cognitive behavioural therapy. Accordingly, an assumption can be made that VR exposure-based therapy may be useful for alleviating FLSA, especially when combined with certain learning approaches. This guided me to integrate learning approaches in the HiVR-based learning design. The specific learning approach – situated learning – underpinning my learning design is illustrated in the following section.

4.2 HiVR and Situated Learning

From both theoretical and pedagogical perspectives, situated learning can be combined with the use of HiVR and this may contribute to FLSA alleviation. In terms of pedagogy, situated learning engages students in experiential language learning tasks within realistic contexts and authentic settings (Felix, 2002). This section presents a detailed introduction to situated learning theory and its utilisation as a model of learning design with HiVR in my study.

4.2.1 Situated Learning Theory

Situated learning is an approach proposed by Lave and Wenger (1991). It views learning as a social, cultural, and contextual activity rather than just an individual, decontextualised and cognitive one. In other words, learning should take place in the situation in which it will be applied in the future rather than be acquired from textbooks. For example, if someone wants to be a cook, they should learn in a kitchen. In foreign language learning, situated learning can be described learning a foreign language while using the language in real-life situations.

Lave and Wenger (1991) believed that learning is a social process in which knowledge is co-constructed by the practitioners in a community rather than an individual acquisition of knowledge from textbooks. In other words, newcomers will master the knowledge and skill when they fully participate in the social practices of a community (Lave & Wenger, 1991). This is related to the notion of legitimate peripheral participation, the process by which newcomers become part of a community of practice (Waite & Pratt, 2015). Apprenticeship could be used to explain legitimate peripheral participation. For example, apprentices are first accepted by the new community and then becoming young masters after learning basic skills. When the young master's skills improve, they will be allowed to enter the core of the community to do important work. All the skills learning processes happen when they conduct their practice in the community.

4.2.2 Situated Learning Theory as a Model for Learning Design

According to Gee (2004), situated learning is rarely used in traditional classroom instruction because of the complexity and difficulty of creating a real-world setting, but technologies provide possibilities for using it. This makes the involvement of HiVR with situated learning more meaningful. In conventional situated learning theory, the primary focus is on participation and identity transformation in a community of practice. However, my study applied situated learning theory as a model of learning design in HiVR.

The benefits of applying situated learning theory as a model of instructional or learning design in HiVR have been investigated by previous researchers. McLellan (1991) stated that situated learning theory seemed to offer a model to aspire to in adapting VR as a tool for training and learning, and exemplified the application of it in the design and use of HiVR for helicopter and pilot training. McLellan (1996) further built a model for situated learning; its key components are stories, reflection, cognitive apprenticeship, collaboration, coaching, multiple practices, articulation of learning skills and technology. Sharma (2016) investigated the effectiveness of this model in alleviating mathematics anxiety among 99 elementary school students. A 2*2 factorial design involving two teaching models (situated learning and traditional method) and two effortful controls (high and low) was conducted in his study. After the treatment, a decrease in mathematics anxiety levels was found in the situated learning group compared to the traditional method group, which indicates the effectiveness of situated learning model in alleviating mathematics anxiety.

In addition, Schell and Black (1997) argued that situated learning increases students' motivation by creating an atmosphere appropriate for learning. Several later studies (Edwards *et al.*, 2008; Ketelhut *et al.*, 2007; Kneebone *et al.*, 2004) also revealed that situated learning could improve students' motivation and interest in learning. Mei and Sheng (2011) designed a virtual hospital situated learning system for medical students to develop knowledge of human organs. The results showed that situated learning increased students motivation by 18.9%. Herrington and Oliver (2000) summarised the characteristics of learning environments following situated learning theory as follows:

- Provide an authentic context that reflects the way the knowledge will be used in real-life;
- Provide authentic activities;
- Provide access to expert performances and the modelling of processes;
- Provide multiple roles and perspectives;
- Support collaborative construction of knowledge;
- Provide coaching and scaffolding at critical times;
- Promote reflection to enable abstractions to be formed;
- Promote articulation to enable tacit knowledge to be made explicit;
- Provide for integrated assessment of learning within the tasks.

Hossainy *et al.* (2012) followed this framework and found significant changes ($p < .001$) in intrinsic motivation among students in the situated learning group. Intrinsic motivation refers to the inner drive that propels a person to pursue an activity rather than being motivated by external products, pressures, or rewards (Fishbach & Woolley, 2022). The improvement in motivation for or interest in learning may alleviate FLSA because this improvement may contribute to a more pleasant experience for foreign language learners.

Based on the above evidence, it can be hypothesised that a VR environment designed by using situated learning theory as a model for learning design may reduce anxiety levels. A detailed description of how situated learning theory informed the learning design for my study can be found in Section 5.3.3.

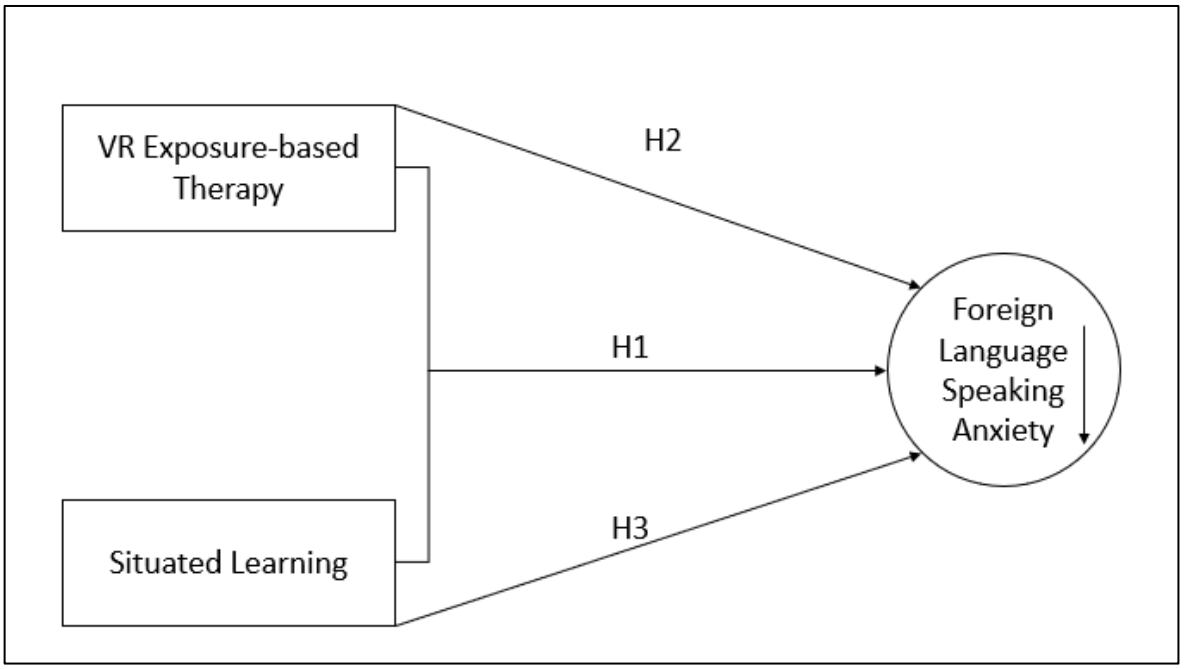
4.3 The Framework and Hypotheses

In conclusion, FLSA is a common issue among foreign language learners, and it needs to be alleviated because it has adverse effects on foreign language learning and use (see Chapter 2). Given all the theoretical and practical evidence, HiVR seems to be a potential solution to cope with FLSA, especially considering the specific causes of FLSA among Chinese EFL learners, such as lack of experience of speaking English in and out of classroom (see Chapter 2). Although the potential for utilising HiVR in foreign language learning has been demonstrated (see Chapter 3), no empirical studies had been conducted on HiVR and FLSA before my fieldwork in 2020. My hypothesis that HiVR may help alleviate FLSA is mainly based on the effectiveness of VR exposure-based therapy in coping with social anxiety disorders (see Section 4.1.2). The main application of VR exposure-based therapy in my study is to provide students with a simulation of real-life scenarios to speak English that can potentially arouse anxiety. However, possible negative feedback is removed by providing positive feedback only. Despite some studies on HiVR and FLSA have been published since 2020, their results are inconsistent. It appears that pedagogy and learning approaches are not taken into consideration in the existing studies on HiVR and FLSA. However, this is not the research gap guided my thesis, as these studies are published after my fieldwork.

In addition, based on OECD (2015)[25], the effectiveness of technology in facilitating learning depends on how the technology is exploited rather than the technology *per se*. My study therefore takes the affordances of HiVR into consideration (see Section 3.2) and utilises situated learning theory as a model for learning design in HiVR (see Section 4.2.2). Specifically, an authentic English-speaking environment was created in HiVR with cooperative learning activities requiring students' manipulation of the virtual environment and communication in English (see Chapter 5). My main assumption was that combining VR exposure-based therapy with situated learning theory might be helpful for alleviating FLSA among Chinese EFL learners. It is also necessary to understand the influence of VR exposure-based therapy and situated learning theory on FLAS. The theoretical framework for the present study is presented in Figure 4-1.

[25] Organization for Economic Cooperation and Development.

Figure 4- 1 *Theoretical framework for my study (my own creation)*



In this framework, the following three hypotheses are proposed. Their relationship to the research design and research questions in my study is demonstrated in the following paragraphs.

> **Hypothesis 1** (**H1**): Combining VR exposure-based therapy and situated learning theory results in a decrease in FLSA levels.
>
> **Hypothesis 2** (**H2**): Simply being exposed to the HiVR environment results in a decrease in FLSA levels.
>
> **Hypothesis 3** (**H3**): Utilising situated learning theory as a model for learning design results in a decrease in FLSA levels.

The above hypotheses led to an experimental research design. Comparison of FLSA levels after treatments can be made using both HiVR and situated learning theory for intervention (H1), as well as HiVR and traditional teaching method (H2), and situated learning theory and a traditional teaching environment (H3). In the Chinese EFL setting, the widely used traditional teaching method is teacher-centred learning (see Section 2.3.3), and the majority of learning is undertaken in classrooms. Therefore, teacher-centred learning and the classroom were taken into consideration in the research design. The following four groups were involved, consisting of three treatment groups and one control group:

SVR group (treatment group): situated learning and HiVR
TVR group (treatment group): teacher-centred learning and HiVR
SC group (treatment group): situated learning and classroom
TC group (control group): teacher-centred learning and classroom

This results in the following two Research Questions (RQs). RQ 1 is intended to address H2, and RQ 2 can be linked to H1 and H3.

RQ 1: To what extent (if any) do the FLSA levels of students who used HiVR change after the intervention?
RQ 2: To what extent (if any) do two learning environments (HiVR and classroom) combined with two learning approaches (situated learning and teacher-centred learning) have different degrees of influence on students' FLSA levels?

In addition, users' perceptions and experiences of applying HiVR for foreign language learning is underexplored but needs to be investigated (see Chapter 3). In my study, Chinese EFL learners' perceptions of the reasons for possible changes in FLSA levels should be examined. I therefore employed a mixed-method research design (see Chapter 5) in order to integrate qualitative data with experimental research.

The following three RQs were proposed based on the qualitative perspective for my research.

RQ 3: How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?
RQ 4: What are students' perceptions of the impact of HiVR on FLSA?
RQ 5: What are students' experiences of using HiVR in oral English learning?

Chapter 5. RESEARCH METHODOLOGY AND METHODS

Previous chapters have illustrated that both statistical evidence and qualitative findings on the impact of High-immersion Virtual Reality (HiVR) on Foreign Language Speaking Anxiety (FLSA) are insufficient. My study attempts to fill this research gap by employing a mixed methods design. In this chapter, the research methodology is explained along with the research setting and participants. The instruments used to collect both quantitative and qualitative data are introduced as well as the interventions conducted in different groups. The final section covers data analysis methods and ethical issues.

5.1 Overview of the Research Design

A mixed-method experimental design (see Section 5.2) was applied in the current research. From a quantitative perspective, it aimed to examine the difference in EFL learners' FLSA levels before (pre-test) and after (post-test) the interventions, and from a qualitative perspective, it aimed to discover their perceptions of HiVR and FLSA and their experience of using HiVR for oral English learning. A 2*2 factorial design (see Section 5.3.1) was applied to the experiment, which involved four groups of Chinese EFL learners (35 students in each group) at a university in two different learning environments (HiVR and classroom) and using two learning approaches (situated learning and teacher-centred learning). The difference in learning environments and learning approaches also guided the 9-session intervention for each group (see Section 5.3.3), which was delivered over a three-month period. The learning content used in HiVR was developed with technical support from a VR company (see Section 5.3.3 for the collaboration details).

Two types of measurements were utilised for measuring FLSA levels: the English speaking anxiety scale (ESAS, see Section 5.3.4) and the self-rating of FLSA levels in specific tasks (see Section 5.3.6). A speaking test (see Section 5.3.5) was conducted to obtain the video recordings for the participants to rate their FLSA levels during the test. The measurement of FLSA and the speaking tests were piloted with 19 EFL learners in a different university ahead of the main study in order to ensure its reliability and validity (see Section 5.3.4.2 and 5.3.5.2). The two types of FLSA levels collected from the pre-test and post-test were compared within and among groups in order to answer RQ 1 and RQ 2 (see below).

Qualitative data was collected through an open-ended questionnaire (see Section 5.4) to gather students' perceptions of HiVR and FLSA and their experience of using HiVR for oral English practice after the intervention for RQ 3, RQ 4 and RQ 5 (see below). The quantitative and qualitative datasets were integrated for the analysis to obtain a comprehensive understanding of the plausible impact that HiVR may have on learners' FLSA. The Research Questions (RQs) are restated below to aid understanding of the research design.

RQ 1: To what extent (if any) do the FLSA levels of students who used HiVR change after the intervention?

RQ 2: To what extent (if any) do two learning environments (HiVR and classroom) combined with two learning approaches (situated learning and teacher-centred learning) have different degrees of influence on students' FLSA levels?

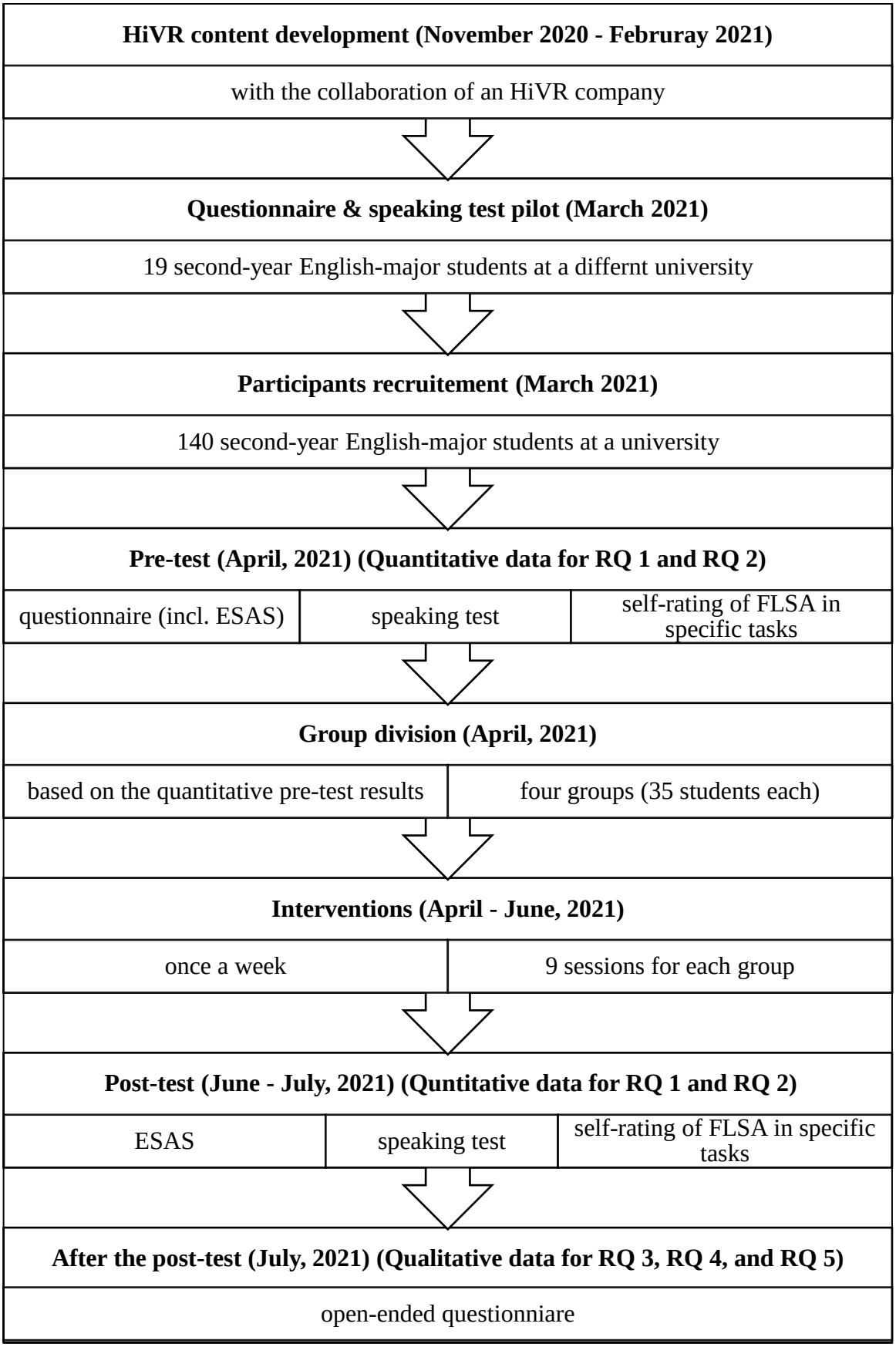
RQ 3: How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?

RQ 4: What are students' perceptions of the impact of HiVR on FLSA?

RQ 5: What are students' experiences of using HiVR in oral English learning?

An overview of the fieldwork of the current research as well as the corresponding Research Questions (RQs) is presented in Figure 5-1. The following sections explain the research methodology and methods in detail.

Figure 5- 1 *Fieldwork procedure and corresponding RQs in my study*

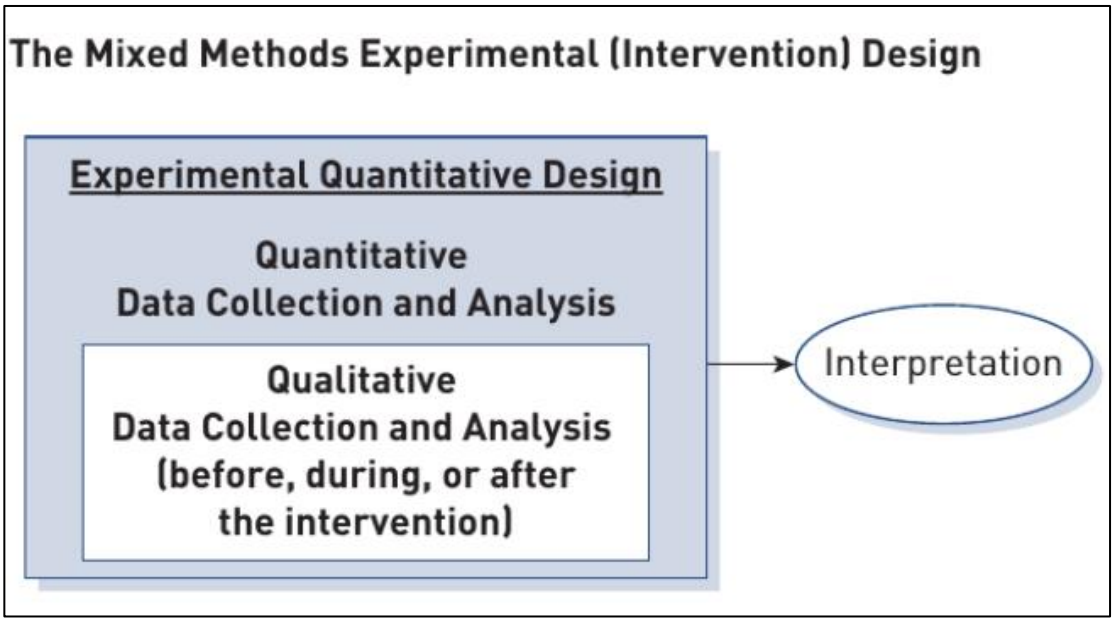


5.2 Mixed Methods Experimental Design

Mixed methods research is described as a third methodology that sits between quantitative and qualitative methods; the researcher obtains both quantitative and qualitative data and then draws inferences by integrating or combining the two databases (Creswell, 2021). The integration of data expands the breadth and depth of understanding a problem. The integration may occur in the process of employing viewpoints, data collection, and data analysis (Johnson, Onwuegbuzie, & Turner, 2007). The insights obtained from mixed methods research go beyond the sum of quantitative and qualitative results, but are a connection or embedment of them (Creswell & Clark, 2017). A formula 1 + 1 = 3 was proposed by Fetters and Freshwater (2015) to describe the integration challenge of mixed methods research. The mixed methods research requires not simply adding separate quantitative and qualitative approaches but adding values by articulating these two approaches and databases to generate research findings. The mixed methods approach incorporates the strengths of the quantitative approach (e.g. generalisation and precision) and the qualitative approach (e.g. in-depth individual perspectives) to produce a more comprehensive understanding and rigorous corroboration of a research problem (Creswell & Clark, 2017).

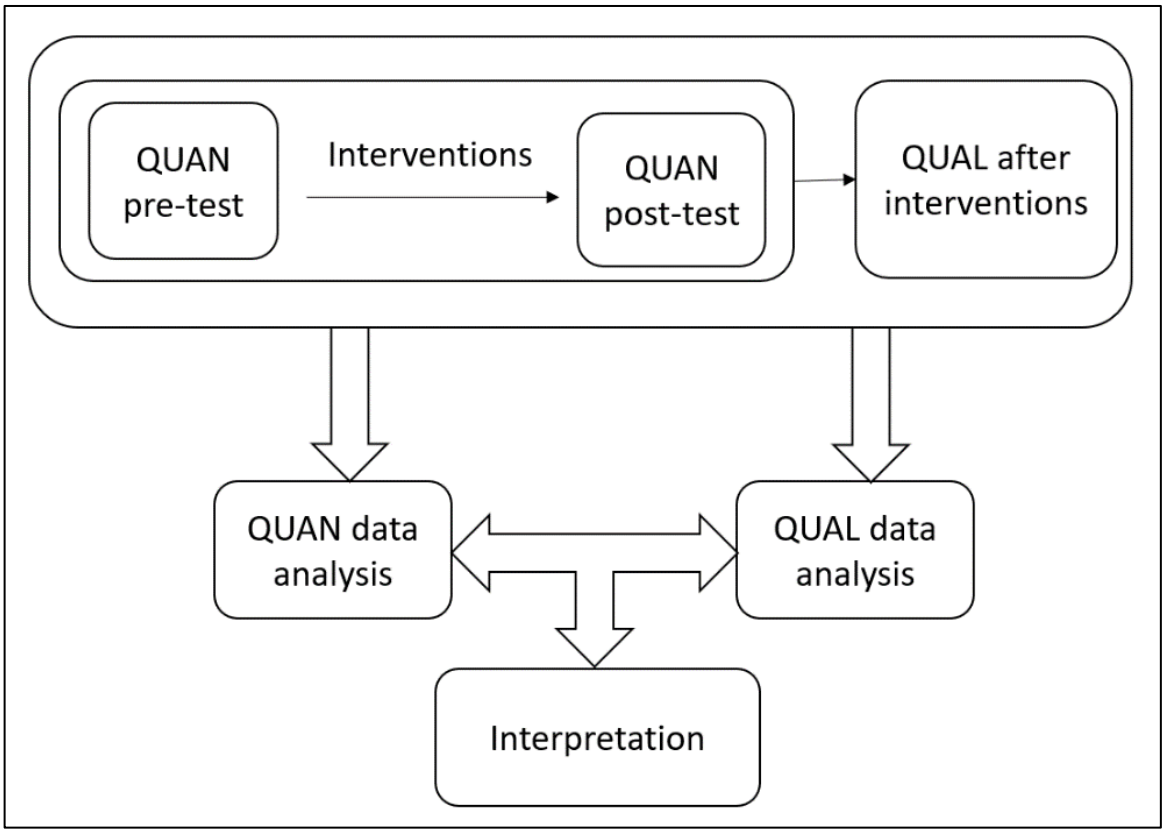
In particular, a mixed methods experimental design (Creswell, 2021) was applied to my study. Mixed methods experimental design (see Figure 5-2) refers to a study that primarily conducts an experiment or an intervention trial and adds qualitative data (before, during or after the intervention) to augment the results (Creswell & Clark, 2017). It is known as a complex mixed methods design, in which core mixed methods designs (convergent, explanatory sequential, and exploratory sequential) are embedded into a primary quantitative or qualitative method, another methodology or a theoretical framework (Creswell, 2021).

Figure 5- 2 *The mixed methods experimental design (Creswell & Clark, 2017, p. 105)*



My study is a complex design, adding convergent and explanatory sequential designs to a quantitative experiment (see Figure 5-3). Convergent design offers a way to compare results by merging two data sets to find convergence or divergence; this sometimes involves data transformation (Creswell, 2021). In particular, I compared and combined students' descriptions of their FLSA levels with the quantitative results of FLSA changes before and after the intervention. In an explanatory sequential design, qualitative data is used to explain the quantitative results in order to understand the unusual or surprising quantitative findings or to explain the quantitative results in more detail (Creswell, 2021). In my study, qualitative data on students' perceptions of their FLSA levels was used to explain the quantitative results of FLSA changes and differences after the intervention. The final interpretation of data analysis was based on integrating the quantitative and qualitative data sets.

Figure 5- 3 *The mixed methods experimental design of my study (own creation)*[26]



5.3 Quantitative Data Collection (RQ 1 and RQ 2)

The quantitative part of my study adopted experimental research, which is based on the deductive approach used to test the hypotheses (Blackstone, 2012). The use of experimentation has a long history in psychology and education, as well as in educational technology, to test hypotheses and interpret the results in terms of why something has occurred (Ross & Morrison, 2004). Through analysing data collected in a controlled environment, the researcher can determine whether the hypotheses are supported by the evidence (Tanner, 2018). Specifically, quantitative data was collected for RQ 1 and RQ 2.

The specific hypotheses underlying the RQ 1 and RQ 2 were as follows:

> **Hypothesis 1:** The FLSA levels of students who used HiVR reduce after the intervention.

[26] In the figure, QUAN refers to quantitative, QUAL refers to qualitative.

Hypothesis 2: The intervention leads to a different level of decrease in FLSA for each group, and the control group (classroom and teacher-centred learning) results in the smallest decrease.

5.3.1 2*2 Factorial Design

The present study employed a 2*2 factorial design (Fisher, 1936), involving two independent variables. These were 1) learning environments (X), with two conditions – in HiVR (X1) and in the classroom (X2); and 2) learning approaches underlying the intervention (Y), again with two conditions – situated learning (Y1) and teacher-centred learning (Y2). The theoretical foundations for carrying out the 2*2 factorial design were illustrated in Section 4.3. Independent variables are the suspected cause under consideration in the research project. The dependent variable, also known as the effect variable, is the level of FLSA in my study. A matrix was created in Table 5-1 to illustrate the 2*2 factorial design, which also indicates the characteristic of each group. Both the sole effects and the interactive effects of the independent variables on the dependent variable (FLSA levels) were examined in my study.

Table 5- 1 The 2*2 factorial design of my study

		Learning approaches (Y)	
		Situated learning (Y1)	Teacher-centred learning (Y2)
Learning environments (X)	In HiVR (X1)	SVR[27] (N = 35)	TVR (N = 35)
	In classrooms (X2)	SC (N = 35)	TC (N = 35)

5.3.2 Participants

This study was conducted at a university in the southeast of China (hereafter, University), which belongs to the three-circle Model of World Englishes (Kachru, 1996) (see Chapter 1) that is an EFL context. The University is a comprehensive university that is mainly focused on teacher education. Permission for recruiting participants and conducting the fieldwork

[27] SVR: situated learning + HiVR; TVR: teacher-centred learning + HiVR; SC: situated learning + classrooms; TC: teacher-centred learning + classrooms.

was obtained from the head of the Foreign Language College at the University, a connection based on my own social network. This University is located in a province where spoken English is undervalued in the education system. It is documented that 71.6% of students in the Foreign Language College at the University are from the same province in which the University is located. This indicates that students' proficiency in spoken English was not well developed before entering the University, although most of them had learnt English since the third year of primary school. This lack of proficiency stems from the fact that spoken English skills are not required for the National College Entrance Examination (Gaokao, see the explanation in Section 1.1); the focus of this examination is on grammar and vocabulary learning, especially for reading and writing (Wang, 2014).

The English-major students in the University with a population of approximately 3,500 were targeted for participant recruitment. Although non-English-major students present the largest group of English learners at university level, it may be very difficult for them to speak English and difficult for them to complete a task wholly in English. Students were chosen from English majors rather than non-English majors to exclude cases where speaking anxiety may be caused purely by limited language proficiency. Although English-major students' oral English proficiency is not well-developed before they enter the university and, their overall English proficiency is often higher than non-English major students because there is a requirement of their English exam scores before being offered to study English major. However, the exam scores are confidential. The other reason for excluding non-English majors is that the population and sample must be as homogeneous as possible with regard to characteristics that could be expected to affect the outcome or response to the intervention in the experiment (Campbell & Stanley, 2015). Non-English-major students had different experiences of English learning at the university as they are from different departments. However, for English majors, their English courses are delivered in a similar manner, and I could coordinate with their tutors to remove similar teaching content or avoid using similar teaching method (e.g. situated learning) as used in my project.

Volunteer sampling was used to recruit participants because students had to know in advance of participating that they would be involved in an intervention. According to Jupp (2006), volunteer sampling is a form of purposive/non-random sampling used when participants agree to be recipients of such treatment or interventions. The major problem with volunteer

sampling is on the generalisation of research findings. Thus, one should be cautious when stating the results of the research (Jupp, 2006).

I went to each class (10 classes in total) to introduce my research and recruited participants find distributing a participant information sheet (see Section 5.9). A total of 172 students filled in the consent forms (see Section 5.9) to show their willingness in participating in the study at the beginning, and they responded to the questionnaire (see Section 5.3.4) before the intervention. However, before the intervention started, 30 students dropped out, so in total 140 participants (132 females, 8 males) participated throughout the study. No participant withdrew after the intervention started, and I collected data from these 140 participants to analyse for my thesis. Four other students joined the study, but only as partners for pair-work in class, so no data was collected from them. The gender imbalance of the sample was due to the fact that English majors in Chinese universities always have more female students than males (Van Der Meij & Zhao, 2010). It is also stated that gender differences does not influence Chinese EFL learners' language anxiety (Jiang & Dewaele, 2019).

The students ranged in age from 18 to 21, with an average age of 19.54. The students had studied English for an average of 10.63 years, and approximately 70% had studied it for 11 years. Only one student visited a foreign country (two weeks in the United States). Twenty students had used VR before, but mainly for games and watching videos, not for learning purposes. In accordance with the Common European Framework of Reference for Languages self-assessment (Council of Europe, 2020), the students' perceived levels of proficiency in spoken English ranged from pre-intermediate (A2) to intermediate (B1). This information about the participants was collected through a questionnaire, which is presented in Section 5.3.4.

Following the 2*2 factorial design addressed in the above section, the participants were divided into four groups of 35 students each in different learning environments (HiVR and classroom) and learning approaches (situated learning and teacher-centred learning). In order to ensure the homogeneity of each group and to alleviate the influence of extraneous factors on the experiment outcome, stratified random sampling (Dörnyei, 2007) was employed to assign the students into different intervention groups. Gender, age, years of English learning, VR experience and English speaking anxiety levels were used as stratifying criteria. The

advantage of stratified random sampling is to ensure no fixing or bias when allocating students into groups (Bryman, 2016). The details of group division and group information can be found in Section 6.1. After dividing students into four groups, they were assigned to corresponding interventions as demonstrated in the following section.

5.3.3 Interventions

The 2*2 factorial design determined the different interventions for each group in my study. It should be noted that the interventions were not strictly purposive for linguistic knowledge learning, but for oral English practice. The intention of the intervention was to decrease students' anxiety about speaking English.

Intervention topics

The topics involved in the interventions were decided according to previous literature (see Chapter 3). The intervention topics were the same for all groups. Krashen and Terrell (1983) outlined some topics and situations that are likely to be most useful to students such as ordering a meal in a restaurant, shopping in a supermarket, finding locations, as well as topics related to vacations, money, and prices, and so forth. Some adaptations of these topics and situations can be found in recent studies on High-immersion Virtual Reality (HiVR) or Low-immersion Virtual Reality (LiVR) and oral English learning. In terms of LiVR, Jung (2002) simulated shopping with friends and consulting a sales assistant. Wang, Petrina, and Feng (2017) created a role-play scenario including a hotel and a restaurant for students to conduct conversations in a virtual environment. Shih and Yang (2008) developed a scenario for complaining about the food in *Second Life*[28], and Chen (2016) generated a session of ordering food in a restaurant created in *Second Life*. Among the limited studies on HiVR and oral interaction, Alfadil (2020) simulated an airport and a café to help Arabic students learn English; Yang *et al.* (2020) involved hotel reservations and restaurant ordering in a HiVR English learning system; Wang, Guo, *et al.* (2021) created a supermarket to allow Chinese EFL learners to practise oral English.

Considering the sequence and frequency of scenarios learners might experience in real life, the airport, hotel, restaurant, and supermarket were selected as the main locations of the

[28] Second Life is an online virtual world developed by Linden Lab. It allows people to create an avatar for themselves and interact with other users.

learning sessions for my study. Three situations were involved in each location for students to experience as shown in Table 5-2. In total, 12 sessions were designed. Due to the pandemic outbreak, the time for the fieldwork was shortened because that semester started a few weeks behind schedule. A total of 9 sessions (including airport, hotel, and supermarket) were delivered to students over three months rather than 12 sessions. The restaurant location was excluded for lack of time because it was intended to be the final scenario. The fieldwork timeline is introduced in Section 5.1. To fit students' proficiency levels, the content of interventions used an intermediate level of English.

Table 5- 2 *Topics for interventions in my study*

Locations	**Topic 1**	**Topic 2**	**Topic 3**
Airport	Flight information consultation	Check-in	Security check
Hotel	Check-in	Reporting problems	Check-out
Supermarket	Shopping with a friend	Asking for assistance	Returning goods
Restaurant (Not used)	*Ordering food*	*Complaining about the food*	*Giving comments and paying the bill*

VR content development

A few commercial HiVR language systems were available when I prepared for my fieldwork in 2019 and 2020, and they did not fit my research needs. I therefore sought collaboration with a VR company through the University of Glasgow and developed some HiVR English learning content with the company's free-of-charge help. A collaboration contract was settled between the University of Glasgow and the VR company. It also demonstrated that the intellectual property of the created VR content is protected by the University of Glasgow and cannot be commercialised without permission and further discussion.

I went to the VR company and worked with the VR programme developers from November 2020 to February 2021. What I provided were the company descriptions of the learning environment, scripts of the learning activities, and all the learning materials involved in the activities including the text and voice recordings of vocabulary and dialogue, and pictures of all virtual objects and scenarios (airport, hotel, supermarket, and restaurant). The company

then built scenarios and 3D models accordingly through *Unity* and *3Ds Max*. The flow of HiVR learning content was generated via a platform developed by the company for creating VR content. I guided and supervised the entire process of building each VR learning session. In order to ensure the smoothness of learning activities, each session was iteratively tested internally between the company and me before being used for the fieldwork. All the HiVR learning sessions were stored on my personal laptop under the name of HiVR English Learning System (VRELS) for my research.

The HiVR device used in my study was a stand-alone headset, the HTC Vive Focus Plus (see Figure 5-4). The University of Glasgow purchased two units of this device. Full-body tracking was accessible in this device via the head-mounted display and two controllers, which meant that students could move, turn around and interact within the virtual environment. Moreover, due to time pressure – my fieldwork had to be completed in a three-month semester – four additional HiVR devices of the same model were borrowed from the VR company.

Figure 5- 4 *HTC Vive Focus Plus, the HiVR device used in my study*[29]

<u>Technical components in VR learning content</u>

Realistic surroundings were created for the four locations (airport, hotel, supermarket, and restaurant) in the VR learning content. Different avatars were created for the corresponding scenarios to represent the students in the HiVR English Learning System (VRELS). An

[29] Retrieved from https://enterprise.vive.com/ca/product/focus-plus/

example, the avatar for a supermarket assistant, is shown in Figure 5-5. Due to the fact that the VR learning content is to simulate scenarios in the United Kingdom for English learning purpose, the avatars were designed with a white people look as it is most Chinese people's impression of people from Western countries. Due to technical limitations in late 2019 and early 2020, the body language and facial expressions of the real person were not transformable to the avatar. However, facial and body animation was created that the avatar would move body, hands, head, mouth, and eyes with the designed movement. I am aware that the synchronisation of the nonverbal cues to avatars is achievable with current HiVR technology at the time of writing the thesis.

Interaction with the environment and virtual objects can be achieved through pointing and clicking on controllers. For example, a head of broccoli would be taken virtually in the user's hand and could be turned around as well as put into a shopping trolley by clicking the controller (see Figure 5-6). The pronunciations of words would be heard when users pointed at corresponding words or items and clicked the controllers (see Figure 5-6). A native English speaker, who was also the examiner for the speaking test (see Section 5.3.5), recorded all the words and sentences for the VRELS. Sentences were mostly used for activity instructions and sample dialogues (see Figure 5-7 for an example). Movement in the virtual environment could be achieved by walking around, but also by clicking on the indication signs (blue arrows on the floor in Figure 5-7).

Immediate feedback with signs and sounds was provided at each step in order to inform the students whether they had solved the tasks correctly and whether their interaction with the environment and virtual objects was successful. However, immediate assessment of students' utterances was not involved in VRELS, as the purpose of learning in VRELS is to help Chinese EFL learners alleviate their English speaking anxiety by removing negative feedback. This is based on the involvement of exposure-based therapy (see Chapter 4) in my learning design. Instead, positive feedback, such as badges and sound of applause would show up automatically when students complete each step of learning activities, such vocabulary learning, dialogue learning, and successfully complete a task with their partners. In addition, there was no time limit for each session.

Figure 5- 5 *The avatar of a supermarket assistant created in VRELS*

Figure 5- 6 *Vocabulary shown in orange with the sound of the pronunciation in VRELS*

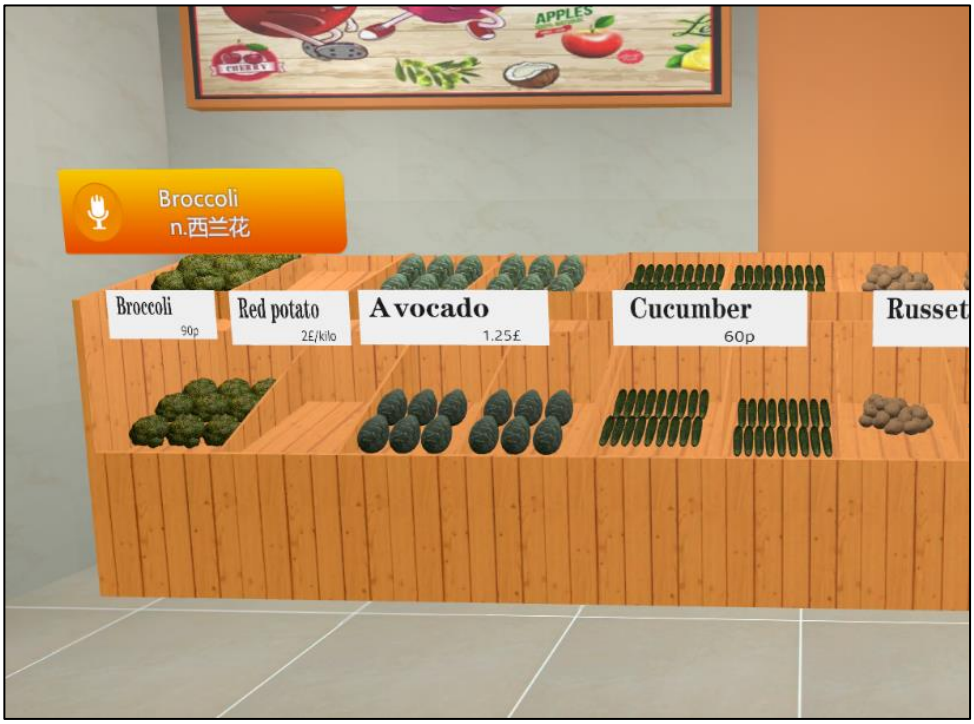


Figure 5- 7 *A sentence example in VRELS for activity instruction*

HiVR learning task flow

According to the 2*2 factorial design (see Section 5.3.1), two groups of students used HiVR to practise English speaking: a situated learning (SVR) group and a teacher-centred learning (TVR) group. The intervention for the SVR group was designed according to the characteristics of using situated learning theory as a model for learning design as proposed by Herrington and Oliver (2000) (see Section 4.2.2). The corresponding characteristics in the model are highlighted in italics in the following paragraphs for understanding.

Before the actual intervention, instruction on the VRELS was provided in both video and text versions to students in HiVR groups (SVR and TVR). Students were allowed to ask any technical questions while using HiVR in the intervention. However, few students were confused by the VRELS, but could start to use it quickly. This may be due to the fact that the university students were born in the Z-era, which means they are often confident users of new technology (Turner, 2015). In the pair-work learning activities for the SVR group, two students were in the same room and were exposed to the HiVR environment through a stand-alone VR headset. Students were paired with partners from different classes (11 classes in total), which means that they were not familiar with their partners in the

intervention. This purposive arrangement was to ensure the arousal of speaking anxiety, as being familiar with the partner may not cause students' anxiety while speaking the foreign language (see Section 2.3.3).

The simulation of the real-world scenarios provided an *authentic context* to show how the conversation would proceed in real life. A sample task of asking for assistance in the supermarket is shown in Figure 5-8. Since my study involved pair-work role-plays for the SVR group, two different characters (Player A and Player B) were created for each session, and students could choose through discussion by themselves at the beginning of the task when wearing the VR headset. After clicking on Player A or Player B, a similar scenario with slightly different information would be presented to students.

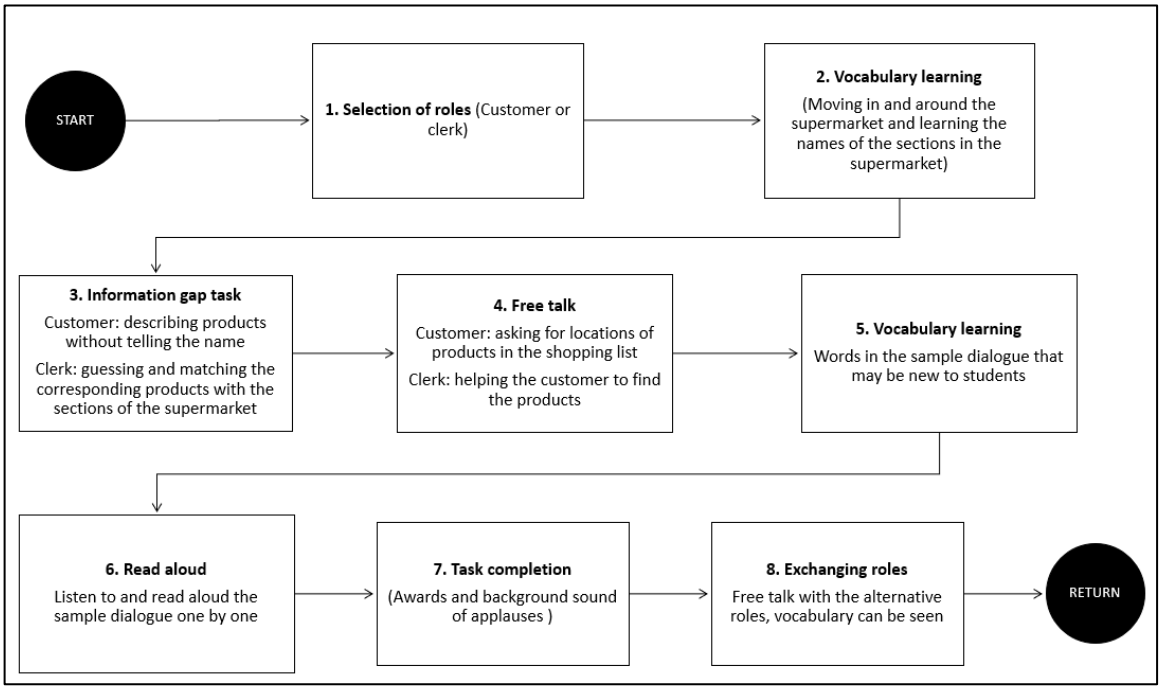
Role-play tasks underlies the task-based teaching approach, in which students need to accomplish an activity through using the language (Ellis, 2003). In role-play tasks, the language learners are exposed to cases or situations in a real or artificial environment, and they have to exhibit in form of roles (Dorathy & Mahalakshmi, 2011). Different social contexts and different social roles can be provided in role-play tasks, which fits the purpose of the current study that providing *authentic activities* for oral English practice.

After selecting roles, students first learn some vocabulary and then move to an information gap task. An information gap task is introduced by Long (1980) to address questions on input and interaction. In an information gap task, language learners do not have the information they need to complete a task and need to communicate with each other to find it (Pica, Kang, & Sauro, 2006). The information gap tasks meet the *collaborative learning*, which is a characteristic of learning environment designed using situated learning theory. In the scenario of asking for assistance in the supermarket, the customer is required to describe the product without mentioning its name. The clerk needs to guess the name and match it with the section of the supermarket where the product can be found.

After completing the information gap task, students then move to a free talk activity. Here they receive a brief guidance about what they have to talk about in the conversation, but they can also talk about anything relevant. Therefore, the free talk activity involved in my study is semi-structured. For example, in the task shown in Figure 5-8, the customer's mission specifies the products they would like to buy from a shopping list. After this, students learn some vocabulary new to them in the sample dialogue and then move to a read-aloud activity.

The sample dialogue learning is *access to expert performance* and *scaffolding* as highlighted in the model of learning design using situated learning theory. In the read-aloud activity, students listen to and read aloud the sample dialogue sentence by sentence. After completing the whole task, students can exchange their roles and have a conversation using the vocabulary provided. However, they are not required go through the whole learning task again but can just hold a conversation using vocabulary or expressions given in the virtual environment.

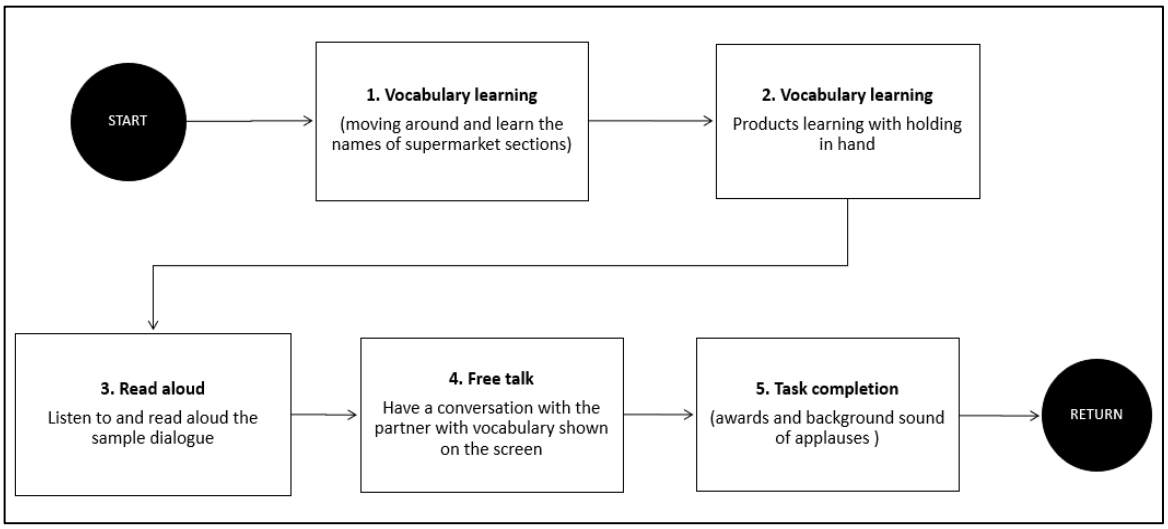
Figure 5- 8 *Task sample for SVR group in VRELS (devised by me)*



The HiVR learning in the TVR (teacher-centred learning + HiVR) group also involves two students in each session, but the VR content is delivered in a traditional way, where students follow the guidance to learn the vocabulary and sample dialogues. Most of the learning process in the TVR group involves clicking on the vocabulary and sentences to listen to them rather than interacting with the virtual environment and communicating with a partner. In other words, vocabulary learning, and dialogue read-aloud are the main forms of learning for the TVR group. A free talk activity with role-play tasks is provided at the end of each session for students to practise what they have learnt in the session. Figure 5-9 shows a sample of learning flow for the TVR group. In the scenario of asking for assistance in a supermarket, students first learn the vocabulary of the supermarket sections and some products. They then listen to a dialogue sample and are required to read it aloud sentence by

sentence. All these steps were conducted individually. The pair of students then discussed which role they would like to take and had a conversation according to the guidance shown in the virtual environment.

Figure 5- 9 *Task sample for TVR group in VRELS (devised by me)*



In the interventions for both the SVR and TVR groups, each learning session contained two to three learning tasks as shown in Figure 5-8 or Figure 5-9. Students could take a break after each task. One task lasted about 10 to 15 minutes, so one session was about 30 to 45 minutes. During the intervention, I was an observer, ensuring students' safety while using HiVR, helping if technical issues occurred, and explaining the task if students did not understand; I did not join in with the learning activities or the conversation.

Classroom groups intervention

There were two other groups of students involved in the classroom teaching: underlying situated learning theory (SC) and the teacher-centred learning approach (TC). Students in the SC (situated learning + classroom) group followed the pair-work procedure designed for the SVR group, but this time in a traditional classroom with me explaining the task and giving out paper copies of the activity's guidelines or pictures and vocabulary. Students first selected the role they would like to take, then I gave them the corresponding learning materials (see Figure 5-10 for an example). They learnt the vocabulary first with me and had a role-play conversation (free talk) according to the guidelines. After completing the conversation, I would not correct them directly but show them a sample dialogue on the screen in the classroom (see Figure 5-11 for an example). The students needed to listen to

and read the sample dialogue aloud following audio recordings provided by an English native speaker who was also the examiner in the speaking test (see Section 5.3.5).

Figure 5- 10 *An example of learning material for SC group (devised by me)*

Supermarket 2-Assistant

Free talk 1

1. Vocabularies

Aisle n.过道

Unfortunately adv.不幸地

Produce n.农产品

Alternative adj.可选择的

Fresh adj.新鲜的

Delivery n.运送

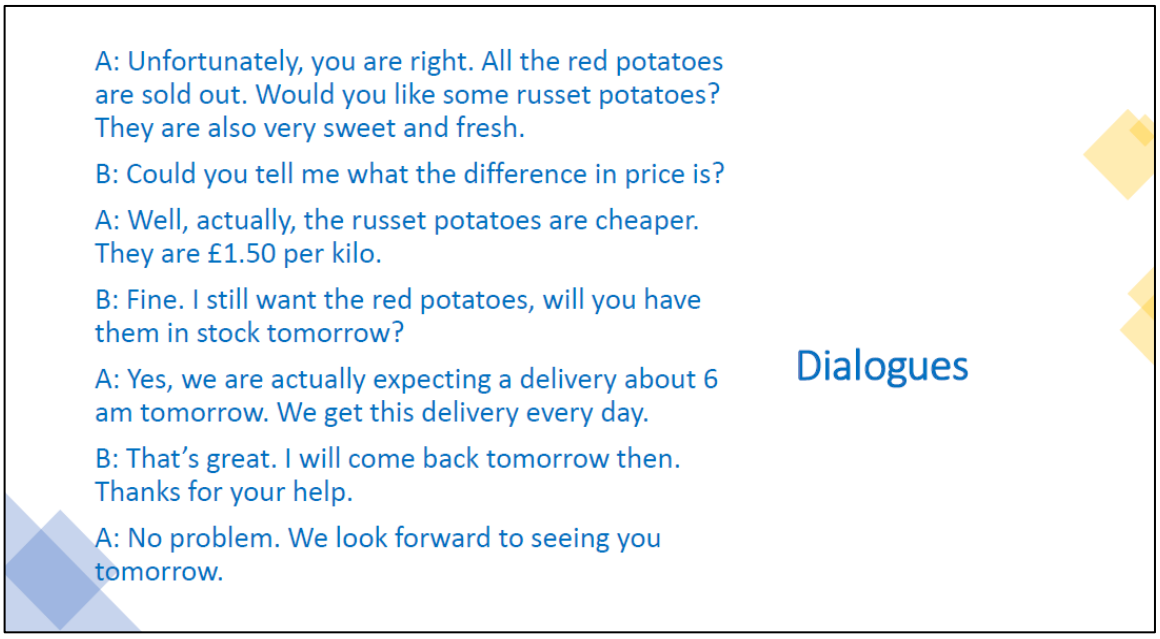
2. Conversation guidelines

a. Tell the customer where the items could be found

b. Offer an alternative option to the customer (Russet potato: £1.25-available, red potato: £1.50-unavailable)

c. Vegetables are delivered every day

Figure 5- 11 *A sample dialogue shown for SC group (devised by me)*



For students in the TC group, the lessons were delivered as a lecture where the students listened to me introducing the vocabulary and explaining the sample dialogue. Similar to the learning flow for the TVR group, a free-talk task was given at the end of each session. All learning materials were presented on PowerPoint (see Figure 5-12 for an example) and showed to the students in the classroom via the screen. To maintain consistency of intervention for each group in the experiment, only external differences were controlled in the learning environment and teaching methods, and students from the SC and TC groups also attended learning sessions two by two as pairs. Each learning session for traditional classroom groups (SC and TC) lasted 30 to 45 minutes. In addition, students who participated in the classroom groups were permitted to try the HiVR intervention after the experiment if they wish.

Figure 5- 12 *Vocabulary learning examples for TC group (devised by me)*

5.3.4 Questionnaires

Adapting the questionnaire

Questionnaires were employed to collect data for RQ1 and RQ2 in my study. Questionnaires can be used to gather a large amount of information quickly and easily (Dörnyei, 2007). The questionnaire employed in the present study consisted of four parts: English learning and using experience (Section A), English speaking proficiency self-assessment (Section B),

technical experience on VR (Section C), and English Speaking Anxiety Scale for Chinese English learners (Section D) (see Appendix 1). The first part of the questionnaire collects background information about participants' English learning experience. The second section of the questionnaire gathers students' self-perceived proficiency in English speaking according to the Common European Framework of Reference for Language. The third section asks about students' technical experience in VR. The last part is the scale used to measure FLSA levels, which contains 27 five-point Likert-type statements.

The development of the 27-item English Speaking Anxiety Scale (ESAS) follows the four categories for FLSA among Chinese English learners summarised in Section 2.3: low English proficiency (items 1–8), lack of self-confidence (items 9–18), lack of experience (items 19–24), and unwillingness to communicate (items 25–27). All the statements were adapted through combining previous measurements, as no scale fitted the exact setting of the present study. Some items were adapted from the Foreign Language Classroom Anxiety Scale (Horwitz, Horwitz, & Cope, 1986) and the adaption by Mahmoodzadeh (2012), although it is focused on classroom circumstances. Some parts of the Foreign Language Speaking Anxiety Scale (He, 2011) were also adapted for my study. Some items on the ESAS were adapted from scales proposed at the early stage of foreign language anxiety investigation, such as Unwillingness to Communicate Scale (Burgoon, 1976) and Language Class Discomfort Scale (Ely, 1986). Oral English Classroom Anxiety Scale for Chinese Learners (Li, 2016) was also taken into consideration, as it includes some cultural impact (e.g. loss of face) among Chinese students. All these measurements of foreign language anxiety or FLSA were reviewed in Section 2.2. Specifically, the ESAS adapts statements from various measurements that are shown in the Table 5-3.

Table 5- 3 *The adaptation of items in ESAS*

Items in ESAS	**Adapted from**
Items 1, 2, 3, 4, 7, 8 and 12	Speaking Anxiety Questionnaire (Mahmoodzadeh, 2012)
Items 5, 17, 19, 21, 25 and 26	Foreign Language Speaking Anxiety Scale (He, 2011)
Items 6 and 10	Foreign Language Classroom Anxiety Scale (Horwitz, 1986)
Items 13	Foreign Language Speaking Anxiety Scale (Balemir, 2009)
Items 9, 11, 14, 22, 23 and 24	Foreign Language Speaking Anxiety Scale (Melouah, 2013)
Items 15 and 16	Oral English Classroom Anxiety Scale for Chinese Learners of English (Li, 2016)
Items 18 and 20	Language Class Discomfort Scale (Ely, 1986)
Item 27	Unwillingness to Communicate Scale (Burgoon, 1976)

As stated above, items in ESAS were borrowed from the existing questionnaires (see Chapter 2). In order to strive towards the validity and reliability of the questionnaire, I piloted it in three steps: peer and expert review, back-translation and collaborative translation, and piloting.

Peer and expert review of the questionnaire

At first, in order to detect potential problems with the questionnaire and reduce errors in data-gathering, the questionnaire was sent to four of my peers, who were Chinese PhD students at the University of Glasgow. Two of them were based at the School of Education and have knowledge in second language acquisition. The other two were from the School of Business and have no expertise in the field of second language acquisition but have knowledge of questionnaire design. The four peers gave their feedback from both inside and

outside the research field, first imagining themselves as participants then thinking about how to help students understand the questionnaire. My colleagues also asked themselves what questions they would ask in the questionnaire if they were the researcher of the present study. They went through each item of the initial questionnaire and gave me written feedback. I discussed their comments with them by email or phone before finalising a version of the questionnaire that incorporated all their feedback.

The peer review helped to refine some items in the questionnaire. The question about perceptions of VR and language learning was removed, because the participants might gain no knowledge of it before the intervention. The peer review also helped to shorten the ESAS from 34 items to 27 items (see Table 5-4 for deleted items). "I feel that not knowing enough vocabulary is the biggest problem preventing me from speaking English easily" was deleted as it is similar to "I become anxious when I get stuck on one or two words in speaking English". They are both about vocabulary and speaking anxiety. The statement "I feel overwhelmed by the number of grammar rules I have to learn to speak English" was deleted as it seemed more about language learning than speaking anxiety. The statements "The more I fail to speak English fluently, the more disappointed I get" and "I get anxious when I find I cannot speak English fluently" were considered by some of my colleagues to have a similar meaning. "I feel shy when I speak English in front of others" and "I don't mind thinking aloud in English, but I feel very uncomfortable when I have to speak to others in it" were removed as they overlapped with the statement "I feel very self-conscious about speaking English in front of others". The other two statements "I am not willing to speak in English when I don't know the scheduled oral activities" and "I am not willing to get involved in the speaking tasks when the topics are not interesting" were assessed by peers as not referring to anxiety towards completing oral tasks, but just unwillingness to communicate.

Feedback from a second language acquisition expert in my supervisory team was sought after the peer review. The question asking participants to give their ages was changed to a multiple-choice question with a range of ages provided, because some students might be reluctant to admit their exact age. The expert also suggested me use the Common European Framework of Reference for Language self-assessment descriptor for speaking proficiency self-assessment as there is no official speaking test for Chinese university students to gain data about their speaking proficiency. With all peers' and expert's suggestions, an original English version of the questionnaire (see Appendix 1) was set to be translated into Chinese.

Table 5- 4 *Items deleted from the ESAS after reviewing*

Deleted item	Statements
1	I feel that not knowing enough vocabulary is the biggest problem preventing me from speaking English easily
2	I feel overwhelmed by the number of grammar rules I have to learn to speak English
3	The more I fail to speak English fluently, the more disappointed I get
4	I feel shy when I speak English in front of others
5	I don't mind thinking aloud in English, but I feel very uncomfortable when I have to speak to others in it
6	I am not willing to speak in English when I don't know the scheduled oral activities
7	I am not willing to get involved in the speaking tasks when the topics are not interesting

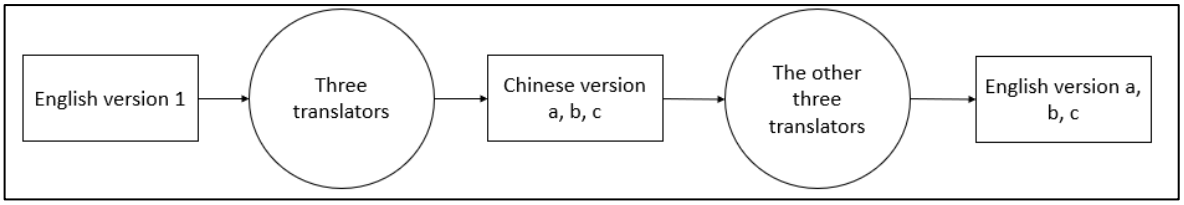
Back-translation and collaborative translation of the questionnaire

In order to ensure the comprehensibility of the questionnaire and minimise possible misunderstanding caused by a foreign language, three steps were taken to develop a final translated version of the questionnaire: back-translation, collaborative translation, and piloting (Douglas & Craig, 2007). Not single approach uncovers every problem, and piloting could show if respondents understood all the material (Brislin & Freimanis, 2001).

Back-translation and the committee translation approach were carried out to ensure the quality of translation. Back-translation involves translating from the original language to the target language then translating the text back to the original language (Chen & Boore, 2010). In the committee approach to translation, a group of people check the translation collaboratively and iteratively (Douglas & Craig, 2007). Six masters' students who had graduated with a translation major at a Chinese university (the translators) and an expert who had 20 years' experience in both questionnaire design and Chinese and English translation were involved in the back-translation and committee translation approach.

In back-translation, I prepared the original English version of the questionnaire for three of the translators to translate into Chinese separately. Next, I sent each of the other three translators one of these three Chinese versions of the questionnaire, and they translated the questionnaire back to English without reference to the original English version. Figure 5-13 illustrates the procedure of back-translation.

Figure 5- 13 *The back-translation procedure in my study*



Back-translation revealed discrepancies between the final English versions (a, b, c) and the original text (English version 1), especially in Sections B and D. This is not a flaw in the method, but a discovery of statements that are not appropriately or accurately translated (Brislin & Freimanis, 2001). I then analysed the English versions (a, b, c) and identified the discrepancies between them and the original one. This determined what changes should be made to the Chinese version.

Following the committee approach to translation (Douglas & Craig, 2007), a committee-based review meeting was then held with the translators, the translation expert and me. We discussed all the discrepancies identified and collaboratively decided on a version for final review. After the review meeting, I acted as the adjudicator and finalised a translated version for review. A second meeting was held to decide whether the translated version was acceptable. Having incorporated their further suggestions, a final version of the questionnaire was then ready to be tested.

Piloting the questionnaire

Piloting a questionnaire is intended to ensure that all items are fully understood by the participants and test the validity and reliability of the questionnaire (Dörnyei & Taguchi, 2009). A group of 19 second-year English-major students (14 females and 5 males) in a university other than the one used for the main study were recruited to take part in piloting the questionnaire. This group of students could be regarded as representatives of the target population in the main study, because they were also English-major students in Chinese

universities from the same province of main study site. Another reason for choosing students from a different university was to make sure the target population would not be affected, for instance by getting an idea of the questionnaire before the main study, which might influence their responses (Ikart, 2018).

Before the participants in the pilot study answered the questionnaire, they were informed of the nature of the pilot study being undertaken and the importance of their comments on each item, and they received a guarantee of anonymity and confidentiality in regard to the use of the information they provided. A participant information sheet (see Section 5.9) was sent to the participants and a consent form (see Section 5.9) was signed by each student. On the basis of students' responses, "speaking English" was reworded as "speaking English with others" (or added to items 3, 18, 20) in the English Speaking Anxiety Scale (ESAS) to make the statements more closely related to the purpose of my study and more comprehensible to the participants. "Tasks" in items 15 and 24 was changed to "interaction tasks" for the same reasons. All these revisions went through the same procedures of back-translation and committee approach to translation as described above and were discussed with my supervisory team. A Chinese version of the questionnaire was finalised and is shown in Appendix 2.

Filling out the questionnaire

The entire questionnaire was transferred to an online version using Microsoft Forms. The link to the questionnaire was sent to each student via QQ (a Chinese social media application) before the intervention. After the intervention, only Section D, English Speaking Anxiety Scale (ESAS) for Chinese English Learners, was sent out to collect students' general FLSA levels. All 140 students responded to these questionnaires before and after the intervention.

5.3.5 English Speaking Test

Retrospective self-report surveys may not be able to detect language anxiety due to the dynamic nature of this emotion (Gregersen, MacIntyre, & Olsen, 2017). For example, anxious students may enjoy some moments of speaking the foreign language, and non-anxious students may feel anxious at times during foreign language communication, such as suddenly struggling to remember a vocabulary item (MacIntyre & Legatto, 2011). Thus, the purpose of the English speaking test is to collect data on students' performance in oral tasks.

The video recordings of their oral performance in the speaking tests would be utilised to obtain data about students' FLSA levels in specific tasks. The design, piloting and implementation of the speaking test are illustrated in this section.

Developing the speaking test

As the aim of my study was to investigate students' FLSA during interactions, a two-way conversation was used for the speaking test rather than a monologue. Specifically, role-play tasks were employed in the speaking test, this is consistent with the type of tasks used for the intervention in situated learning groups (SVR and SC). Role-play tasks would simulate various communication situations that the test-takers might meet outside the test (Luoma, 2004). In a speaking test, the role-play tasks can be taken by two examinees or between an examinee and an examiner (Luoma, 2004). Due to the fact that an interlocutor may influence both a student's oral performance in a two-way conversation and the student's anxiety levels, the role-play task was taken between one English native speaker (the examiner) and a student (the examinee) in the speaking test rather than between two students in order to ensure the equivalence of each test. Since it is a role-play task, the examinee and the examiner both would take roles and a new simulated relationship with each other to complete a communication (Luoma, 2004). Specifically, role-play tasks from the American Council on the Teaching of Foreign Languages Oral Proficiency Interview were adapted for the speaking test. The Oral Proficiency Interview is a research-backed valid test for spontaneous unrehearsed language ability assessment (Tschirner, Bärenfänger, & Wanner, 2012). From the examiner's perspective, the steps for conducting an Oral Proficiency Interview are[30]:

1. Introduction: confirm the examinee's identity; read a statement in English introducing the interview process
2. Warm-up: discuss a few general topics
3. Interview: explore the highest proficiency level of the examinee; explore a level above the language proficiency in order to gather evidence that the examinee cannot sustain performance at that level

[30] Retrieved from https://www.actfl.org/assessment-research-and-development/actfl-assessments/actfl-postsecondary-assessments/oral-proficiency-interview-opi

4. Cool-down: ask a few questions that the examinee can easily perform; end the interview

The intermediate level was selected for the speaking test employed in my study based on participants perceived spoken English proficiency level (A2 to B1). The intermediate Oral Proficiency Interview includes some simple topics related to daily activities. The original 11 role-play prompts for the intermediate Oral Proficiency Interview can be retrieved from the website. Due to the fact that my study aimed to examine students' FLSA rather than their oral performance, there are no criteria for marking it. The way students themselves rating their FLSA level is illustrated in Section 5.3.6.

Piloting the speaking test

In order to investigate the feasibility of the Oral Proficiency Interview with Chinese EFL learners, it was piloted at the university where the questionnaire was piloted (see Section 5.3.4.4). A total of 18 second-year English major students (13 females and 5 males) who had taken the questionnaire (19 students in total) were involved. One student did not make the speaking test because it conflicted with another exam.

Based on the Oral Proficiency Interview, a one-minute preparation was added into the test procedure, as it might be difficult for the participants to speak immediately after being introduced to the task. This minute allowed them to think about the question and prepare for the role-play tasks. The role-play conversation was set to around five minutes. This is based on the study of Gregersen (2005), in which the first four minutes of a foreign language presentation was used to evaluate student anxiety levels. With regard to the prompts, *city or country* was specified as *a British town*, and some specific situations, such as the shopping mall, music festival, renting a property, were fixed rather than an open choice. The task in the restaurant was removed from the original version of topics in the Oral Proficiency Interview because a similar situation was designed for the intervention. It was removed due to the familiarity of the task might influence students' anxiety levels, although the restaurant scenario was not implemented in the intervention at the end.

The speaking test procedure of the Oral Proficiency Interview was adapted as follows in my study:

1. Ice breaker: greet the student and confirm identity; introduce the task and show the prompt
2. One-minute preparation: allow the student to understand and prepare for the task (pen and paper provided)
3. Role-play: have a conversation with the student for around five minutes
4. Close the conversation

A video camera facing the students recorded the speaking test. Each student would watch their own video-recordings of the speaking tests and rate their FLSA levels during it. Audio recordings were made on a smartphone during the speaking tests, in case the camera did not work, or the data extracted from the camera had problems.

No information about the speaking tasks content was given prior to the test in order to avoid students' preparation before the test and to ensure their answers are spontaneous during the test. At the beginning of the speaking test, the examinees selected a number from 1 to 10 randomly for the prompts they would take for the role-play task. After selecting one number, the corresponding task was removed from the rest prompts for the next student in order to make sure that each of the 10 role-play tasks are piloted. The duration of the speaking test was intended to be around five minutes. The day after the test, each student was briefly interviewed about their perceptions of the speaking test. The following three questions were asked in the interviews:

1. What do you think about the difficulty of the speaking test?
2. What do you think about the length of the test?
3. Do you have any other comments on the test?

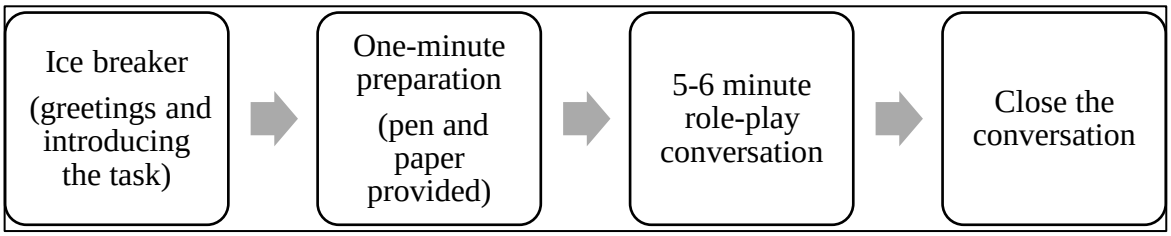
The pilot testing lasted three hours in total, and the length of each student's test ranged from 3 minutes 40 seconds to 6 minutes 15 seconds, with an average of 4 minutes and 53 seconds. Ten students found the length of the test was fine, but six students thought it was too short, and one student thought it was too long. The recordings of students who thought the test was too short were double-checked and were found to be less than five minutes in duration. The role-play conversation for the main study was therefore set to last between five minutes and six minutes.

Although one of the students had difficulty understanding the tasks, the rest perceived the task difficulty as reasonable. The aim of the speaking test was to analyse anxiety levels rather than investigate their oral proficiency, so the task difficulty was not analysed through score distribution, and it was not necessary to calculate the intercorrelation and reliability of scores (Liao & Wei, 2010). To avoid students not understanding the task type in the main study, a brief introduction to role-play tasks, though not detailed task information was provided to students in Chinese before the test. It was also found that most students found it harder to ask questions than answer questions. So, the role taking in the prompts were switched to offer more opportunities for students to speak by answering questions (see Appendix 3). The wording in the prompts was slightly changed for better understanding of the task. Another problem that occurred in the pilot study was that the examiner sometimes interrupted the students unintentionally, which caused students to pause. The solution to this was that the examiner would only speak when she was sure the student had finished their utterances. For example, the examiner would wait for a couple of seconds before her turn in conversation to see whether students finished their utterance or not.

Conducting the speaking test

In the main study, the speaking tests were conducted twice on each student during the experiment, before (pre-test) and after (post-test) the intervention. The tests were carried out in the same classroom at the University each time. The examiner was the same English native speaker as introduced in the above section. One camera was set up facing students to record their performance in the speaking tests. As with the pilot study, students randomly selected a number from 1 to 10 before entering the room in order to avoid them knowing the prompt prior to taking the test. The same speaking prompts were used for the pre-test and the post-test, but for each student the post-test task was different from the pre-test one. This is because task repetition may influence oral performance and anxiety levels during taking the test (Amiryousefi, 2016). The test procedure is shown in Figure 5-14 below.

Figure 5- 14 *The speaking test procedure in my study*



5.3.6 Self-rating of FLSA Levels in Specific Tasks

Based on the dynamic perspective of FLSA (see Chapter 2), a self-rating method was employed to measure FLSA levels in the speaking test. This method was inspired by the *anxometer*, which was developed by MacIntyre and Gardner (1991a) and adapted from the idiodynamic approach created by MacIntyre and Legatto (2011). The *anxometer* consists of a scale from one to ten printed on a thermometer figure (MacIntyre & Gardner, 1991a). The idiodynamic approach is realised through software that presented video recordings of students' performance in a speaking task and required them to rate their anxiety levels from -5 to +5 each second. A detailed introduction to these two approaches can be found in Section 2.2.2.

The scale I adapted for self-rating of FLSA levels in specific tasks was an 8-item 10-point scale (see Appendix 4), in which 1 means not anxious at all, while 10 means very anxious. As mentioned in Section 5.3.5, the speaking test was recorded via the camera and the first four minutes of it would be used for data analysis in order to answer RQ1 and RQ2.I therefore divided the four-minute task into eight sections of 30 seconds each. The scale was translated into Chinese, and an online version was created via Microsoft Forms (see Appendix 5). Students watched the video recording of their performance immediately after the test and rated their anxiety levels using the scale. The responses of the 140 participants were gathered for both pre-test and post-test.

5.4 Qualitative Data Collection (RQ 3, RQ 4 and RQ 5)

After the second speaking test (post-test), the students' (N = 140) perceptions were collected via open-ended questionnaires in order to answer RQ3, RQ4, and RQ5. The open-ended questionnaire included five questions. The first three questions were for all participants and focused on students' perceived FLSA, and the last two were just for HiVR groups (SVR and

TVR), focusing on students' perceptions of HiVR and FLSA as well as their experience in using HiVR for oral English practice. The open-ended questions and the corresponding Research Questions (RQs) they aimed to collect data for is shown in Table 5-5.

Unlike closed-ended questions, open-ended questions do not provide answer choices, but allow participants to respond in their own words (Albudaiwi, 2017). Questionnaires containing open-ended questions are often used to collect qualitative data. The first reason for using the open-ended questionnaire in my study was that a closed-ended questionnaire might not completely represent participants' thoughts (Albudaiwi, 2017). Using open-ended questions can enrich the diversity of answers, although it requires extensive coding (Reja *et al.*, 2003).

With regard to the questionnaire adapted for general FLSA measures (ESAS, see Section 5.3.4) in my study, participants were just able to choose *strongly agree (5), agree (4), neither agree or disagree (3), disagree (2),* or *strongly disagree (1)* to the statements provided, and their opinions other than these could not be expressed. Therefore, two open-ended questions asking about students' FLSA before and after the intervention were added in order to gather richer data that could widen my understanding of FLSA among Chinese EFL learners as well as the influence of interventions on their FLSA levels. Similarly, the third question asks about students' actual experience of anxiety during the speaking tests. The two questions on the impact of HiVR on FLSA and the experience of using HiVR for foreign language learning were designed to be easily understood in order to avoid irrelevant responses.

Table 5- 5 *Questions in the open-ended questionnaire and corresponding research questions*

Open-ended questions	Corresponding RQs
1. *Please describe your anxiety about speaking English with others BEFORE the learning sessions.*	**RQ 3:** How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?
2. *Please describe your anxiety about speaking English with others AFTER the learning sessions.*	
3. *In fact, did you feel more at ease in the second speaking test compared to the first one? What made you feel this way?*	
4. *What do you think was the impact of immersive VR on your anxiety about speaking English?*	**RQ 4:** What are students' perceptions of the impact of HiVR on FLSA?
5. *In your experience, what is your opinion of the advantages and disadvantages of using immersive VR for the oral English course?*	**RQ 5:** What are students' experiences of using HiVR in oral English learning?

Another advantage of an open-ended questionnaire is that it brings a sense of individuality to each participant (Allen, 2017). Allen (2017) stated that the ability to express opinions would be improved when receiving more personalised opportunities in an investigation. While interviews can also provide a sense of individuality, it would be very time-consuming in my study to interview 140 participants. A single open-ended questionnaire could be sent via the Internet to all participants at once, which saved a lot of time.

One of the disadvantages of using an open-ended questionnaire is low quality responses (Reja *et al.*, 2003). It was found by Couper (2001) that making a longer text box available in an online questionnaire resulted in a higher number of characters being typed in. In order to ensure the quality of the responses, a minimum of 30 Chinese characters were set to the response box. I translated the open-ended questions myself because the questions are easy

to understand. A Chinese version of the open-ended questionnaire can be found in Appendix 6.

5.5 Quantitative Data Analysis

As mentioned in Section 5.2, the qualitative data in my study was used to compare with and/or explain the quantitative results. Therefore, the quantitative data was analysed before the qualitative data, and the results of the two data sets were integrated for a comprehensive answer to the research questions. The quantitative data analysis methods and procedure are introduced in the following sections.

5.5.1 Analysing Questionnaires

Data screening and preparing

SPSS 27 and SPSS AMOS 28 were used to analyse quantitative data in my study. Some responses to the questionnaire (see Section 5.3.4) were first transformed into SPSS. With regard to gender, 1 was used to represent female and 2 to represent male. Section A of the questionnaire asked about the time students started to learn English, and the response was entered as how many years they have been learning English. Using the Common European Framework of Reference for Languages self-assessment of speaking proficiency (Section B), A1 was entered as 1, A2 as 2, B1 as 3, and so forth. Therefore, a higher score refers to a higher level of perceived oral English proficiency. In terms of the VR experience (Section C), 1 refers to the answer Yes, and 2 refers to the answer No. The demographic information sections of the questionnaire (Section A, B, and C) were used for group division as discussed in Section 5.3.2.

Section D of the questionnaire – the English Speaking Anxiety Scale (ESAS) of Chinese English Learners – measures participants' FLSA levels retrospectively. It is a 27-item 5-point Likert scale, and responses were transformed to scores as follows: Strongly disagree – 1; Disagree – 2; Neither agree or disagree – 3; Agree – 4; Strongly agree – 5. The sum of the students' responses to ESAS then could be calculated as their FLSA level. Smaller scores mean less anxiety and larger scores indicate more anxiety.

The questionnaire responses were first exported as a Microsoft Excel document from Microsoft Forms. In order to ensure there were no mistakes when entering the data, I used the replace function in Microsoft Excel to transform responses into numbers automatically. I then double checked all data with the original version of the responses from participants. Afterwards the dataset was copied into SPSS for statistical analysis.

Normality

Checking normality of or the distribution of data is important before starting analysing the data, because it determines whether parametric or nonparametric statistical techniques should be utilised for data analysis (Woodrow, 2014). If the data was normally distributed, parametric analysis methods should be used (Bryman, 2016). The nonparametric approach could be used to achieve a similar analysis if the data was not normally distributed (Bryman, 2016). Skewness and kurtosis are often used for normality checks in SPSS. In my study, the range of ± 2 was applied to skewness value as a normal distribution, and ± 7 for kurtosis as indicated by Hair (2009). A scatter plot and a histogram could also be used to assess whether data distribution was normal. Details of normality checks can be found in Chapter 6.

Reliability

The reliability of a scale refers to the consistency and dependability of the measurements (Faul, 2018). It concerns whether or not similar results would be obtained when using the scale to measure the same construct multiple times. There are many types of reliability that could be estimated in different ways, such as inter-rater reliability, test-retest reliability, split-half reliability and internal consistency reliability (Faul, 2018). In terms of the ESAS, internal consistency was examined because it reveals the consistency between different items in the scale. The most widely used estimate for internal consistency reliability is given by Cronbach's alpha coefficient (Cronbach, 1951). Cronbach's alpha has a theoretical range from 0.0 to 1.0. According to Pallant (2020), for a multi-item scale, the value of Cronbach's alpha above .7 is acceptable, and above .8 is preferable.

Validity

Validity of a scale concerns the accuracy of the measurement (Faul, 2018). In other words, the validity of a scale examines to what extent the scale measures what it is supposed to

measure. The main types of validity are content validity, criterion validity and construct validity (Pallant, 2020). Content validity refers to the degree to which a scale reflects the domain of the content it is intended to measure (Carmines & Zeller, 1979). This is usually ensured during the development of the scale rather than by evaluating it after construction. My method for adapting the ESAS is presented in Section 5.3.4. Following suggestions for ensuring content validity by Zeller (1979), I first summarised the four domains of FLSA (lack of experience, lack of self-confidence, low English proficiency, and unwillingness to communicate) that are relevant to Chinese EFL learners based on examining previous literature. Second, I specified the items related to the corresponding domains of FLSA from existing scales and adapted them for my study by rewording the statements. However, in social science research, it is not possible to sample content. Instead, it is necessary to formulate a set of items that have the potential to reflect the content of a given theoretical concept (Zeller, 1979). Thus, there seems to be no criterion for assessing the content validity of ESAS, but the content validity was ensured while developing the Likert scale.

The construct validity of a scale is measurable, however. Construct validity is concerned with the extent to which a scale adequately represents the underlying theory for the concept or construct that it is intended to measure (Faul, 2018). Confirmatory factor analysis can be used to investigate construct validity (also known as internal validity) of a measurement. It investigates the relationship between observed measures or *indicators* (e.g. items in the scale) and latent variables or *factors* (Brown & Moore, 2012). "A factor is an unobservable variable that influence more than one observed measure and which accounts for the correlations among observed measures" (Brown & Moore, 2012, p. 2). In other words, confirmatory factor analysis could reveal the commonalities within the data through finding clusters of intercorrelated observed variables, which can help understand the validity of the construct (Atkinson *et al.*, 2011). This is because if items correlate highly with each other, it means that they have potential to measure the same construct as they share a common cause (Keith, 2019).

Confirmatory factor analysis is a main type of factor analysis, which is a reduction technique to uncover a smaller number of latent variables (factors) by studying the variation and covariation among a set of observed variables (Long, 1983). When applying confirmatory factor analysis as an analytic tool for construct validation, convergent and discriminant validity would both be assessed for a set of related constructs. Convergent validity refers to

the closeness with which a measure relates to (or converges on) the construct that it is purported to measure, and discriminant validity refers to the degree to which a measure does not measure (or discriminates from) other constructs that it is not supposed to measure (Brown & Moore, 2012). Different from other analytic methods, measurement error can be estimated in confirmatory factor analysis (Brown & Moore, 2012).

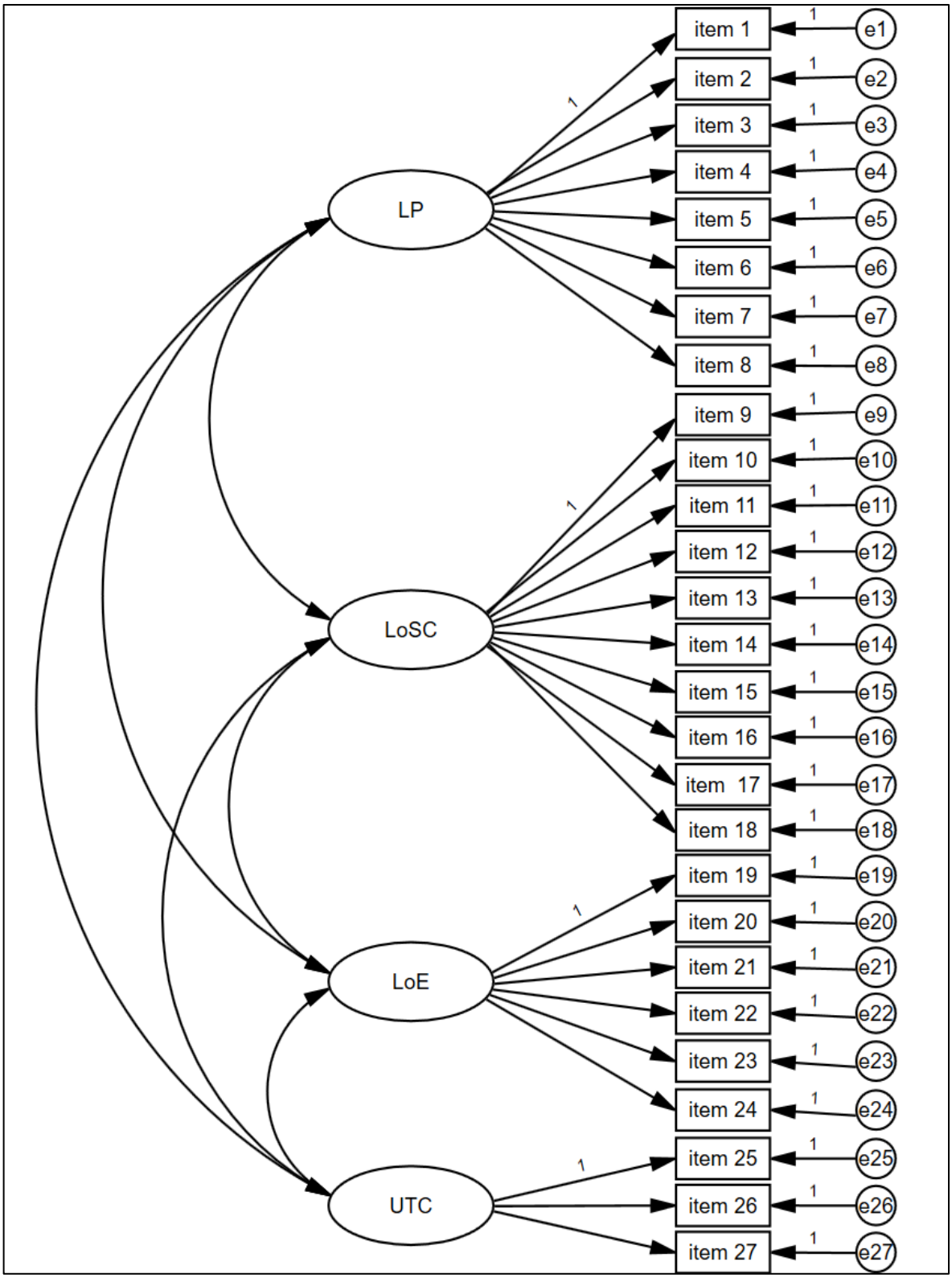
Other than confirmatory factor analysis, another main type of factor analysis is exploratory factor analysis. Exploratory factor analysis is often used to explore the interrelationship among a set of variables when no specifications are made in regard to the number of factors at the early stage of the research (Ockey, 2013). Exploratory factor analysis is most often used to determine the number of factors that can account for items of an instrument (Ockey, 2013). However, confirmatory factor analysis is usually used when the number of factors is specified in advance and the underlying structure has been established (Pallant, 2020). In my case, exploratory factor analysis might help to identify potential factors for the 27-item ESAS rather than the four-factor construct I developed based on the previous literature. Barlett's test of sphericity and Kaiser-Meyer-Olkin were conducted prior to exploratory factor analysis to assess the suitability of the dataset for factor analysis. The significance of Barlett's test of sphericity should be smaller than .05 and the recommended Kaiser-Meyer-Olkin value is above .6 (Kaiser, 1970). The process of conducting exploratory factor analysis is as follows (Ockey, 2013):

a. Screening and preparing the data
b. Extracting the factors
c. Retain the number of factors for the solution
d. Rotating steps b and c
e. Interpreting the solution.

In step b, eigenvalues exceeding 1 were often extracted as factors through principal components analysis, and the percentage that each factor explains the variance could be found. However, if the cumulative explanation of the factors to the whole dataset is no lower than 60%, the correlation between each item and the corresponding factor should be examined (Hair, 2009). The cue-off point for factor loading is .03 (Stevens, 2012). Scree test (Cattell, 1966) was used in step c to determine the number of factors.

Both confirmatory factor analysis and exploratory factor analysis were conducted to investigate the validity of the ESAS in my study via SPSS AMOS 28. An initial confirmatory factor model was first built in AMOS (see Figure 5-15). The latent variables are those in big circles, the observed variables are in rectangles, and the small circles are labelled with e1 to e27 representing the potential error. The size of factor loading is an important consideration of communality in CFA, which is the path of coefficients leading from the latent variable (factors) to the observed variables. According to Hair (2009), factor loading of standardised estimates should be greater than .50, and ideally above .70. The square of a standardised factor loading means how much variation in an item is explained by the latent factor (Hair, 2009). For example, if a factor loading is .55, its square is .30, which means the factor is explaining 30% of the variation in the item, and the other 70% are error variances. The double-headed arrows connecting two exogenous variables imply correlation. The results of confirmatory factor analysis and exploratory factor analysis can be found in Section 6.2.1.2.

Figure 5- 15 *Initial ESAS confirmatory factor model*[31]



[31] LP = low English proficiency, LoSC = Lack of self-confidence, LoE = Lack of experience, UTC = unwillingness to communicate.

Group comparison

There are two types of group comparison in my study: a) FLSA changes before (pre-test) and after (post-test) the intervention with each group; b) differences among four groups in terms of FLSA levels after the intervention (post-test). When the data is normally distributed, parametric approaches are used to analyse the data. A paired-sample *t*-test is commonly used to compare two lots of data from the same group in pre- and post-test experiments (Woodrow, 2014). If the *p* value in the *t*-test results is lower than .05, it indicates significant differences in the data (Pallant, 2020). Analysis of variance (ANOVA) is often used to investigate the difference of means in more than two groups (Woodrow, 2014). One-way between-group ANOVA involves one independent variable with different levels and one dependent variable (Pallant, 2020). In my study, the independent variable corresponds to the interventions for four groups, and the dependent variable is the FLSA level. A *p* value smaller than .05 in ANOVA indicates significant differences among groups (Pallant, 2020). If so, *post-hoc* tests have to be conducted to see which specific groups are different from the other (Pallant, 2020).

When the data is not normally distributed, the Wilcoxon Signed Rank Test can be used to compare two paired groups. This is a non-parametric alternative to a paired-sample *t*-test (Wasserman, 2006). The Kruskal-Wallis Test is an alternative to ANOVA that can be used to investigate differences in more than two groups, when the data is not normally distributed (Wasserman, 2006).

5.5.2 Analysing Self-rating of FLSA Levels in Specific Tasks

Analysis of the self-rating of FLSA levels in specific tasks was similar to that for ESAS. The self-rating of FLSA levels in specific tasks (see Section 5.3.6) was obtained from an 8-item 10-point scale where students rated their anxiety levels in each 30-second period while watching a video recording of their performance in the speaking test. The average scores of the self-ratings were used for analysis. Due to the fact that these FLSA self-ratings examined the fluctuation of FLSA levels during a speaking test, there were no issues of reliability or validity as there were for ESAS. However, the normality check was conducted using the tests of skewness and kurtosis. When the data is normally distributed, paired-sample *t*-tests are applied for comparison of the FLSA levels in each group before and after the intervention. ANOVA and *post-hoc* tests are used to compare difference in FLSA levels in the post-test

among four groups. If the data is not normally disturbed, non-parametric approaches, such as the Wilcoxon Signed Rank Test and the Kruskal-Wallis Test, are used as alternatives.

5.6 Qualitative Data Analysis

5.6.1 Content Analysis

Content analysis was employed to analyse the qualitative data in my study. This technique is used to systematically compress text of words into fewer categories via coding (Stemler, 2000). Content analysis is not only a quantitative form of counting the frequency of words or presenting the relationship between variables; it also involves a qualitative way of encoding and categorising data with the deep meaning of texts produced by the participants (Baxter, 2009). Both quantitative and qualitative approaches to content analysis were applied to analyse the answers collected from the open-ended questionnaire (see Section 5.4) in my study. This is because combining quantitative and qualitative content analysis may overcome the weakness of each individual approach. For example, quantitative content analysis is more reliable, transparent and generalisable, but qualitative content analysis is more valid, flexible and useful for answering 'why' questions (Baxter, 2009).

Quantitative content analysis

Quantitative content analysis originates in communication science research and has been used in the field of educational technology since 1999 (Rourke & Anderson, 2004). According to Riff, Lacy, and Fico (2014, p. 20),

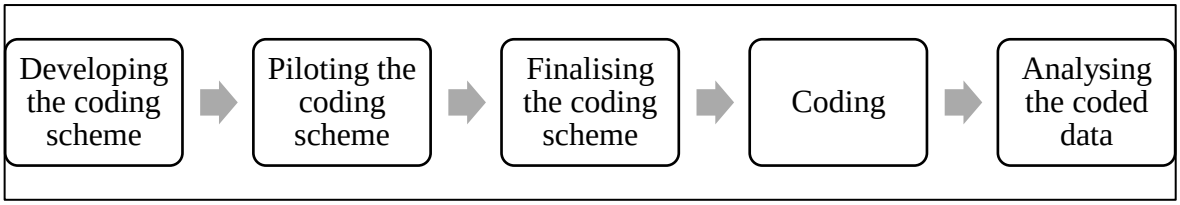
"Quantitative content analysis is the systematic and replicable examination of symbols of communication, which have been assigned numeric values according to valid measurement rules, and the analysis of relationships involving those values using statistical methods, in order to describe the communication, draw inferences about its meaning, or infer from the communication to its context, both of production and consumption".

Quantitative content analysis differs from other quantitative research methods in the application of a coding scheme and the way reliability is assessed (White & Marsh, 2006). It is sometimes treated as a research design rather than just a data analysis method. However, in my study, it was employed as a way to analyse the qualitative data collected from the

open-ended questionnaire in order to understand students' FLSA. The first step in utilising quantitative content analysis is to develop a coding scheme, which can be developed inductively from the data (White & Marsh, 2006). A coding form or coding schedule needs to be prepared to record details of the codes applied to the data during the coding process. The details of my coding scheme can be found along with the corresponding open-ended questions in Chapter 6. Following the suggestions by Neuendorf (2017), in order to ensure the feasibility of the cording scheme, I randomly selected some responses from the database and piloted to see if there were any problems with the initial coding forms.

Although it is better to involve multiple coders to make sure the coding scheme can be applied in a reliable way (Neuendorf, 2017), I was the only person who could access all the data collected for my study because of the ethical issues (see Section 5.9). I imported the written text of the responses into *NVivo* 12. Descriptive statistics, such as frequency counts, were used to summarise findings from the sample. Figure 5-16 shows the process of quantitative content analysis in my study.

Figure 5- 16 *Inductive approach of quantitative content analysis in my study*

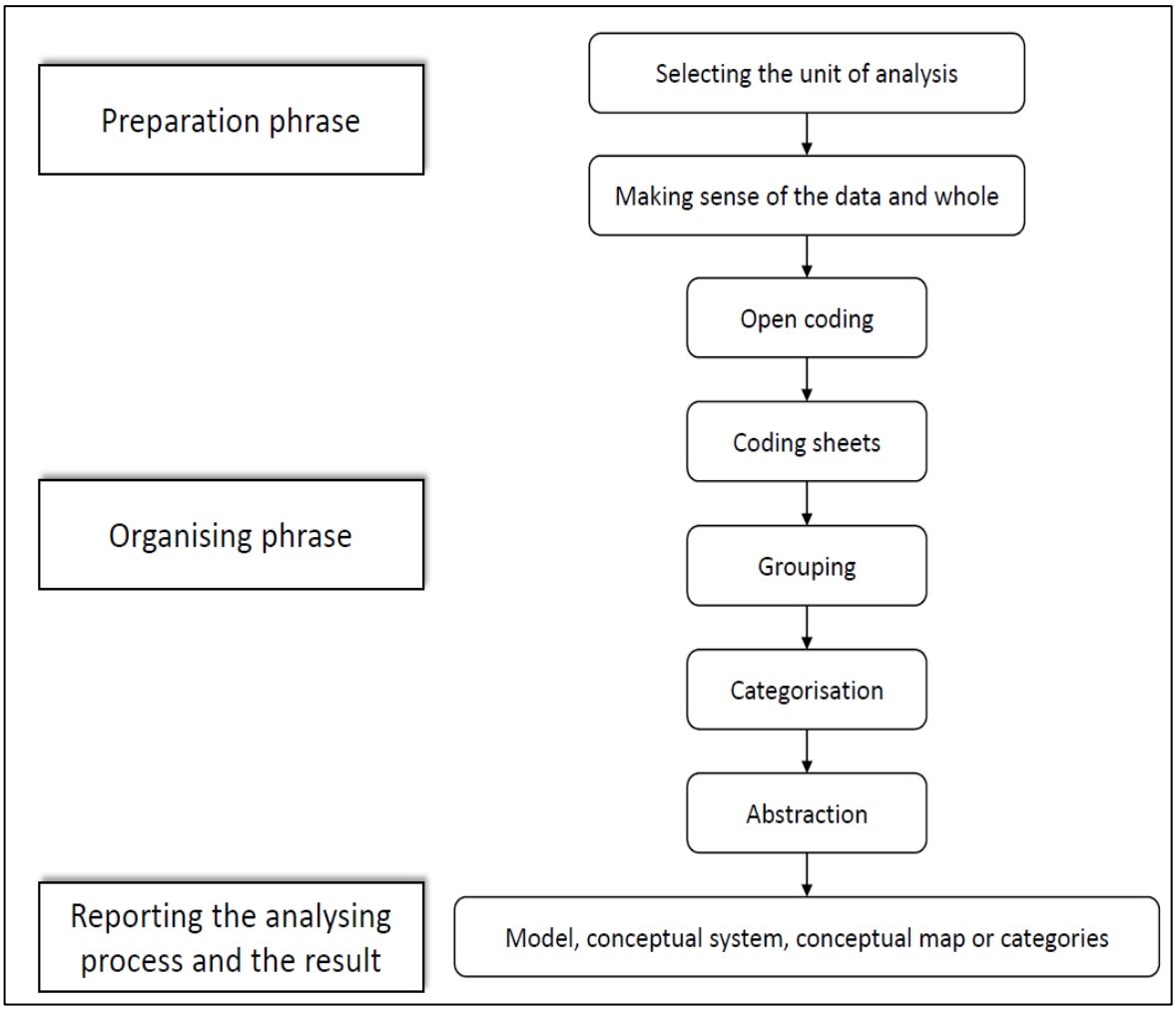


Qualitative content analysis

In a similar vein, an inductive method was applied with the qualitative content analysis in my study because it can be used to generate a general statement on unknown aspects of the topic. I followed the analytic process of qualitative content analysis (see Figure 5-17) proposed by Elo and Kyngäs (2008). First, I familiarised myself with the responses to the questionnaire and determined a unit of analysis for each question (see Chapter 6). Then I coded the text via an open coding approach, which segments data into meaningful expressions and then describes them as a single concept. A coding sheet was created after open coding, following which the notes could be grouped and used to generate categories. The purpose of creating categories is not simply to bring similar observations together but to provide a means of describing the phenomenon, enhance understanding and broaden

knowledge (Elo & Kyngäs, 2008). I, as the researcher, decided what content was to be put in the same category while interpreting the data. After the categorisation, I formulated a general description through generating categories, which was followed by sub-categories. The inductive process was iteratively applied for generating a model to help answer the research questions.

Figure 5- 17 *Inductive approaches to qualitative content analysis (adapted from Elo & Kyngäs, 2008)*



5.6.2 Reliability, Validity and Trustworthiness

Reliability in content analysis mainly refers to the *reproducibility*, which is defined as "the degree to which a process can be replicated by different analysts working under varying conditions, at different locations, or using different but functionally equivalent instruments" (Krippendorff, 2018, p. 215). A single researcher coding the text may produce highly subjective judgements, which may reduce the reliability of the content analysis (Oleinik *et*

al., 2014). However, as mentioned above, I was the only one who could access my data. I could not involve other coders because of ethical issues. Therefore, practising and piloting coding were applied to improve the reliability of the content analysis in my study.

The term of validity in content analysis refers to "the extent to which a measuring procedure represents the intended, and only the intended, concept" (Neuendorf, 2017, p. 122). Other than the internal validity and external validity previously explained in Section 5.5.1, *replicability* is important to consider in the content analysis procedure to ensure external validity of the analysis. Replicability refers to replication of the study with a different set of messages. This can be ensured by providing a full report of all content analysis procedures (Neuendorf, 2017). Validity in content analysis can be assessed in several ways, such as face validity, criterion validity, content validity, and construct validity (White & Marsh, 2006). Face validity refers to the extent to which a measurement covers essential aspects of the concept being measured, which can be corroborated by working backwards from the measurement to determine the concept being measured (Neuendorf, 2017). I therefore used iterative coding, reviewing the text, and coding several times. The other three types of validity did not seem to be applicable to the quantitative content analysis in my study.

The most widely used criterion for evaluating qualitative analysis is *trustworthiness*, developed by Lincoln and Guba (1985). The aim of trustworthiness is to confirm that the qualitative research findings are worth paying attention to. This is especially important for inductive content analysis, because categories are drawn from the data rather than theory (Elo *et al.*, 2014). According to Elo *et al.* (2014), the trustworthiness in qualitative content analysis should be considered in all three phases of preparation (data collection method, sampling strategy, and selecting the unit of analysis), organisation (categorisation and abstraction, interpretation, and representativeness) and reporting (reporting results and reporting the analysis process). For the preparation phase of my study, the data collection method and the sampling strategy were illustrated in Section 5.3 and Section 5.4. Details of the analytic process, including the selection of the unit of analysis, categorisation, interpretation, and representativeness, are illustrated in the analysis of each open-ended question in Chapter 6. Direct citations of the original data were used to enhance the trustworthiness of the analysis (Elo & Kyngäs, 2008).

5.7 Integration of Data Collection and Analysis

The integration of quantitative and qualitative data can enhance the value of mixed methods research. This can be implemented in the stages of research design and research methods as well as interpreting and reporting the results (Fetters, Curry, & Creswell, 2013). As mentioned in Section 5.2, my research involved a mixed methods experimental design. The integration of research methods was achieved by *merging* the data and "bringing the two databases together for analysis and for comparison" (Fetters, Curry, & Creswell, 2013, p. 2140). Specifically, questions (questions 1, 2 and 3 in the open-ended questionnaire) were used to collect qualitative data on FLSA levels, data that was also collected through the ESAS and self-rating of FLSA in the speaking tests in a quantitative way. The two data sets were compared to see whether the difference in FLSA levels found in a quantitative manner were reflected in qualitative analysis of the textual data. This would provide a comprehensive understanding of the impacts of HiVR on FLSA. Table 5-6 shows how data collection and data analysis were integrated in my mixed methods study. In Chapter 6, a joint display of results is presented to illuminate the integration in the stage of interpreting and reporting results. Joint displays allow researchers to draw conclusions beyond those which can be gained from separate quantitative and qualitative results by presenting the data visually (Fetters, Curry, & Creswell, 2013).

Table 5- 6 *The integration of data collection and analysis in my study*

Data collected	**Analysis methods**	**Goal**	**Integration**	**Corresponding RQs**
ESAS total scores	ANOVA, *t*-tests, *post-hoc* test	To investigate the change of ESAS scores before and after the intervention	To compare with students' descriptions of general FLSA before and after the intervention	RQ1 and RQ2
Average FLSA levels in specific tasks	ANOVA, *t*-tests, *post-hoc* test	To investigate the change of self-ratings of FLSA levels in two speaking tests	To compare with reflections of FLSA levels in the speaking test	RQ1 and RQ2
Written responses in the open-ended questionnaire	Quantitative content analysis	To investigate the descriptions of difference in FLSA from the qualitative data	To compare with the quantitative results of RQ1 and RQ2	RQ3
	Qualitative content analysis	To understand student's FLSA before and after the intervention	To support, explain or refute the quantitative results of RQ1 and RQ2	RQ3
			To further understand FLSA among Chinese EFL learners	
		To understand students' perceptions of HiVR and FLSA as well as their experiences of using HiVR for foreign language learning	N/A	RQ4 and RQ5

5.8 Ethical Considerations

Before the fieldwork was conducted, ethics approval was obtained from the College of Social Sciences, University of Glasgow. The idea of informed consent is crucial in social science research in order to provide information for the participants that allows them to decide whether to volunteer for the research (Cohen, Manion, & Morrison, 2017). The aims of my study and my contact details were provided in the paper versions of the Plain Language Statements (see Appendix 7 and 8) and the Informed Consent Forms (see Appendix 9 and 10) while recruiting participants. All the documents were reviewed and approved by the Research Ethics Committee at the University of Glasgow. Permission to access the university in China was gained from the head of the Foreign Language College, as illustrated in Section 5.3. All participants signed the consent form and returned it to me to confirm their willingness to participate in the study.

All data collected from the participants was stored on secure personal computers and on a password-protected hard disk. To ensure the confidentiality and anonymity of participants, pseudonyms were used in storing the data as well as in reporting the results of data analysis. For example, SVR1 refers to Student 1 from the SVR group. Although the video recordings could identify the participants, the recordings were not involved in the data analysis but just shown to the corresponding participant for self-rating of FLSA levels. In addition, only my supervisors and I have access to the raw data.

5.9 Summary

This chapter presented an overview of the research design and the data collection procedure at the beginning (see Section 5.1). The mixed methods experimental design of my research was explained in Section 5.2. The setting (a Chinese university) and participants (140 students in total) as well as the sampling approach and group division method were demonstrated in Section 5.3. The instruments (ESAS, self-rating of FLSA levels in specific tasks, and open-ended questionnaires), the interventions for different groups (four groups in total) were introduced in Sections 5.3 and 5.4. The data analysis methods, along with the integration of the quantitative and qualitative data sets, were presented in Sections 5.5, 5.6 and 5.7. In addition, the issues of reliability, validity and trustworthiness were discussed in Sections 5.5 and 5.6 and the research ethics were covered in Section 5.8. The results of the data analysis will be presented in the following chapter.

Chapter 6. RESULTS

This chapter presents the results of the quantitative and qualitative data analysis and the integration of the analysis in both databases in order to generate a comprehensive understanding of the impacts that High-immersion Virtual Reality (HiVR) may have on Foreign Language Speaking Anxiety (FLSA) levels among Chinese EFL (English as a Foreign Language) learners. As stated in Chapter 5, the instruments employed for quantitative data collection included the English Speaking Anxiety Scale (ESAS), and the self-ratings of FLSA levels in the role-play English speaking tests. Qualitative data included texts collected through open-ended questionnaires after the experiment. The Research Questions (RQs) examined in this study are shown below. Figure 6-1 shows the data analysis procedure and corresponding RQs that drives the structure of this chapter. Specifically, RQ 1 and RQ 2 were first examined quantitatively. Students' perceptions of their FLSA changes were analysed in order to compare with and/or to explain the quantitative results of RQ 1 and RQ 2 in order to answer RQ 3. Qualitative analysis was further conducted in order to understand students' perceptions of HiVR and FLSA for RQ 4 as well as experiences of using HiVR for oral English learning for RQ 5.

RQ 1: To what extent (if any) do the FLSA levels of students who used HiVR change after the intervention?

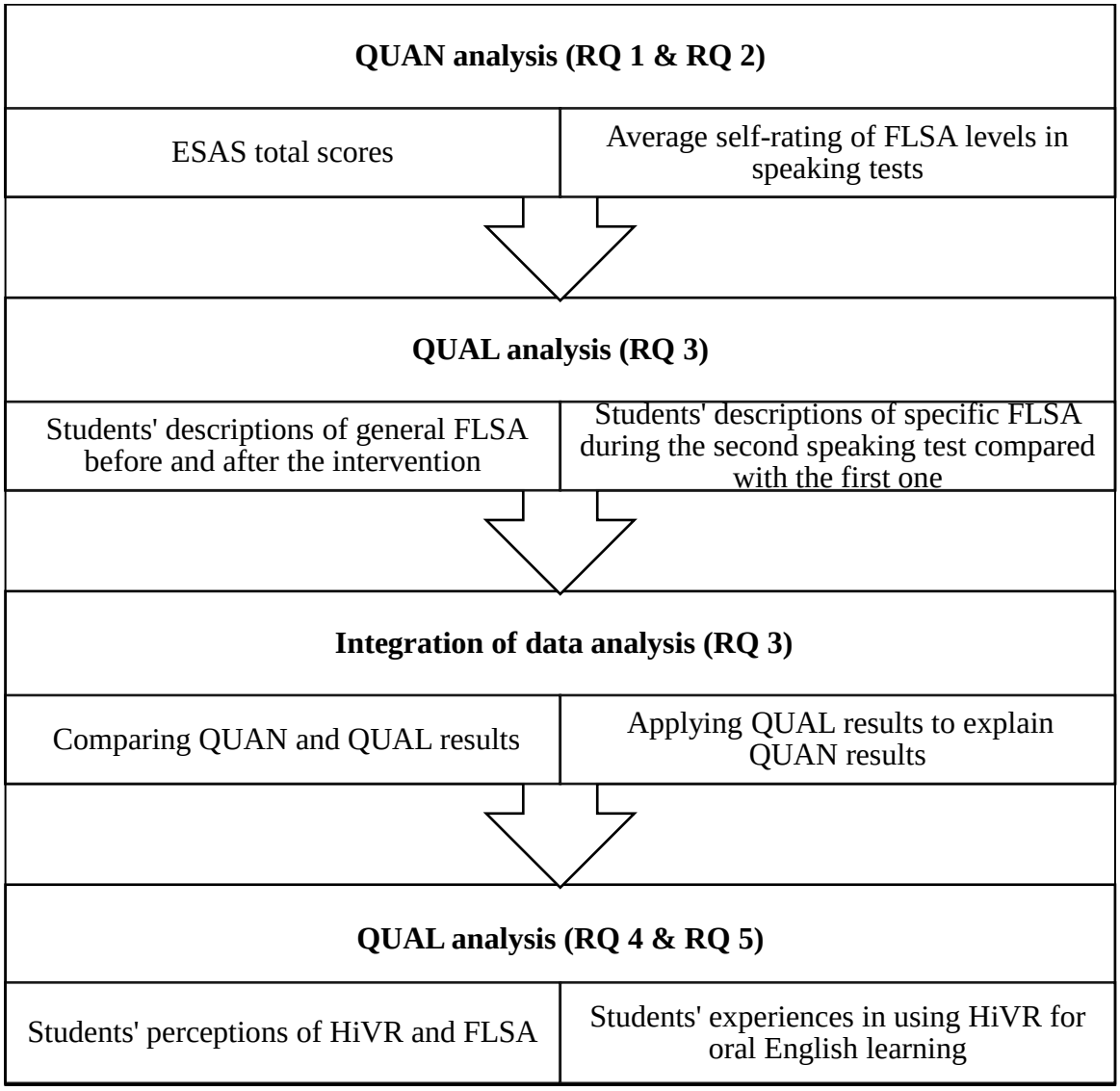
RQ 2: To what extent (if any) do two learning environments (HiVR and classroom) combined with two learning approaches (situated learning and teacher-centred learning) have different degrees of influence on students' FLSA levels?

RQ 3: How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?

RQ 4: What are students' perceptions of the impact of HiVR on FLSA?

RQ 5: What are students' experiences of using HiVR in oral English learning?

Figure 6- 1 *Data analysis procedure of my study*



6.1 Group Division

A total of 140 university students participated in my study, and they were divided into four groups (35 students each). These groups experienced different learning environments (HiVR and classroom) and learning approaches (situated learning and teacher-centred learning). Based on the 2*2 factorial design (see Section 5.3.1), the groups were named according to the initial letters of the different learning approaches and learning environments as follows:

SVR group: situated learning and HiVR

TVR group: teacher-centred learning and HiVR

SC group: situated learning and classroom

TC group: teacher-centred learning and classroom

The method for dividing students into different groups was briefly illustrated in Section 5.3.2. Specifically, students were first divided into three strata according to their total scores of English speaking anxiety scale (ESAS) before the intervention. The strata were low anxiety (48 students, score range 52–83), moderate anxiety (48 students, score range 84–97) and high anxiety (44 students, score range 98–123). From each stratum (high, middle, and low anxiety) 25% of students were then randomly assigned to different groups using SPSS. For instance, 25% of students in the high anxiety stratum were randomly allocated to the SVR group, another 25% in the high anxiety stratum were randomly allocated to the TVR group, and so forth.

Average self-rating of FLSA levels in specific tasks was considered afterwards to slightly switch the allocation of participants among groups in order to retain the equivalence of each group in FLSA levels before the experiment. The other criteria (gender, age, years of learning English and VR experience) were also taken into account during group division. As shown in Table 6-1, there were 33 females and 2 males in each group. Nearly half the students (16–18 students) in each group were 18–19 years old, and nearly another half (17–19 students) in each group were 20–21 years old. The average number of years of learning English for each group was between 10.5 and 11 years. In each group, five students had used HiVR before, and the rest 30 students had no experience of HiVR. The mean of ESAS score in each group ranged from 89.34 to 89.74, and the average self-rating of FLSA levels in speaking tests was from 5.82 to 5.85.

Table 6- 1 *Group information at the start of the experiment (N = 140)*

	SVR (N = 35)	**TVR (N = 35)**	**SC (N = 35)**	**TC (N = 35)**
Gender (female/male)	33/2	33/2	33/2	33/2
Age (18–19/20–21)	17/18	18/17	16/19	16/19
English learning years (M[32]/SD[33])	10.37/1.80	10.66/1.41	10.31/1.81	11.06/1.62
VR experience (has/has not)	5/30	5/30	5/30	5/30
ESAS total scores (M/SD)	89.34/14.51	89.34/15.39	89.46/15.32	89.74/14.15
Self-ratings of FLSA in the speaking test(M/SD)	5.82/1.77	5.83/1.44	5.83/1.61	5.85/1.54

6.2 Quantitative Data Analysis (RQ 1 and RQ 2)

Two types of FLSA level (ESAS total scores and average self-rating of FLSA in speaking tests) were collected for RQ 1 and RQ 2 (see below). In my study, the alpha level to report the statistical significance was set to .05, and the *p* values reported were all two-tailed.

> **RQ 1:** To what extent (if any) do the FLSA levels of students who used HiVR change after the intervention?
> **RQ 2:** To what extent (if any) do two learning environments (HiVR and classroom) combined with two learning approaches (situated learning and teacher-centred learning) have different degrees of influence on students' FLSA levels?

6.2.1 English Speaking Anxiety Scale

Descriptive normality and reliability analysis

The English Speaking Anxiety Scale (ESAS, see Appendix 1) for Chinese EFL learners contains 27 items with a 5-point Likert scale. Descriptive statistics of ESAS scores for the whole group and for each group for both pre-test and post-test can be found in Table 6-2.

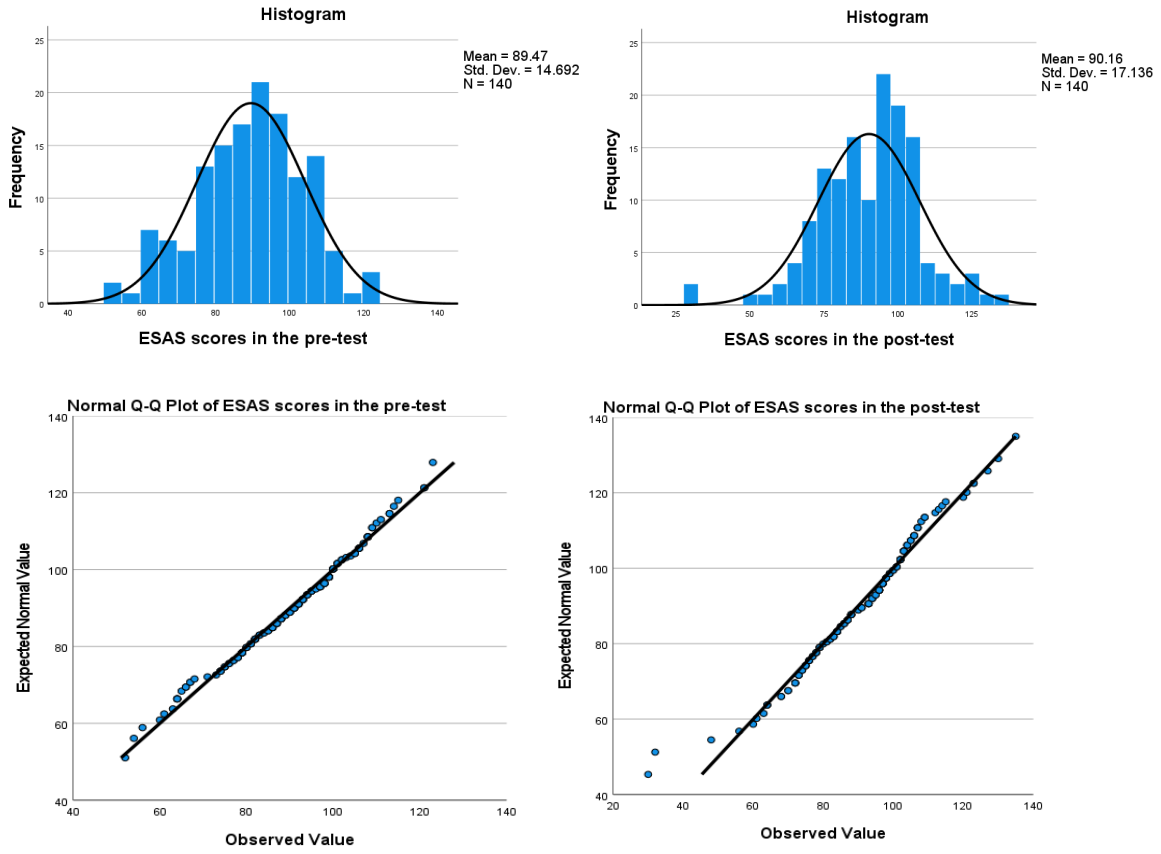
[32] Mean
[33] Standard Deviation

Skewness and kurtosis were tested for the normality of ESAS scores in both pre- and post-test. It seems ESAS scores in both the pre-test (skewness = -.251; kurtosis = -.234) and the post-test (skewness = -.439; kurtosis = 1.271) were normally distributed. This is because skewness between -2 to +2 and kurtosis between -7 to +7 are claimed to be acceptable values for normal distribution (Hair, 2009). The histograms and Q-Q plots (see Figure 6-2) were created to further determine the distribution of the dataset. The pre- and post-test ESAS scores seemed to fit the normal curve in the histograms and the data was for the most part close to the straight line in the Q-Q plot, which suggests a normal distribution. Therefore, a parametric approach was used to analyse ESAS scores. The Cronbach's alpha level was .933 and .947 respectively for the ESAS scores in pre-test and post-test, which indicated that 27 items in the ESAS are highly related with each other. Thus, the construct of the ESAS is reliable.

Table 6- 2 *Descriptive statistics of ESAS scores in pre- and post-test*

	Pre-test		**Post-test**	
Groups	**M**	**SD**	**M**	**SD**
Entire groups (N = 140)	89.47	14.69	90.16	17.14
SVR (N = 35)	89.34	14.51	89.91	15.79
TVR (N = 35)	89.34	15.39	90.57	18.06
SC (N = 35)	89.46	15.32	89.20	19.89
TC (N = 35)	89.74	14.15	90.94	15.09

Figure 6- 2 *Histograms and Q-Q plots of ESAS scores in the pre- and post-test (N = 140)*



Factor analysis

In order to assess the construct validity of ESAS, a confirmatory factor analysis was first performed on the four-factor structure (see below) with the remaining 27 items. Based on previous literature (see Section 2.3), the model hypothesised four categories of FLSA:

1. Low Proficiency (LP),
2. Lack of Self-confidence (LoSC),
3. Lack of Experience (LoE),
4. Unwillingness to Communicate (UTC).

The standardised estimates results (see Figure 6-3) showed that the factor loading size of item 3, item 6 and item 27 were lower than .50, which means these items were unlikely to be communal with their corresponding factor. According to the accepted values for each parameter displayed in Table 6-3, the model fit of the original model was poor (CMIN/DF = 1.894, GFI = .742, CFI = .821, RMSEA = .08, SRMR = .077). This indicates that the analysis did not fully support a four-category model as a definitive explanation of the ESAS

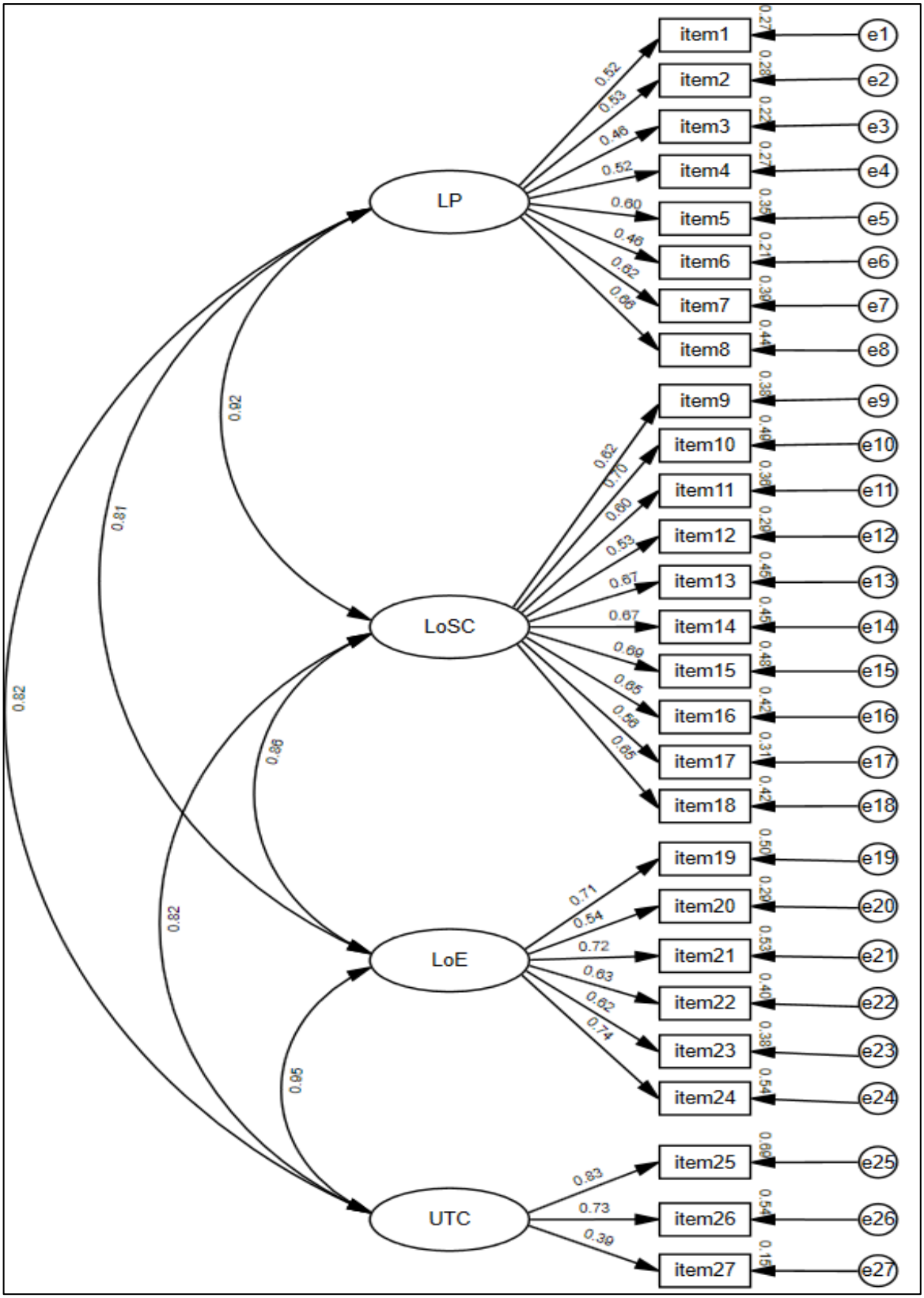
data in my study. Moreover, the correlation between each pair of two factors was statistically significantly large (larger than .8), which may be interpreted to mean the four categories are highly related with each other. Thus, an exploratory factor analysis was conducted to investigate other possible constructures in order to better identify the dimensions explaining the ESAS under consideration.

Table 6- 3 *Model fit accepted values of confirmatory factor analysis adapted from Uedufy*

Parameters	Accepted value	Reference
Chi-square divided by Degree of Freedom (CMIN/DF)	≤ 3	(Tabachnick, Fidell, & Ullman, 2007)
Goodness of Fit Index (GFI)	≥ .90	(Hu & Bentler, 1999)
Comparative Fit Index (CFI)	≥ .90	(West, Meserve, & Stanovich, 2012)
Root Mean Square Error of Approximation (RMSEA)	≤ .05	(MacCallum, Browne, & Sugawara, 1996)
Standardized Root Mean Squared Residual (SRMR)	≤ .05	(Diamantopoulos, 1994)

The 27-item ESAS was first assessed for its suitability for exploratory factor analysis. The result of Barlett's test of sphericity was highly significant ($p < .001$), and the Kaiser-Meyer-Olkin value of .904 exceeded the recommended value of .6 (Kaiser, 1970), which meant the dataset was suitable for factor analysis. Through principal components analysis, it was revealed that five eigenvalues exceeded 1, explaining 37.50%, 7.82%, 4.78%, 4.24% and 4.17% of the variance, respectively. The cumulative explanation was 58.51%, so the correlation between each item in the ESAS was examined. It is recommended by Hair (2009) that if the cumulative explanation is lower than 60%, it means the construct is not valid, and revisiting the data is needed. The correlation between item 6 "It frightens me when I do not understand what others are saying in English" and other items in the domain of low English proficiency, as well as that between item 27 "Generally, I feel nervous when I have to speak to others" and the other items in the domain of unwillingness to communicate, was lower than .3. Consequently, item 6 and item 27 were removed from the ESAS for the analysis of the next step.

Figure 6- 3 *Standardised parameter estimates of confirmatory factor analysis of ESAS with the initial model*



After removing item 6 and item 27, principal components analysis revealed five eigenvalues exceeding 1 and explained 60.53% of the variances in total. A break was observed in the scree plot after the third component. According to the scree test, a three-factor solution was retained, which explained a total of 51.86% of the variance. As shown in Appendix 11, a total of 13 items were included in Factor 1 (F1), and they are mostly related to concerns regarding low language proficiency. A total of nine items loaded onto Factor 2 (F2) were mainly concerned with low self-confidence. There were three items in Factor 3 (F3) which focused on pronunciation difficulty. Therefore, according to the three-factor solution, the factors that had an impact on students' FLSA may be summarised as low proficiency, low self-confidence, and pronunciation difficulties.

Nevertheless, the confirmatory factor analysis results of the above third-factor model did not produce good model fit (CMIN/DF = 1.808, GFI = .788, CFI = .855, RMSEA = .076, SRMR = .072). Items with factor loading lower than .50 in the three-factor confirmatory factor analysis model were sequentially removed from the ESAS for further analysis. After applying several iterations of exploratory factor analysis and confirmatory factor analysis to examine the structure of ESAS, no acceptable model was found. This indicates that the concepts of ESAS were highly related to each other and supported the high reliability of ESAS as an instrument to measure FLSA levels in my case. Therefore, the ESAS was used as a single scale rather than breaking it down in several factors for further analysis in this thesis. Moreover, further investigation into the construct of ESAS is informed by qualitative data analysis in my study (see Section 6.3.1.1).

Comparison of total ESAS scores

A paired-sample *t*-test was employed to compare the total scores of the ESAS before and after the intervention. The statistical results are presented in Table 6-4. For the participants as a whole (N = 140), the ESAS scores showed no statistically significant changes ($p = .488$) and were very close to each other. Within each group, the results of paired-sample *t*-tests also revealed no significant difference in ESAS scores (SVR: $p = .755$, TVR: $p = .510$, SC: $p = .913$, TC: $p = .534$) between pre-test and post-test. It seems that neither HiVR nor situated learning theory influenced the FLSA levels measured by ESAS.

Table 6- 4 *Paired-sample t-test results comparing ESAS scores among whole participants and in each group in the pre-test and post-test*

	Pre-test		**Post-test**				
	M	**SD**	**M**	**SD**	***t* (*df*)**	***P* value**	**Cohen's *d***
Entire group	89.47	14.69	90.16	17.136	*t* (*139*) = -.695	.488	-.059
SVR	89.34	14.51	89.91	15.79	*t* (*34*) = -.315	.755	-.053
TVR	89.34	15.39	90.57	18.06	*t* (*34*) = -.666	.510	-.113
SC	89.46	15.32	89.20	19.89	*t* (*34*) = .110	.913	.019
TC	89.74	14.15	90.94	15.09	*t* (*34*) = -.628	.534	-.106

A one-way analysis of variance (ANOVA) test was employed to examine the difference in post-test ESAS scores in the four groups. The results of the one-way ANOVA test (see Table 6-5) showed no statistically significant differences (F (3,136) = .069, p = .977) among the four groups' post-test ESAS scores. This might indicate that different learning environments (HiVR or classroom) and learning approaches (situated learning or teacher-centred learning) made no difference to student's FLSA levels measured via ESAS.

Table 6- 5 *ANOVA comparing post-test ESAS scores among four groups*

Between groups	***df***	***F***	***p***	**Partial eta squared (*η2*)**
ESAS scores	3	.069	.977	.002

To sum up, ESAS seemed to offer a highly reliable measurement of FLSA levels in my study, but its construct validity needs further investigation. The statistical results of the ESAS total scores answered the RQs as follows:

> The answer to RQ 1: students' FLSA levels measured by ESAS did not show statistically significant changes after using HiVR;
>
> The answer to RQ 2: no statistically significant difference was found among the four groups in terms of FLSA measured by ESAS after the intervention.

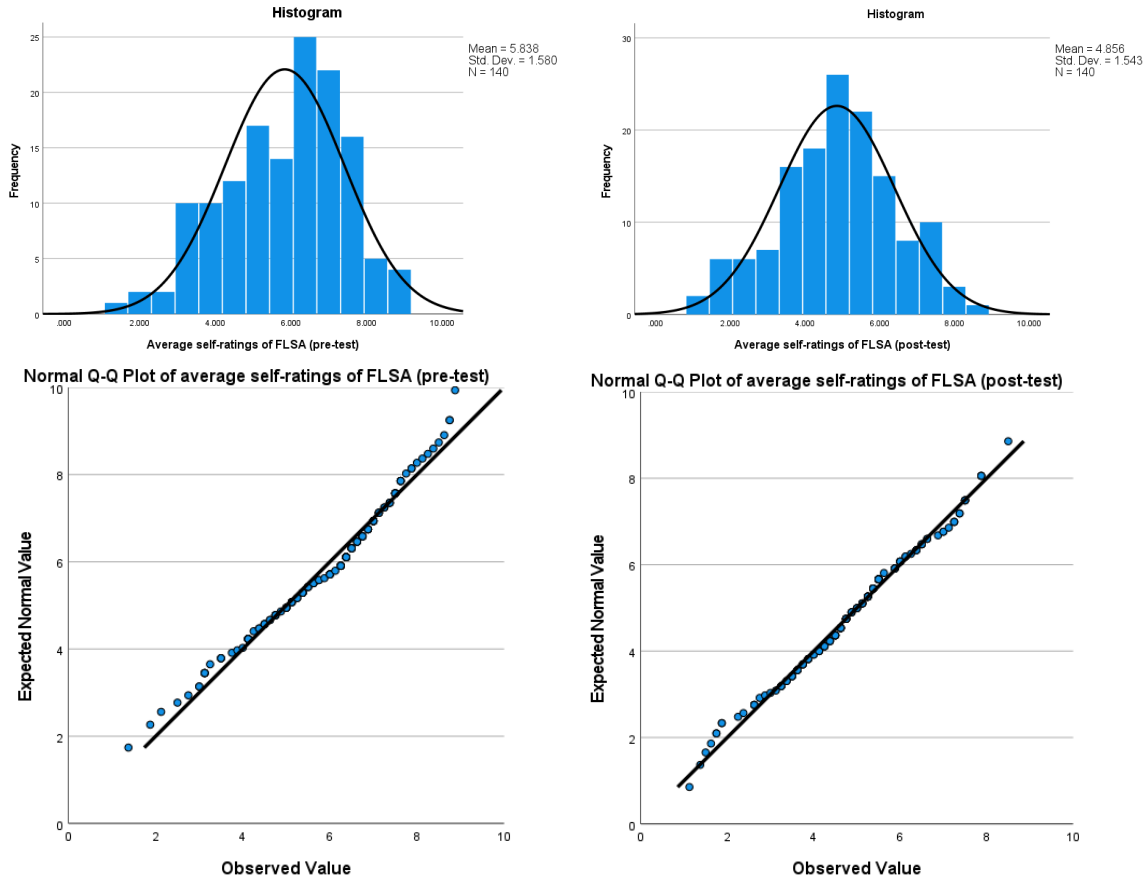
6.2.2 Self-rating of FLSA in Specific Tasks

The self-ratings of FLSA levels in specific tasks were obtained from an 8-item 10-point scale with which students rated their anxiety levels every 30 seconds while watching a four-minute video recording of their performance in the speaking tests. Table 6-6 displays the descriptive statistics of the average self-rating of FLSA levels in pre-test and post-test for the entire participants and for each group. All the self-rating scores lie in an approximately normal distribution according to the results of skewness and kurtosis (see Table 6-6) as well as histograms and Q-Q plots (see Figure 6-4). Therefore, parametric approaches were employed for the analysis of self-ratings of FLSA levels in specific tasks.

Table 6- 6 *Descriptive statistics of self-ratings of FLSA in first and second speaking tests*

	Pre-test				**Post-test**			
	M	**SD**	**Skewness**	**Kurtosis**	**M**	**SD**	**Skewness**	**Kurtosis**
Entire group (N =140)	5.84	1.58	-.457	-.257	4.86	1.54	-.111	-.170
SVR (N = 35)	5.85	1.45	.046	-.024	4.99	1.13	.089	-.446
TVR (N = 35)	5.82	1.77	-.591	-.442	5.21	1.61	-.287	.331
SC (N = 35)	5.83	1.61	-.455	-.789	4.28	1.76	.287	-.251
TC (N = 35)	5.86	1.54	-.705	.695	4.95	1.50	-.132	-.439

Figure 6- 4 *Histograms and Q-Q plots of self-ratings of FLSA in first and second speaking tests (N = 140)*



Based on the results of paired-sample *t*-tests (Table 6-7), the average self-ratings of FLSA levels decreased significantly (t (*139*) = 6.828, $p < .001$, Cohen's d = *.577*) from pre-test (M = 5.84, SD = 1.58) to post-test (M = 4.86, SD = 1.54) for all participants.

Table 6- 7 *Paired-sample t-test results comparing self-ratings of FLSA in first and second speaking tests (N = 140)*

	Pre-test		**Post-test**				
	M	**SD**	**M**	**SD**	***t* (*df*)**	***P* value**	**Cohen's *d***
Average scores	5.84	1.58	4.86	1.54	t (*139*) = 6.828	< .001	.577

A comparison was then conducted on the average self-ratings of FLSA in speaking tests within each group. According to the paired-sample *t*-test results shown in Table 6-8, a significant decrease in average self-rating of FLSA levels was found in the SVR group (t (*34*) = 3.375, p = .002, Cohen's d = .571) and the TVR group (t (*34*) = 2.693, p = .005,

Cohen's d = .455) after the intervention. The self-ratings of FLSA in the SC group (t (*34*) = 4.496, p < .001, Cohen's d = .760) and the TC group (t (*34*) = 3.048, p = .004, Cohen's d = .515) also reduced significantly after the intervention.

Table 6- 8 *Paired-sample t-test results comparing self-ratings of FLSA between first and second speaking tests in each group*

	Pre-test		**Post-test**				
	M	**SD**	**M**	**SD**	***t* (*df*)**	***P* value**	**Cohen's *d***
SVR (N = 35)	5.85	1.45	4.99	1.13	*t* (*34*) = 3.375	.002	.571
TVR (N = 35)	5.82	1.77	5.21	1.61	*t* (*34*) = 2.693	.005	.455
SC (N = 35)	5.83	1.61	4.28	1.76	*t* (*34*) = 4.496	< .001	.760
TC (N = 35)	5.86	1.54	4.95	1.50	*t* (*34*) = 3.048	.004	.515

Figure 6-5 shows the change in self-rating of FLSA levels in each group before and after the intervention. It is notable that the SC group reduced the most (1.55) after the intervention, followed by the TC group with 0.91 in the self-rating of FLSA in specific tasks. The decrease in the SVR group reached 0.86, and the TVR group seemed to reduce the least (0.61) of the four groups.

Figure 6- 5 *Difference in average self-rating of FLSA in speaking tests in each group*

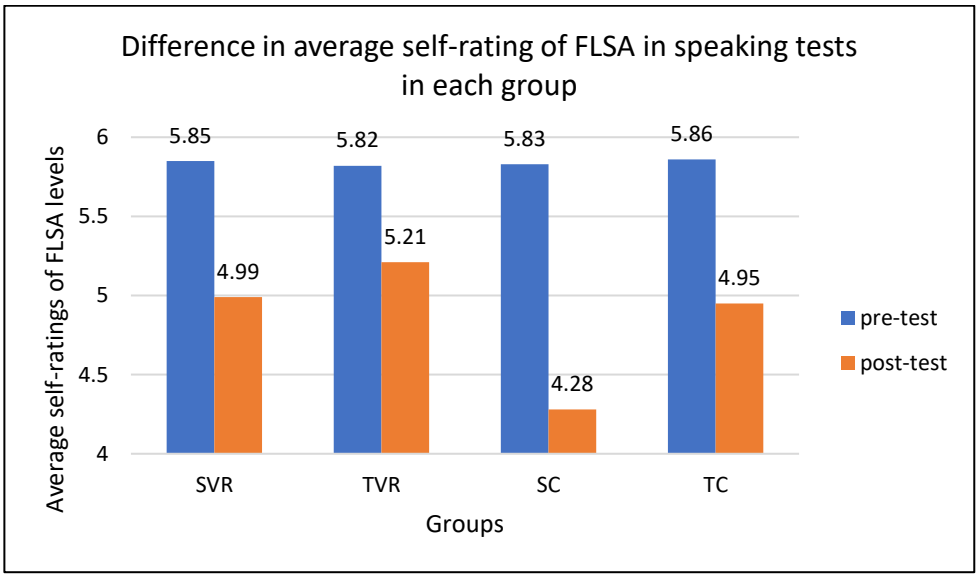


Figure 6-6 shows that the SC group seemed to remain the lowest self-rating of FLSA levels during the whole four-minute English speaking test. However, the one-way between-group ANOVA results (see Table 6-9) showed that no statistically significant difference in FLSA levels between each group were found after the intervention (F (3, 136) = 2.471, p = .065). This might be interpreted as that the difference in learning environment (HiVR or classroom) and learning approach (situated learning or teacher-centred learning) did not result in any statistically significant differences in average self-ratings of FLSA among the four groups in the second speaking test, although the decrease in level of specific FLSA in each group varied.

Figure 6- 6 *Self-rating of FLSA levels in each time interval of the second speaking test in each group*

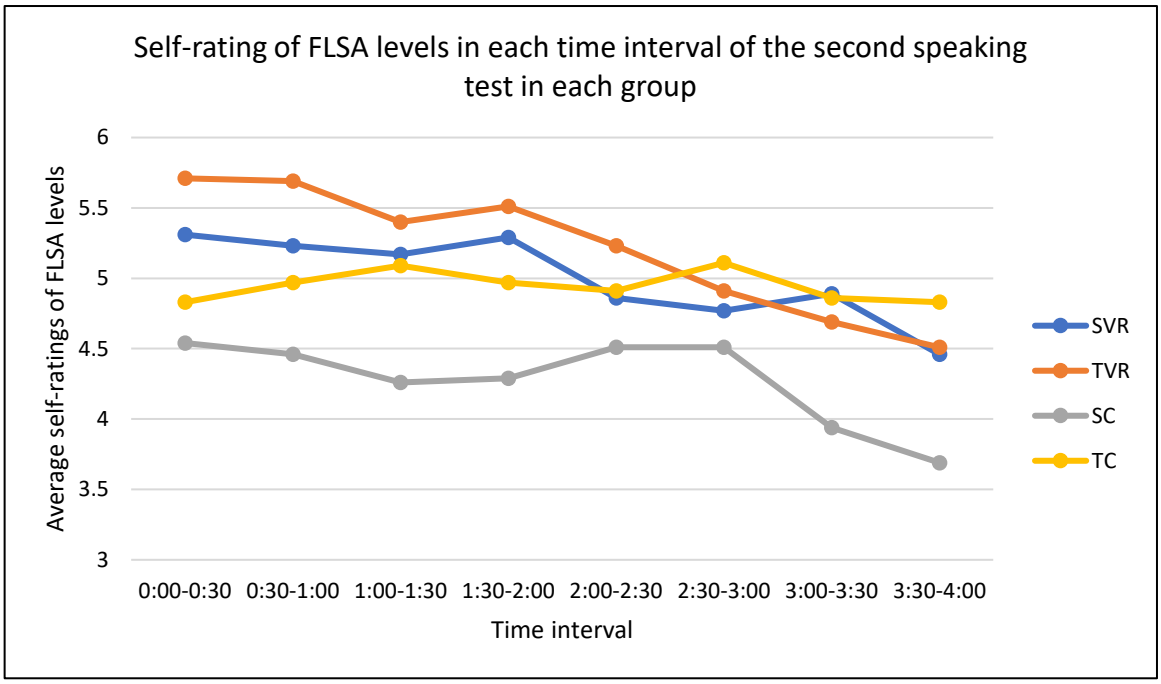


Table 6- 9 *ANOVA comparing average self-ratings of FLSA in specific tasks in the four groups*

Between groups	***df***	**F**	***p***	***Partial eta squared (η2)***
Self-rating of FLSA in specific tasks	3	2.471	.065	.052

Above all, in terms of self-ratings of FLSA levels in the speaking tests, it seemed that students' FLSA levels decreased significantly after the intervention in each learning

environment (HiVR or classroom) and learning approach (situated learning or teacher-centred learning), but none of these conditions made a difference to the self-rating of FLSA levels in the second speaking tests. The RQs could be answered as follows:

> The answer to RQ 1: students' self-ratings of FLSA levels in the speaking tests witnessed a statistically significant decrease after using HiVR;
> The answer to RQ 2: no statistically significant differences were found in the four groups in terms of average self-rating of FLSA levels in the second speaking tests.

6.2.3 Summary of Quantitative Results

Table 6-10 summarises the quantitative results for my study with regard to RQ 1 and RQ 2. Combining the results of ESAS scores and self-ratings of FLSA levels in the speaking tests, contradictory statistical results were revealed in relation to RQ 1. For the HiVR groups (SVR and TVR), no statistically significant changes in ESAS scores were found before and after the intervention, but students' average self-ratings of FLSA levels in the speaking tests reduced significantly in the second speaking test compared with the first. Moreover, self-rating of FLSA in the speaking tests reduced the most in the SC group, and the least in the TVR group. A possible reason for difference in results of ESAS and self-ratings of FLSA levels in the speaking tests is that they are measuring different types of FLSA. Further explanation will be sought by analysing qualitative data (see Section 6.3).

With regard to RQ 2, the results were consistent. None of the four groups showed a statistically significant difference in FLSA levels after the intervention. In other words, after a nine-session intervention, neither the learning environments (HiVR and classroom) nor the learning approaches (situated learning and teacher-centred learning) resulted in statistically significant differences in students' FLSA levels.

Table 6- 10 *Quantitative results summary*

RQs	Instrument	Data type	Groups	Results
RQ1	ESAS	Total ESAS scores before and after the intervention	SVR TVR	No statistically significant changes
	Self-rating of FLSA in specific tasks	Average FLSA levels in the first and second speaking test		Statistically significant decrease in the second speaking test (SC group decreased the most, TVR group decreased the least)
RQ2	ESAS	Total ESAS scores after the intervention	SVR TVR SC TC	No statistically significant differences
	Self-rating of FLSA in specific tasks	Average FLSA levels in the second speaking test		No statistically significant differences

6.3 Qualitative Data Analysis

Qualitative data was collected from an open-ended questionnaire with five questions (see Section 5.4). The first three questions were analysed to compare with or explain quantitative results in order to answer RQ 3. The fourth question aimed to investigate students' perceptions of HiVR and FLSA for RQ 4. The last question obtained data about students' experiences of using HiVR for oral English learning to answer RQ 5. The Research Questions (RQs) required qualitative data are as follows:

> **RQ 3:** How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?
> **RQ 4:** What are students' perceptions of the impact of HiVR on FLSA?
> **RQ 5:** What are students' experiences of using HiVR in oral English learning?

All responses to the open-ended questionnaire were analysed via *NVivo* 12. The analysis was based on the Chinese version of responses but translated to English for the understanding in this thesis. I, as the researcher, am a native speaker of Chinese and a proficient user of English, which ensures the quality of Chinese-English translation. The analysis of each

question was conducted through the content analysis approach including quantitative content analysis (see Section 5.6.1.1) and inductive qualitative content analysis (see Section 5.6.1.2).

6.3.1 General FLSA Change (RQ 3)

The following sections (Section 6.3.1, 6.3.2, and 6.3.3) of qualitative analysis answer RQ 3 (see below). In data presentation, the participants were coded using the name of their group and a randomly assigned number between 1 and 35. For example, SVR1 means the first student in the SVR group.

> **RQ 3:** How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?

Question 1 (Q1): Please describe your anxiety about speaking English with others BEFORE the learning sessions.

Q1 aims to obtain students' perceptions of their general FLSA before the intervention. The description of students' feelings was coded according to the scheme (see Table 6-11) as anxious or non-anxious. Some students expressed their anxiety explicitly, such as *I am anxious/nervous about…*, or *I am not confident about…*, but some mentioned it implicitly, saying *I am worried about …, I am fear of …,* and *I cannot ….* Non-anxious students expressed it positively (e.g. *relaxing, enjoyable, interesting*) or negatively (e.g. *not anxious, not nervous, not fear of*).

Table 6- 11 *Coding scheme for Q1*

Categories	**Responses examples**
Anxious	*Anxious, nervous, worried about, fear of, not confident, embarrassed, panic*
Non-anxious	*Not anxious, not nervous, not fear of, enjoyable, interesting, relaxing*

Based on quantitative content analysis of the responses to the open-ended questionnaire, it was found that 93.57% of the participants (N = 131) were anxious about speaking English

at the start of the experiment. Two or three students in each group were not anxious about speaking English at all. Generally, anxious students stated that they were nervous before speaking English (e.g. SVR10, SVR31, TVR29, TVR33, SC14, SC24, TC25, TC18), felt anxious while speaking (e.g. SVR4, SVR18, TVR1, TVR4, SC34, SC33, TC1, TC9), or even could not speak English at all (e.g. SVR31, SVR5, TVR9, TC7).

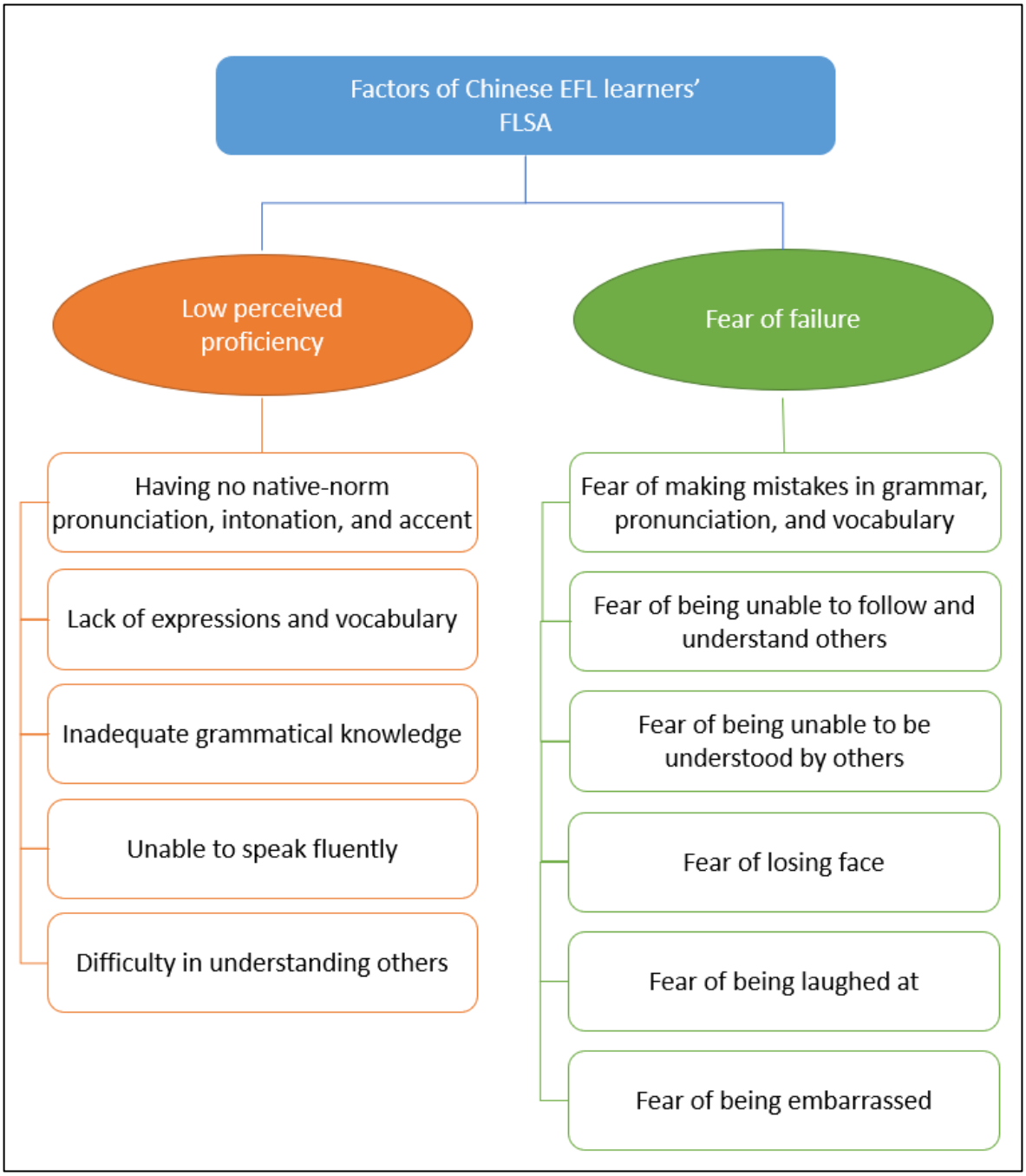
However, some students (N = 35) thought their anxiety levels might vary in different situations, such as facing different interlocutors, talking about different topics, and being involved in a different environment. This is consistent with my distinction between general FLSA and specific FLSA (see Section 2.1.2), which demonstrates that foreign language learners may be anxious in some foreign language speaking situations but non-anxious in others. In my study, students (SVR21, SVR11, TVR10, SC18, TC23, TC5) expressed that when they were not familiar with the interlocutor, they felt more anxious; but if the interlocutors were their friends or classmates, they might feel less anxious. The interlocutor's English proficiency may also influence students' FLSA levels. Students were worried about speaking with those who were at a higher proficiency level (SVR12, TVR32, TVR35, SC6, TC21) or native speakers (SVR30, SC21, TC28), but they were not anxious while speaking with people at a similar proficiency level (e.g. their classmates).

With regard to the causes of FLSA, data reported by Chinese EFL students in my study were different from previous literature (see Section 2.3.3) to some degree. This is mainly because Q1 in my study collected data about students' descriptions on FLSA rather than their reflections of the causes of FLSA.

Figure 6-7 summarises the factors for FLSA expressed by my participants from two perspectives: individual (low perceived proficiency) and social (fear of failure). Specifically, low perceived proficiency is related to students' pronunciation, intonation and accent, expressions and vocabulary, grammatical knowledge, fluency, and listening comprehension ability. Fear of failure encompasses students' fear of making mistakes in grammar, pronunciation, and vocabulary; fear of being unable to understand the interlocutor; and fear of the interlocutor being unable to understand their utterances. Moreover, students may fear losing face, being laughed at or embarrassed, forgetting what they know and panicking while speaking of interlocutors with higher proficiency. Additionally, the causes of FLSA summarised in Figure 6-7 may interweave with each other. A detailed explanation of the

causes of FLSA among EFL learners who participated in my study is presented in the following paragraphs.

Figure 6- 7 *Factors for Chinese EFL learners' FLSA (my own creation based on data analysis)*



Low perceived proficiency

Low perceived proficiency refers to the fact that anxious students view their English proficiency level as low, no matter what their exact proficiency level is. In other words,

participants in my study were anxious about speaking English because they thought they had not acquired enough linguistic knowledge and/or did not have adequate competence to speak English. This is considered an individual factor for FLSA related to the internal perspectives of the foreign language learners. It is similar to the low self-esteem found as a factor for FLSA by previous scholars (see Section 2.3.1.1). Students underestimate their linguistic skills or speaking abilities, and become worried about their deficiencies, such as not having a native-like pronunciation and accent, which resulted in increased levels of FLSA (Alnahidh & Altalhab, 2020; Kitano, 2001). I did not categorise these factors as Low Proficiency, as in Section 2.3.3, but as Low Perceived Proficiency. This is because it is about how students view their own English speaking proficiency rather than the exact proficiency assessed objectively.

In my study, about 64 students felt they had difficulty expressing ideas properly in English, because they lacked vocabulary. For example, SVR34 expressed: *While speaking English with others, I worry the most is the amount of my vocabulary is insufficient to express my ideas clearly.* Similarly, TVR12 stated: *I often felt anxious while I could not find words to express myself in English.* This was also mentioned by students in the SC and TC groups. For example, SC35 could not plan the language in a short time, which results in difficulty in expressing herself. TC15 felt anxious because she did not have enough vocabulary to express her meaning in English. In addition, the lack of vocabulary or expressions led to students' unwillingness to speak English (SVR9, TVR9), panic while speaking (TVR12, TVR6, SC33, TC15), and do not know what to say (SVR4, SVR23, SVR9, TVR13, TVR29, SC35, SC11, SC1, TC32, TC19, TC7).

About 37 students were concerned that their pronunciation, intonation, and accent were not native-like, which resulted in experiencing FLSA. Native-like refers to speaking as a native speaker of English. It seems that correctness of pronunciation and native-like intonation concerned many participants (e.g. SVR10, TVR1, TVR17, SC6, SC35, TC14), and the concern of the pronunciation and intonation results in their fear of not being understood by others. One of the factors mentioned by SVR33 relating to her anxiety in speaking English was: *Concerning my pronunciation is incorrect, and my intonation is not native-like*. SVR14 worried when her pronunciation was different from others, and she would think she had pronounced the word wrongly. SVR7 even regarded her pronunciation as awkward: *I am afraid that I had a very bad pronunciation, sounds awkward and difficult for others to*

understand. With regard to accent, TC28 was afraid that the difference in accent would result in misunderstanding, which caused her to feel nervous while communicating with native speakers of English.

A total of 39 participants mentioned their concern about the use of grammatical knowledge in order to speak correctly. SVR4, SVR33, and SC5 all stated: *I often worried that I would make grammatical mistakes.* These concerns about grammar also cause fear of not being understood by others (e.g. SC7, SC24, TC2). Moreover, approximately 20 students (e.g. SVR1, TVR6, SC14, SC23, TC1, TC26) expressed they could not speak fluently, which resulted in nervousness or anxiety while speaking.

Fear of failure

Fear of failure can be interpreted as anxious students' concern about failure and mistakes in the interaction. It can be seen as a social factor related to the external environment (e.g. interlocutors, audiences). Forty-two participants mentioned that they feared making mistakes, resulting in embarrassment and loss of face. Face is how one is perceived by others, which is important in Chinese culture (see Section 2.3.3 for a detailed explanation). For example, TVR4 stated her fear of making mistakes:

> *I am very anxious because I fear making mistakes while communicating with others, which would give a bad impression to others. I think it is very embarrassing to make mistakes, so I am not willing to speak English in front of others in order to save my face* (TVR4).

Approximately 20 students (e.g. SVR33, SVR25, TVR13, SC7, TC2, TC28) were worried about making grammatical mistakes, 11 students (e.g. SVR6, SC34, SC30, TC9) were concerned about making pronunciation mistakes, and 3 students (SVR31, SVR8, TVR22) feared using incorrect vocabularies. In addition, 31 participants feared not being understood by an interlocutor because of their grammar (e.g. SVR3, SC7), pronunciation (e.g. TVR17, TVR1, TVR2, SC35, SC7, TC28), or vocabulary (e.g. TVR7, SC24). For example, TVR1 mentioned her fear of not being understood by others:

> *I always worry about my pronunciation problems. This is because I don't have standard pronunciation, I fear that others can't understand me when I speak English. Also, my vocabulary is not enough to express myself. This will make it difficult for others to understand me* (TVR1).

It can be seen from TVR1's response that low perceived proficiency seems to interweave with students' fear of making mistakes.

The other concern mentioned by participants was the fear of not understanding others' utterances. 18 students (e.g. SVR28, TVR3, SC35, SC20, TC25) mentioned that they were anxious because they feared not understanding their interlocutor in the conversation. SC20, for example, stated:

> *I am worried that my pronunciation is not correct every time when I speak English. This may make others can't understand me. Moreover, I am afraid I can't understand others, either, which will result in a failure in communication* (SC20).

The above types of anxiety derive from the pressure of the social environment, which was usually accompanied by fear of being laughed at, fear of losing face, and fear of being embarrassed. SC8 was very anxious about speaking English, and expressed her fear of being laughed at by others: *I will avoid speaking English because I think my oral English is very poor, and I am afraid that others will laugh at me if I made mistakes.* Similarly, TC6 stated her avoidance in speaking English: *I rarely communicate with others in English because I am always concerned about making mistakes, and if others cannot understand me, I fear being laughed at by them.* Students feared being embarrassed when they could not find the words during the interaction (e.g. SVR25, SVR15), and had no confidence in being understood by others (e.g. SVR5, TC14).

Above all, the main causes of FLSA expressed by my participants included low perceived oral English proficiency and fear of failure, which overlap each other. However, these two causes covers the individual (internal) and social (external) factors, which is consistent with using VR Exposure-based Therapy and situated learning theory (see Chapter 4) in my study to attempt to alleviate FLSA both internally and externally.

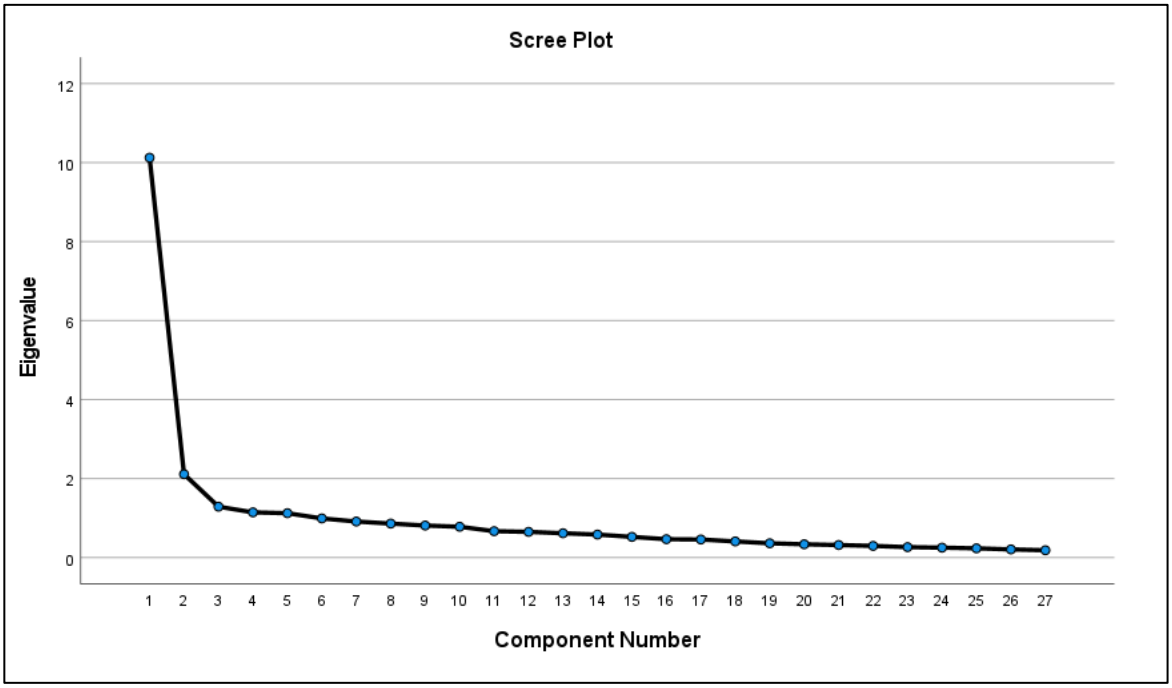
Refining ESAS

As stated in Section 6.2.1.2, factor analysis of the English speaking anxiety scale (ESAS) did not result in an acceptable solution. The two causes of FLSA (low perceived proficiency and fear of failure) identified from the qualitative data could be applied to refine the ESAS in my study. Identifying factor models for ESAS is important because it may help to

understand Chinese EFL learners' FLSA and contribute to a valid construct for a scale to measure it.

A two-factor extraction was conducted via the exploratory factor analysis with principal component analysis (see Section 5.5.1). The results showed that two eigenvalues exceeded explaining 45.31% of the data in total. This was lower than the 60% suggested by Hair (2009) as a valid construct value for cumulative explaining. Therefore, the two-factor model was not accepted as a solution. Moreover, the scree plot (see Figure 6-8) with extracting two factors showed the curve change the direction to horizon after Component 2. It seems Component 1 explains much of the variation in the ESAS than the remaining components. Therefore, the ESAS may be seen as a single scale. This is consistent with the results found in Section 6.2.1.2. However, further investigation of the construct of ESAS is essential, involving a larger data set and obtaining more in-depth data on the causes of FLSA among Chinese EFL learners.

Figure 6- 8 *Scree plot of EFA results with extracting two factors for ESAS*



Question 2 (Q2): Please describe your anxiety about speaking English with others AFTER the learning sessions.

The purpose of Q2 is to gather students' reflections on their FLSA after the intervention in order to investigate whether the intervention has affected students' anxiety about speaking

English. It was noticed that students' answers to Q2 were mainly compared with their descriptions of FLSA before the intervention. Thus, the description of FLSA changes was coded according to Table 6-12 as less anxious, less anxious to some degree and no change. Some students expressed their lessened anxiety positively, using phrases such as *I am more confident to..., I am willing to…* and some negatively such as *I am not worried about*, *I am not so anxious or nervous*. With regard to less anxious to some degree, the typical responses were *my anxiety decreased at some degrees*, *I am a bit more confident to …*, and *I am still anxious but less than before*. For those who experienced no change in FLSA level, some responses were direct (e.g. *did not change, I did not find any change*) but some were indirect (e.g. *I still feel anxious/nervous, I still fear of …*).

Table 6- 12 *Coding scheme for Q2*

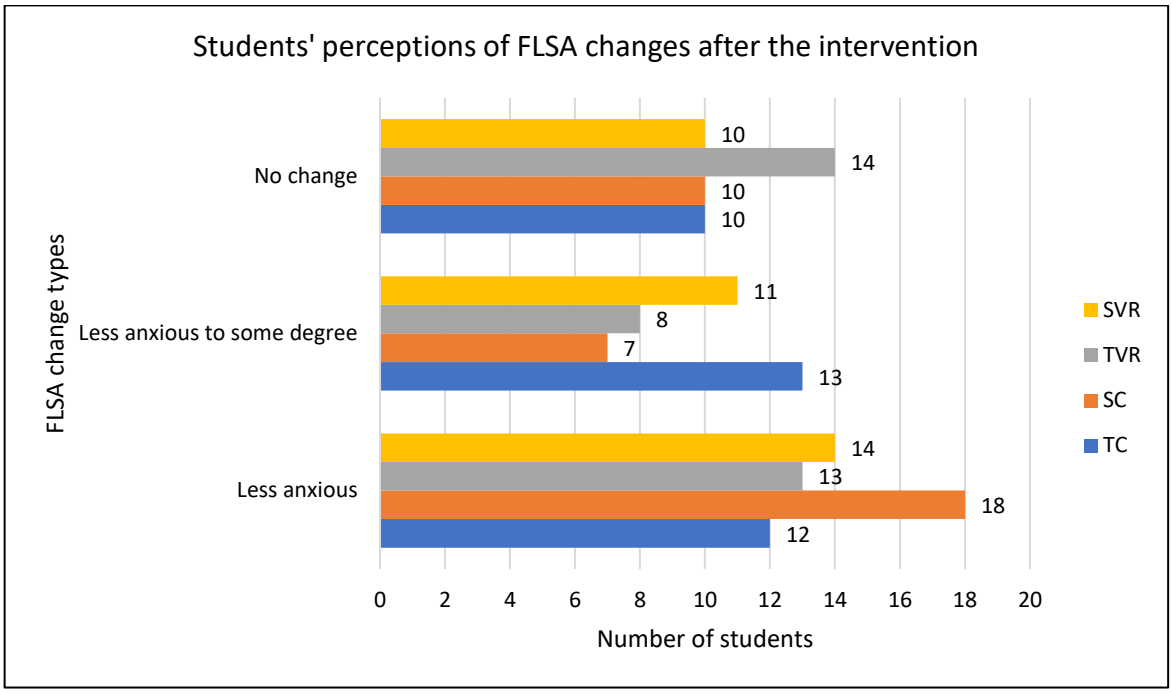
Categories	**Responses examples**
Less anxious	*Nervous/anxious situations reduced, not worried about…, my anxiety levels reduced, I am not so anxious/nervous, more confident, brave/willing to speak English*
Less anxious to some degree	*My anxiety decreased to some degree, I am a bit more confident to …, I am still anxious but a bit less than before*
No change	*Still feel anxious/nervous, still worried about…, still nervous, fear of …, the anxiety about speaking English did not change, I did not find any changes, it is still difficult for me to speak English*

Based on the quantitative content analysis of the responses to the open-ended questionnaire, 40.71% of the students (N = 57) reported a decrease in FLSA level after the intervention, while 27.86% of the students (N = 39) reported feeling less anxious about English speaking some extent. However, about 31.43% of the participants (N = 44) believed there was no change in their FLSA level before and after the intervention. The results are shown in Figure 6-9.

Although it was not possible to examine the statistical significance of FLSA changes with the qualitative data, it is notable that more than 28% of students in each group reported they did not perceive changes in their FLSA levels. This supports the ESAS results showing no

significant changes before and after the intervention. Moreover, the similar number of students who perceived no changes in their FLSA (10 students in each of the SVR, SC, and TC group, and 14 students in TVR group) was mainly consistent with the ESAS result that four groups had no statistically significant difference in FLSA level after the intervention. The specific reason for no statistically significant change in ESAS scores can be found in the qualitative data. This result confirms the necessity of employing the mixed methods approach in my study in order to understand the FLSA changes through different lenses.

Figure 6- 9 *FLSA changes after the intervention based on the qualitative data*



Based on students' responses, the reasons why they perceived no change in general FLSA after the intervention were mainly to do with no perceived improvement in oral English proficiency. Students were still concerned about their vocabulary (e.g. SVR31, SVR34, TVR10, SC11, SC5, TC15), grammar (e.g. SVR25, TVR8, SC1, TC34, TC17), pronunciation (e.g. SVR30, TVR26, SC20, SC9), and listening ability (e.g. TVR30, TVR3, TC11) after the intervention. As the expressions of students' concern about language proficiency were similar to those discussed in Section 6.3.1.1, examples of their responses were not repeated here.

By contrast, some students benefited from the learning sessions in aspect of improving their oral proficiency, such as in the aspect of English expressions (e.g. SVR21, SVR11, SVR1,

TVR25, TVR21, SC6, SC34, TC2). However, other aspects of linguistic competence only received little perceived improvement, including fluency (e.g. SVR27, TVR35, SC28), grammar (SVR4) and pronunciation (TVR17). In terms of expressions, some students (e.g. SVR15, TVR31, TVR32) found a solution during the intervention: they would find a way to express their meaning using simpler words or by paraphrasing it. In particular, SVR9 reported that she was aware that pronunciation and intonation were not the most important elements in communication; the meaning of the utterance was the most important. Acceptance of their own proficiency was also mentioned by some participants, so it seems that some students changed their beliefs about language learning after the intervention. For example, TC19 stated her positive attitude towards her oral English proficiency:

> *I started to accept my elementary level of oral English. Even though I speak slowly, sometimes with incorrect words or non-standard pronunciation, this doesn't matter. What I need to do is express myself bravely. I believe that my oral English will improve after a certain amount of practice* (TC19).

This response is highlighted here because changing foreign language learner's negative emotion about foreign language learning to positive emotions is the expected outcome of my research. Although about one third of the students in each group did not perceive any decrease in their FLSA levels, some of the participants benefited from the intervention.

Factors related to the fear of failure were also mentioned by many participants (e.g. SVR34, SVR12, TVR3, SC35, TC4, TC11), who did not see any reduction in FLSA level after the intervention. In a similar vein, the fear of failure was mostly about students' fear of making mistakes, fear of not being understood, and fear of not understanding others as stated in Section 6.3.1.1. Examples of their responses were therefore omitted here to avoid repetition.

In contrast, some students did not fear of failure after the intervention. For example, SVR3 expressed her attitude towards mistakes as: *I was not nervous as before when I made mistakes, and I would not pay much attention if I made mistakes in grammar during the conversation.* A similar point was made by SVR25, SVR24, TVR2, SC3. Moreover, making mistakes is normal, as SVR7 responded: *Everyone will make mistakes more or less, and oral English learning is a long-term process that needs massive practice*. This also can be seen as a change of learners' attitudes towards the foreign language, which helps to alleviate FLSA.

Where there was no change in FLSA level, some participants in my study attributed this to the shortness of the intervention period. SVR22 stated: *I still feel anxious, and I think improving oral English proficiency needs long-term practice, which then can result in the reduction of FLSA.* SVR14 held the same view: *The duration of learning sessions is not long enough to alleviate my FLSA, because I think the key to reducing FLSA levels is improving oral proficiency, which is difficult to achieve in a short time.* The insufficient number of learning sessions was also mentioned by SVR27, TVR1, TVR26. However, TVR28 held the opposite opinion: she felt less and less anxious after practising speaking many times in the intervention. SC22 also expressed that she benefited from the practice opportunities offered by the intervention, which was enough for her to improve oral fluency.

To sum up, the majority of participants were anxious about speaking English before the experiment. Their concerns related to FLSA, revealed in their responses to the open-ended questionnaire in my study, could be categorised as low perceived proficiency and fear of failure. After the intervention, most of the students experienced a decrease in their general FLSA levels to some extent, but about one third of students in each group did not perceive any change to their general FLSA. This might be interpreted as supporting for the quantitative results of ESAS scores in that no statistically significant changes were revealed. In addition, proficiency level and failure in communication were still the main concerns among participants who did not perceive a reduction in their FLSA levels after the intervention. The shortness of intervention period was also mentioned by a few participants as a reason for no change in their FLSA levels.

6.3.2 FLSA Change in the Speaking Tests (RQ3)

Question 3 (Q3): In fact, did you feel more at ease in the second speaking test compared to the first one? What made you feel this way?

Q3 aims to gather students' perceptions of the changes in their FLSA levels during the speaking tests. Most students gave short answers to this question before explaining the reasons. As shown in Table 6-13, responses such as *Yes, indeed, I felt more relaxed, I am not so anxious as in the first speaking test* were coded as 'more relaxed'. Answers such as *I felt a bit more relaxed, I felt relaxed to some degree in the second speaking test* were coded as 'a bit more relaxed.' Comments such as *I did not feel a big difference in the first and second speaking test, I felt the same degree of anxiety as last time,* and *I did not feel more*

relaxed were coded as 'the same level anxiety'. Comments such as *No, not really, I felt even more anxious than last time, I felt more panic* were coded as 'more anxious'. *I have never felt anxious in speaking tests* and *It was as relaxed as before for me* were coded as 'the same relaxed'.

Table 6- 13 *Coding scheme for Q3*

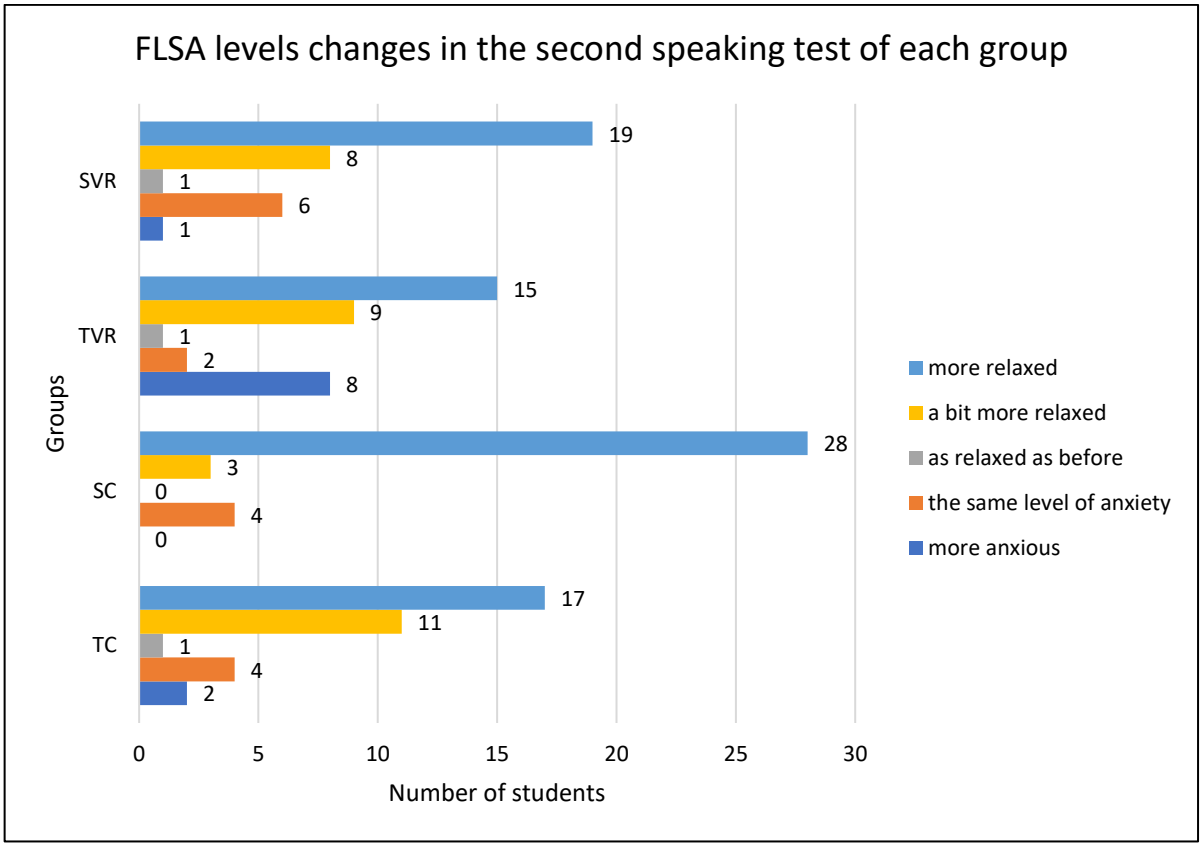
Categories	**Responses examples**
More relaxed	*Yes; indeed; I felt more relaxed; I am not so anxious as in the first speaking test*
A bit more relaxed	*I felt a bit more relaxed; I felt relaxed to some degree in the second speaking test*
The same level of anxiety	*I did not feel a substantial difference between the first and second speaking test; I felt the same degree of anxiety as last time; I did not feel more relaxed*
More anxious	*No; Not really; I felt even more anxious than last time; I felt more panic*
As relaxed as before	*I have never felt anxious in speaking tests; It was as relaxed as before for me*

According to the quantitative content analysis of the texts obtained from the open-ended questionnaire, it was found that of the 140 participants, 78.57% felt more relaxed (N = 79) or a bit more relaxed (N = 31) in the post-test (see Figure 6-10). 13.57% of the students' FLSA levels seemed not to change, either staying the same relaxed as before (N = 3) or the same anxious as before (N = 16).

Moreover, the SC group contained the most students (N = 28) who felt more relaxed in the second speaking test among the four groups, and it was the only group where no students felt more anxious than before the experiment. The TVR group had the smallest number of students (N = 24) who felt more relaxed or a bit more relaxed, but the most students (N = 8) who felt more anxious in the second speaking test. This may mean that students in the TVR group experienced the least alleviation in FLSA in the second speaking test. This is consistent with the quantitative findings with regard to the self-ratings of FLSA levels in

specific tasks (see Section 6.2.2); the SC group witnessed the most decrease in the post-test compared to the pre-test, and the TVR group witnessed the least. That is to say, adopting situated learning in the classroom seemed to help the most in reducing students' FLSA levels in the speaking tests.

Figure 6- 10 *FLSA levels changes in the second speaking test of each group*



In terms of the reasons for the FLSA changes in the second speaking test, students' responses were analysed via an inductive qualitative content analysis approach (see Section 5.6.1.2). The key reasons were categorised in four groups: test familiarity, perceived proficiency improvement, increase in experience, self-confidence development (see Figure 6-11). Test familiarity was the main reason, contributing to the decrease in FLSA in more than half of the students (N = 73). Other than responses on improving perceived proficiency (N = 51) and building self-confidence (N = 73), increase in experience was mentioned by 66 students. However, for those who did not experience any FLSA reduction in second speaking test, reasons included short duration of the intervention, unfamiliarity with the topics, higher expectation of performance, negative feedback from the examiner, and some personal issues.

Figure 6- 11 *Main reasons for FLSA changes in the second speaking tests based on analysing the qualitative data*

Based on students' responses, the reasons for experiencing lower anxiety in the second speaking test could be categorised as: test familiarity, perceived proficiency improvement, increase in experience, and self-confidence development.

Test familiarity

In terms of familiarity with the test, students' responses included being familiar with the interlocutor, the test type, and the test process. This may explain why ESAS results did not show statistically significant differences after the intervention, but self-ratings of FLSA levels in the speaking tests decreased significantly. The ESAS measured general FLSA without considering the task type, but the self-ratings of FLSA were related to anxiety in the specific tasks.

The following statements show the general positive influence of test familiarity on student's FLSA levels. *The first speaking test was unknown which resulted in anxiety, but after experiencing it once, my anxiety reduced* (TVR6). *Based on the experience of the first speaking test, I knew what I needed to do, so I did not feel nervous anymore* (SC32).

Other students pointed out specific aspects of the familiarity of the tests that helped reduce their anxiety levels. Knowing how the test process worked appeared to be the most important factor contributing to students' (N = 25) feeling less anxious during the speaking test. SVR21 stated that he had no knowledge of what the test process would involve in the first test, but once he knew what to expect, he became more relaxed in the second speaking test. Another student expressed similar feelings:

> *For the first speaking test, I did not know what would happen nor what the process would be, which made me feel nervous. After understanding the type of the test, I could prepare for the conversation better, so my anxiety levels reduced in the second speaking test* (TVR8).

Additionally, students experienced lower FLSA levels in the speaking test because of familiarising themselves with the task type. This particularly benefited students from the SVR and SC groups, as the task type in the intervention for these two groups was similar to that of the speaking test. SC12 stated the benefits of the same task type in the speaking test and the lessons to her lower levels of FLSA:

> *For the first speaking test, I didn't know what was going on and what the task would be like. So, when I knew it was a role-play task as I had practised in the English course, I felt more relaxed* (SC12).

SVR1 also expressed: *The task type in the English speaking course is similar to that in the speaking test, so I felt a bit more relaxed*. SC11 explained her lessened anxiety during the

second speaking test: *We have come across several topics in this type of speaking task in class, which is why I felt more relaxed in the second speaking test*.

Familiarity with the interlocutor is another specific contribution to the decrease in FLSA levels. This is consistent with the responses into Q1 showing that students were less anxious while speaking to someone familiar (see Section 6.3.1). For example, SVR13 thought: *It might be due to it being the second time I communicated with the examiner, which made me feel more familiar with her and more relaxed.* Similarly, TVR24 expressed: *In the second speaking test, I was more relaxed as it was not the first time meeting the examiner. I felt familiar with the method*.

Thus, being familiar with the test process, the test type and the interlocutor seemed to be the main reasons for students experiencing lower anxiety in the second speaking test. There will be further discussion on test familiarity and the alleviation of FLSA in specific tasks in Section 7.2.2.1.

Perceived proficiency improvement

As a result of the intervention, 48 students reported that the improvement in their English proficiency helped to reduce their anxiety about speaking English in the test. Acquiring more expressions in English was regarded as the benefit of the intervention by most students (N = 27). Speaking skills improvement was also regarded by 10 students as a reason for FLSA reduction in the second speaking test. A total of 11 students perceived improvements in accent, pronunciation, fluency, and grammar.

SVR32 attributed her lessened anxiety in the second speaking test to the acquisition of expressions: *In the course, I became more familiar with different English expressions, which I could use in the speaking test.* Similarly, SVR4 explained that he experienced less anxiety in the speaking test:

> *The English speaking practice in the immersive environment provided me with many new expressions, which helped to relieve my anxiety as I could use the expressions I learnt in the conversation* (SVR4).

In the TVR group, a few students (TVR18, TVR31, TVR28, TVR29) stated that they had learnt some English expressions in the course, which helped them feel relaxed during the

second speaking test. In the traditional classroom groups (SC and TC), this benefit was also mentioned by some students (SC5, SC26, SC14, TC3, TC32). Other than contributing to alleviating specific FLSA in the speaking test, the improvement in English expressions also lessened to some students' general FLSA (see Section 6.3.1.2). Therefore, it seems acquiring more English expressions is important for helping Chinese EFL learners cope with FLSA.

With regard to speaking skills, students stated that they had learned how to continue a conversation in the intervention. This helped to reduce their FLSA levels in the second speaking test. SVR9 mentioned that she was able to understand the logic of how English is spoken through practising in the course. This resulted in her no longer thinking about the answers in Chinese and translating them into English while speaking. She therefore spoke English in a more natural way, which made her feel more relaxed during the speaking test.

SVR19 also specified her improvement in speaking skills:

> *Every time I practise English speaking in the VR course, I understand how a topic can be discussed and expanded in an English way. The next time I am in an English conversation with others, I will think about what I have learnt, and continue the conversation in a similar way by expanding the content rather than talking about very simple aspects* (SVR19).

In the course, students from the TVR group (TVR26, TVR2) also mentioned that they had learned how to answer questions and express their own opinions during the course, which alleviated their anxiety about speaking English. In the SC group, this point was mentioned by SC17:

> *I become less nervous because I know what to say, and my logic to speak English becomes clearer after practising. I can follow the topic and continue the conversation with others, and sometimes I can lead the conversation by myself. This alleviated my anxiety significantly* (SC17).

In addition, some students (e.g. TVR31, SC14, SC5) felt more relaxed in the second speaking test because they thought their pronunciation of some English words had been corrected in the course.

Increase in experience

Some students (N = 51) stated that they experienced less anxiety during the second speaking test because they had adequate practice during the intervention. Specifically, they encountered various interlocutors and topics, which increased their experience of speaking English. Many students in four groups (e.g. SVR11, SVR15, SVR26, TVR13, TVR19, TVR29, SC1, SC26, SC24, TC16, TC28) noted that they were used to this type of conversation due to the practice in the intervention.

Some students in the TC group stated that they had benefited from the English course due to the particular task type of role-play. For example, TC11 reported:

> *I have practised many role-play tasks with different partners during the English course. It is similar to the task used in the speaking test. This is why my anxiety in the second speaking test reduced* (TC11).

Similarly, TC10 noted:

> *Compared to the role-play task in the first speaking test, I was calmer in the second one. I believe this should be attributed to a number of role-play task practices I have done in the course* (TC10).

Nonetheless, the specific task type (role-play) was not mentioned by participants in the other three groups (SVR, TVR, SC). This may be because the exact task type was not explicitly emphasised in the intervention for these three groups as much as for the TC group. In the TC group, the instructor (me) told the students the exact task type (role-play) that was being used in class. However, it is notable that, even without knowing the exact task type, students in these three groups (SVR, TVR, TC) still experienced a reduction in their FLSA levels as a result of the increased practice offered by the course.

Interestingly, TVR7 did not understand the topic in the second speaking test well, but she still felt less anxious during it. She explained why:

> *Because I have had many opportunities to practise spontaneous English speaking in the course, I felt a bit more relaxed in the second speaking test. Although I did not understand the topic clearly, I could continue the conversation with the examiner* (TVR7).

The above results support the findings of Amiryousefi (2016) that task repetition helped reduce FLSA levels in similar tasks. My results are also consistent with the finding by

Thrasher (2022) that an increase in speaking the target language led to a lower FLSA level. A detailed discussion of the potential of task repetition for alleviating FLSA is presented in Section 7.2.2.1.

Self-confidence development

About 56 students mentioned that they felt less anxious in the second speaking test because their self-confidence in speaking English was built during the intervention. Some students remarked on their increased self-confidence: *I become more confident than before* (SVR24), and *I feel more confident to speak English and feel less anxious* (SVR6). TVR27 described her increased self-confidence in detail:

> *Psychologically, I felt more relaxed during the speaking test. As I have practised many times in the course, I believe that I have the ability to communicate in English. So, I became more confident during the second speaking test* (TVR27).

This point was also raised by some students in the SC and TC groups (e.g. SC3, SC20, TC25, TC31). Interestingly, TC23 was not familiar with the topic of the second speaking test, but she still felt less anxious because her confidence was built during the course. This is important because the familiarity with the topic can influence FLSA levels positively (see Section 2.3.3). However, after the intervention used in my study, students may feel similar confidence when encountering unfamiliar topics. TC23 responded to this question with:

> *Personally, I am braver about speaking English after taking the course. Although I was not familiar with the topic in the second speaking test, I did not feel so anxious, and I could continue the conversation confidently* (TC23).

The development in self-confidence also contributed to some students' (e.g. SVR13, TVR13, TC14) willingness to communicate, which helped to reduce their anxiety levels in the speaking test. SVR7 attributed her relaxation in the second speaking test to her willingness to communicate: *I am not as anxious as in the first speaking test and become willing to take the speaking test after taking the English course.*

Self-confidence grew from no longer being afraid of making mistakes after the intervention. As stated by TVR11, *I realise that making some minor mistakes is not a problem, as long as the meaning of my sentence can be understood by the listeners*. In a similar manner, TVR9 mentioned: *I felt more relaxed because I knew I only needed to speak, and did not need to*

worry about making mistakes. SC8 was also very anxious in her first speaking test, but she became less anxious in the second one because she did not care so much about making mistakes.

It is interesting that the HiVR technology was not directly mentioned by any participants in the SVR and TVR groups as contributing to their FLSA decrease in the speaking tests. The majority of responses explaining lowered FLSA in the speaking tests were related to the content of learning sessions rather than the learning environment. Specifically, they mentioned being familiar with the test, including the test type, test process and the interlocuters/examiners. Additionally, students perceived English proficiency improved, their experience of the same type of speaking task increased, and their self-confidence in speaking English was built during the intervention. All these helped to lower students' FLSA levels in the second speaking test.

Based on the collected data, reasons for experiencing the same or higher anxiety levels in the second speaking test could be categorised into five groups: short duration of the intervention, unfamiliarity with the topics, higher expectations, negative feedback from the examiner, and personal issues. Three students (SVR15, SVR34, TVR10) thought the duration of the course was too short to improve their English proficiency, and they did not find it alleviate their FLSA. This was also reported by a few students in their assessment of their general FLSA levels after the intervention (see Section 6.3.1.2). Therefore, longer duration of the intervention should be considered in future studies.

In addition, 20 students found that the topics in the second speaking test were difficult or unfamiliar, which resulted in higher levels of anxiety. For example, SVR25 expressed: *I did not know what to say at the second speaking test because the topic was difficult for me, which made me feel very anxious.* TVR30 explained that she was more nervous in the second speaking test because she did not understand the topic. TC32 specified the reason for her anxiety in the second speaking test: *The topic of the first speaking test was familiar to me, but the second one was not. I felt very nervous and could not express myself fluently.*

Other than difficulty or unfamiliarity with the topics, a few students (N = 3) had higher expectations for their oral performance during the speaking tests because they had attended the intervention. Some students attributed their higher anxiety in the second speaking test to

personal reasons, such as feeling not well (TVR4), being tired (SVR22), and being unable to focus (SC10).

The influence of the interlocutor (examiner) on students' anxiety levels was also noted by some students (N = 8). TVR10 reported that she found the examiner more serious the second time, which made her nervous. Two students (TVR23, SVR25) had difficulty understanding the accent of the examiner. Other students (e.g. SVR35, TVR17, TC19) became anxious during the speaking test just because the examiner was a native speaker of English.

In conclusion, the majority of students (N = 110) felt less anxious in the second speaking test than in the first one. The SC group had the largest number of students (N = 28) who experienced less anxiety during the second speaking test, while the TVR group had the most students (N = 8) who experienced higher anxiety during the post-test. This is consistent with the quantitative results of the self-ratings of FLSA in the speaking test, which revealed that the FLSA levels in the SC group reduced the most (from 5.83 to 4.28) and the levels in the TVR group reduced the least (from 5.82 to 5.21) (see Section 6.2.2). Familiarity with the test, including being familiar with the test process, test type and examiner, was the main reason for students feeling less anxious during the test. The increase in practice opportunities in the intervention improved students' perceived proficiency, developed their self-confidence, and gave them experience of similar tasks. With regard to students (N = 30) who experienced the same or higher anxiety levels in the second speaking test, unfamiliarity with the topic was the reason given by the highest number of students (N = 20). Other reasons given included the short duration of the intervention, higher expectations of their performance, negative feedback from the examiner, and some personal issues.

6.3.3 Integration of Data (RQ 3)

Overall, 140 students responded to the ESAS before and after the intervention and rated their anxiety levels for the first and second English speaking tests. This generated the quantitative data for their general and specific FLSA levels respectively. All participants answered the first three questions in the open-ended questionnaire about their general FLSA changes before and after the intervention as well as specific FLSA during the second English speaking test. The joint display depicted in Table 6-14 shows how the quantitative results were integrated with the qualitative results. This answers the second part of RQ3 (see below) that how qualitative results of Q1, Q2 and Q3 explain or refute the quantitative results.

RQ 3: How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?

The ESAS scores demonstrated that interventions using HiVR did not result in statistically significant changes, no matter which type of learning approach (situated learning or teacher-centred learning) was applied [quantitative results, hereafter, QUAN] (see Section 6.2.1.3). Moreover, all four groups showed no difference in their general FLSA levels measured by ESAS after the intervention [QUAN] (see Section 6.2.1.3). These findings were supported by the qualitative data that about one third of the participants in each group did not perceive changes in their general FLSA levels [qualitative results, hereafter, QUAL] (see Section 6.3.1.2). Therefore, it seems that neither the learning environment (HiVR or classroom) nor the learning approach (situated learning or teacher-centred learning) makes a difference to students' general FLSA levels in my study.

In addition, the main causes of Chinese EFL learners' FLSA were low perceived proficiency and fear of failure [QUAL] (see Section 6.3.1.1), based on the qualitative data analysis. However, there was not enough evidence to prove the two-factor model through the exploratory factor analysis results of the ESAS data set [QUAN] (see Section 6.3.1.1). Further investigation of the construct of FLSA among Chinese EFL learners is essential.

With regard to the self-ratings of FLSA levels in the speaking test, both the SVR and TVR groups witnessed statistically significant decreases in FLSA level in the second speaking test compared with the first [QUAN] (see Section 6.2.2.2). However, the SC group witnessed the largest decrease in specific FLSA levels and the TVR group seemed to decrease the least of the four groups [QUAN] (see Section 6.2.2.2). These findings were supported by the [QUAL] findings in Section 6.3.2 that over 68% of students in the SVR and TVR groups mentioned that they felt less anxious in the second speaking test. The SC group contained the most students (N = 28) expressing this feeling, and the TVR group contained the least (N = 17). Nevertheless, the self-ratings of FLSA levels in the second speaking test were not statistically significantly different among the four groups [QUAN] (see Section 6.2.2.2), which means the degree of decrease in specific FLSA levels in each group was not statistically different. That is to say, learning environments (HiVR and classroom) and learning approaches (situated learning and teacher-centred learning) did not result in

statistically different influences on students' specific FLSA levels in the speaking test in my study.

According to students' responses to the open-ended questionnaires, the main reason for a specific FLSA decrease in the second speaking test was that students were more familiar with the speaking test, in aspects of test type, test process and examiner [QUAL] (see Section 6.3.2.1). Increase in experience, oral proficiency improvement and self-confidence development also contributed to some participants' FLSA decrease in the second speaking test [QUAL] (see Section 6.3.2.1). Some reasons that were attributed to the increase or no change in specific FLSA in the second speaking test should also be noted. These reasons include unfamiliarity with the topic, short duration of the intervention, higher expectations of the oral performance, negative feedback from the examiner and some personal issues [QUAL] (see Section 6.3.2.2).

Table 6- 14 *Joint display of quantitative and qualitative findings of my study for RQ3*

Groups	**QUAN**	**QUAL**	**Integration**
	FLSA levels	**Open-ended questionnaires**	**MMR analysis**
	ESAS total scores		
SVR TVR	89.34 (pre); 89.91 (post) 89.34 (pre); 90.57 (post) No statistically significant changes (RQ1)	28.57% of students in SVR (N = 10) and 40% of students in TVR (N = 14) perceived no change in their general FLSA levels after the intervention.	The total ESAS score change [QUAN] was supported by the [QUAL] findings that around 30% to 40% of students in the SVR and TVR groups perceived no change in their general FLSA levels after the intervention.
SVR TVR SC TC	89.91 (post) 90.57 (post) 89.20 (post) 90.94 (post) No statistically significant differences (RQ2)	The number of students who stated that they felt less anxious and a bit less anxious after the intervention were similar in each group: SVR (N = 25), TVR (N = 21), SC (N = 25), TC (N = 25).	No statistically significant differences in post-test ESAS scores [QUAN] were supported by the [QUAL] finding that a similar number of students in each group stated they were less anxious or a bit less anxious after the intervention.
	Self-ratings of FLSA levels in specific tasks		

SVR TVR	5.85 (pre); 4.99 (post) 5.82 (pre); 5.21 (post) Statistically significant decrease (RQ1)	77.14% of students in SVR (N = 27) felt less anxious in the second speaking test. 68.57% of students in TVR (N = 24) felt less anxious in the second speaking test.	The significant decrease in self-ratings of FLSA levels in the speaking test [QUAN] was supported by the [QUAL] finding that over 68% of students in SVR and TVR felt less anxious in the second speaking test.
SVR TVR SC TC	4.99 (post) 5.21 (post) 4.28 (post) 4.95 (post) No statistically significant differences found among the four groups' post-test specific FLSA levels, but SC decreased the most and TVR decreased the least (RQ2)	SC contained the most students (N = 31) who felt less anxious in the second speaking test, and TVR contained the least. Test familiarity was reported as the main reason for decrease in specific FLSA levels', followed by increase in experience, proficiency improvement, and self-confidence development.	Although no statistically significant differences in self-ratings of FLSA level in the second speaking test were found in the four groups [QUAN], the [QUAL] finding supported the fact that the SC group contained the most students experiencing less anxiety in the second speaking test, and the TVR group contained the least.

6.3.4 HiVR and FLSA (RQ 4)

Question 4 (Q4): What do you think was the impact of immersive VR on your anxiety about speaking English?

Q4 in the open-ended questionnaire aims to investigate students' perceptions of HiVR and FLSA in order to answer RQ4 (see below). The open-ended questions in this (Q4) and the following section (Q5) (see Section 6.3.5) were only delivered to the HiVR groups (SVR and TVR), not to the entire sample.

RQ4: What are students' perceptions of the impact of HiVR on FLSA?

Inductive qualitative content analysis (see Section 5.7.1.2) was applied to analyse the data. The coding scheme for Q4 is presented in Table 6-15. Responses mentioning a decrease in FLSA levels or improvement in self-confidence in English speaking were coded as positive perceptions. Responses stating no change or no big difference in FLSA levels were regarded as negative perceptions.

Table 6- 15 *Coding scheme for Q4*

Categories	**Response examples**
Positive perceptions	*Alleviated/reduced my anxiety, less anxious, not as anxious as before, more relaxed, have positive influence, have/improve confidence in*
Negative perceptions	*No noteworthy influence, no change, seem to be no big difference*

Answers to Q4 and Q5 were missing from TVR1, possibly because she did not select the correction option to indicate that she was from the HiVR group. In the online version of my open-ended questionnaire, three questions were displayed for non-HiVR groups, while five questions would be displayed for HiVR groups. It is possible that TVR1 made a mistake when selecting which group she belonged to, resulting in her not being able to see the latter two questions (Q4 and Q5). Responses from 69 students were analysed for these two questions. The quantitative content analysis of the responses revealed that 89.86% of students (N = 62) in the SVR and TVR groups believed HiVR had a positive influence on

alleviating their FLSA. In total, the SVR group had more students (N = 32) holding a positive position than the TVR group (N = 24). There are overlaps in responses, so the total number of responses for each element does not equate to the exact number of students in each group (N = 35). Generally, students' positive perceptions of HiVR and FLSA can be summarised as follows: 1) creating an authentic environment, 2) raising concentration on the task, 3) improving perceived oral proficiency, 4) providing more practice, 5) being interesting and relaxing, 6) being unable to see the real person. Among the seven students (two students in SVR, five students in TVR) who believed HiVR did not help with their FLSA, the reasons were mainly the short duration of the intervention and being unable to see others.

Students' positive perceptions of HiVR and FLSA could be categorised as: 1) creating an authentic environment, 2) raising concentration on the task, 3) improving perceived oral proficiency, 4) providing more practice, 5) being interesting and relaxing, and 6) unable to see the real person.

Creating an authentic environment

A total of 20 students commented that HiVR learning sessions reduced their anxiety about speaking English by creating an authentic environment. This is in line with the use of situated learning theory as a model for learning design, which provides an authentic context and authentic activities for learning (see Section 4.2). Several students in the SVR group (e.g. SVR10, SVR22, SVR27) and the TVR group (e.g. TVR5, TVR24, TVR31) reported that they benefitted from being immersed in an environment that simulated real-world situations while using HiVR. Taking the response from SVR27 as an example, he mentioned that *Immersive VR eased my anxiety levels because the virtual but authentic environment in it prepared me for the possible situations that I may face in a real-life conversation.* SVR7 reported that she felt less anxious due to the immersive VR environment, which made it felt as though she was in a real-life scenario. This response could be linked to the sense of presence, which is an affordance of HiVR in learning (see Section 3.2.1). More discussion on this can be found in Chapter 7. A similar point was made by SVR34:

> *By utilising immersive VR, we are able to solve a major problem of speaking English outside of class, namely the environment. Immersive VR provides us with various authentic environments, and we can be immersed in them while using it, which helps to reduce my anxiety about speaking English* (SVR34).

TVR19 felt that HiVR provided an authentic scenario for the conversation, which enabled her to understand the current situation for the conversation immediately. Similar comments were found in the response of TVR14: *Immersive VR facilitates an experience of some situations, which alleviates my anxiety when confronted with similar circumstances in reality*. These are all benefits of creating an authentic environment in HiVR on alleviating FLSA.

Raising concentration on the task

In addition to the authentic environment, some students reported the HiVR sessions helped them focus on the tasks, which resulted in a lower level of FLSA. SVR1 mentioned a benefit of HiVR for concentrating on a task as follows.

> *Observing the immersive virtual environment enables me to focus on the learning tasks, which enhances my concentration on speaking rather than feeling anxious about looking at others' eyes* (SVR1).

SVR2 attributed the positive effect that HiVR had on reducing FLSA to the enhancement in task concentration as below:

> *While using HiVR, my attention was drawn to completing the task through communicating with my partner. I sometimes forgot that I was doing an English course, which resulted in lower anxiety about my speaking proficiency* (SVR2).

Similar points were mentioned by several other students (e.g. SVR24, SVR33, SVR7). This is related to the potential of HiVR to enhance learning engagement as discussed in Section 3.3.1.2, and will be further discussed in Chapter 7.

In addition, the benefit of focusing on the task while using HiVR to find relevant expressions was noted by some students in the SVR group (e.g. SVR3, SVR9). A few students in the TVR group also mentioned this point. For example, TVR15 stated:

> *I was fully engaged in the virtual situation while using immersive VR and was able to find the words to communicate. My anxiety levels were reduced because of this. I could try to express myself rather than having nothing to say, unless I did not understand the sentences at all* (TVR15).

Improving perceived oral proficiency

Many students thought that HiVR learning sessions improved their oral proficiency, which resulted in a decrease in their FLSA. The perceived enhancement in oral proficiency is in line with the findings of previous studies (Liaw, 2019; Soto *et al.*, 2020) on HiVR and oral proficiency (see Section 3.3.1.4). Most students (e.g. SVR26, SVR12, SVR23, SVR31, TVR28, and TVR26) found the HiVR course helped them acquire more expressions and vocabulary and develop their speaking skills. This is consistent with the responses they provided for Q2 (see Section 6.3.1.2) and Q3 (see Section 6.3.2.1), so it is not repeated here. Nevertheless, it should be highlighted that the improvement in oral proficiency was based on students' own perceptions rather than objective assessments of linguistic knowledge.

In addition to learning new expressions, SVR4 and SVR8 noted that acquiring new sentence structures also helped to alleviate their FLSA. Four students (SVR20, TVR5, TVR17, TVR21) also mentioned that HiVR was beneficial to their memory, which improved their proficiency in English and reduced their FLSA levels accordingly. TVR21 gave a detailed response:

> *Through immersive VR, I am able to connect the English expressions to real-life scenarios, which enhances my ability to remember what I have learnt. Moreover, after applying what I know in the VR course via communicating with my partner to solve a real-life problem, I can remember more about the learning content* (TVR21).

To sum up, learning sessions in HiVR seemed to improve students' perceived language proficiency in expressions and vocabulary (e.g. SVR26, SVR12, TVR28 and TVR26), sentence structures (e.g. SVR4, SVR8), speaking skills (e.g. SVR15, TVR26), and memorisation of the learning content (e.g. SVR20, TVR5, TVR21). As perceived language proficiency was the main factor for FLSA among Chinese EFL learners (see Section 6.3.1.1), this could help to alleviate FLSA.

Providing more practice

Some students mentioned the increase in opportunities to practise English while attending HiVR learning sessions as a benefit that helped to reduce their FLSA levels. This is consistent with increase in experience as reflected by participants for Q3 as a reason for feeling less anxious during the second speaking test (see Section 6.3.2.1). Similar responses are not repeated here, but further discussion can be found in Chapter 7.

TVR24 expressed that the authentic environment in HiVR and the specific topics of each learning task helped her to practice oral English. This resulted in her perceiving a decrease in FLSA levels. TVR27 also mentioned that the HiVR environment provided more situations in which she could practise speaking than the classroom setting:

> *Immersive VR had a positive effect on my anxiety. Most of our speaking practice was implemented in a classroom setting, which provided limited situations. However, immersive VR simulates more real-life situations. The more I practise it, the more I know how to speak in real life* (TVR27).

Hence, based on my data, the simulation of various real-life situations in HiVR brought more opportunities for Chinese EFL learners to practice oral English, which seemed to reduce their FLSA levels.

Being interesting and relaxing

According to students' responses, another way in which HiVR helped to alleviate FLSA was that its learning content was interesting and relaxing. This is consistent with the potential of HiVR to provide a low-anxiety environment (see Section 3.3.1.3). Moreover, it could be attributed to using VR Exposure-based Therapy (see Section 4.1) in my learning design, which excluded negative feedback or response. More discussion of this will be presented in Chapter 7.

SVR18 stated that she felt more relaxed because the conversation situation is closer to real life in HiVR rather than in a textbook. SVR32 found the HiVR learning to be interesting, because: *It offers a wide range of scenarios and topics, combined with several conversation examples. In addition, I can also converse freely with my partner in the VR course.* SVR9 mentioned that it was more enjoyable to speak English in the HiVR environment rather than in the classroom. However, only one student in TVR group (TVR6) mentioned this point. This may explain why the TVR group contained fewer students who reported feeling less anxious after the intervention (see Section 6.3.2). This may be due to the different learning approaches used in the HiVR learning design for the SVR and TVR groups. Specifically, it seems my participants perceived situated learning approach more interesting and enjoyable.

According to students' responses, the main reason they found the HiVR lessons low-anxiety was that they were less afraid of making mistakes while using HiVR. This may be attributed

to anonymity in the HiVR environment (discussed in the next section in detail) and the high engagement in learning. For example, TVR9 stated:

> *When I face the avatars, I do not have to worry about making mistakes or the need to speak fluently. I am able to engage fully in the scenario and speak as freely as possible* (TVR9).

Similarly, SVR35 expressed: *While using VR, I focused more on how to communicate rather than being afraid of making mistakes, which reduced my anxiety.*

The relaxing experience of speaking English in HiVR also benefited from the appropriate difficulty of the tasks and familiarity with the interlocutors. For example, SVR5 explained her increase in confidence due to the appropriate level of task difficulty:

> *This immersive VR course simulates real-life situations, and the difficulty of the tasks is appropriate for me. In my opinion, it is easier than taking traditional oral English classes, and I do not feel as exhausted as when taking traditional oral English classes. Furthermore, my interlocutors are students at similar English proficiency levels to mine, so I am able to complete each class in a relaxed manner. As a result, I am becoming more relaxed about speaking English* (SVR5).

To sum up, based on students' responses, the relaxing and interesting learning content in HiVR designed for my study underlying VR exposure-based therapy brought most of them a low-anxiety environment. This helped to decrease their FLSA levels. However, the differences in opinions from those in the SVR and TVR groups showed that the situated learning approach may contribute more to this than the teacher-centred approach.

Being unable to see the real person

According to the responses obtained from the Q4 in the open-ended questionnaire, the fact that being unable to see the actual person in HiVR helped alleviate some anxiety about speaking English. This could be linked with anonymity, which is highlighted as a feature of the virtual environment contributing to the decrease in language anxiety levels (Hammick & Lee, 2014). In particular, with avatars, students usually could not be identified in the HiVR environment nor see the exact interlocutor. Despite, in my fieldwork, students were gathered in the same room when they were taking the HiVR learning sessions and they met the real person who they were talking to before wearing the head-mounted display, this seemed not to influence their experience of anonymity in a virtual environment. SVR21 commented that

I could only see the environment and virtual people in VR, which is less stressful than communicating with people face-to-face. A similar point was made by TVR35; she could only hear the voice rather than see the real person, which helped her to focus on the virtual environment. This lowered her FLSA level.

Many students (SVR35, SVR16, TVR20, TVR25) felt less anxious because they could not see the interlocutor's facial expressions in HiVR. According to TVR7, who is a student with social anxiety disorders, the HiVR learning sessions have greatly reduced her anxiety about speaking English due to this. It seems that when wearing the head-mounted display, students were not distracted from the real world but still had a high sense of presence, even though they met the real interlocutors before starting the HiVR session. However, there were some opposite opinions from students, and these are illustrated in the section below (Section 6.3.4.2).

In conclusion, my participants' positive perception of HiVR and FLSA revealed that HiVR could help creating an interesting, relaxing, authentic environment for oral English practice. This helped to raise learners' concentration on the tasks and had the potential to improve students' perceived language proficiency as well as providing more practice opportunities. Moreover, the anonymity of using HiVR was seen by some students as a benefit for FLSA alleviation, but other participants felt the opposite (Section 6.3.4.2).

Despite the majority of participants viewing the HiVR learning sessions as beneficial for alleviating their FLSA levels, a few (N = 7) students viewed them negatively. These opinions can be roughly categorised into two groups: short duration of the intervention and the inability to see the real person.

Four students gave the short duration of the intervention as a reason for their perception that HiVR lessons did not make a difference to their FLSA. This reason was also mentioned by some participants to explain why they perceived no change in their general FLSA (see Section 6.3.1.2). As this is not directly related to HiVR technology or the learning content, it is not repeated in this section.

In addition, being unable to see the real person was regarded negatively by some students. SVR21 mentioned both the positive and negative effects of not being able to read facial expressions in HiVR:

> *Immersive VR is an advanced technology, in which we cannot see the face of the interlocutor, but only hear their voice. This reduced my anxiety about speaking English. However, sometimes it is a disadvantage that we cannot see the interlocutor because we cannot tell their facial expressions in order to understand the intention of their utterances* (SVR21).

TVR29 also mentioned that she preferred face-to-face communication rather than with avatars, although she knew it was a real person speaking. This is consistent with the theme 'lack of paralinguistic cues' found by Kaplan-Rakowski and Gruber (2021) as a disadvantage of HiVR in foreign language learning.

A different effect of FLSA while using HiVR and in real life was illuminated by TVR30:

> *The immersive VR makes me, who has social anxiety disorders, feel relaxed while using it because I cannot see others' faces. However, when removing VR headsets and going back to the real life, my anxiety comes back* (TVR30).

This is contradictory to the finding by Kampmann *et al.* (2016) that the effects of VR Exposure-based Therapy could be transit to the real-life situations. More discussion on the lack of facial expressions in HiVR and the transition to real-life scenarios is presented in Chapter 7.

Above all, the answer to RQ4 would be that most students had a positive view of the impact of HiVR on FLSA, although a few students found no difference in their FLSA after taking the HiVR English course. The positive perceptions of HiVR and FLSA could be summarised as 1) creating an authentic environment, 2) raising concentration on the task, 3) improving perceived oral proficiency, 4) providing more practice, 5) being interesting and relaxing, 6) being unable to see the real person. The negative opinions about HiVR and FLSA mainly concerned the short duration of the intervention and the lack of facial expressions.

6.3.5 HiVR in Oral English Learning (RQ 5)

Question 5 (Q5): In your experience, what is your opinion of the advantages and disadvantages of using immersive VR for the oral English course?

Q5 aims to collect data on students' perceptions of the advantages and disadvantages of using HiVR for oral English learning based on their experiences in the intervention. This

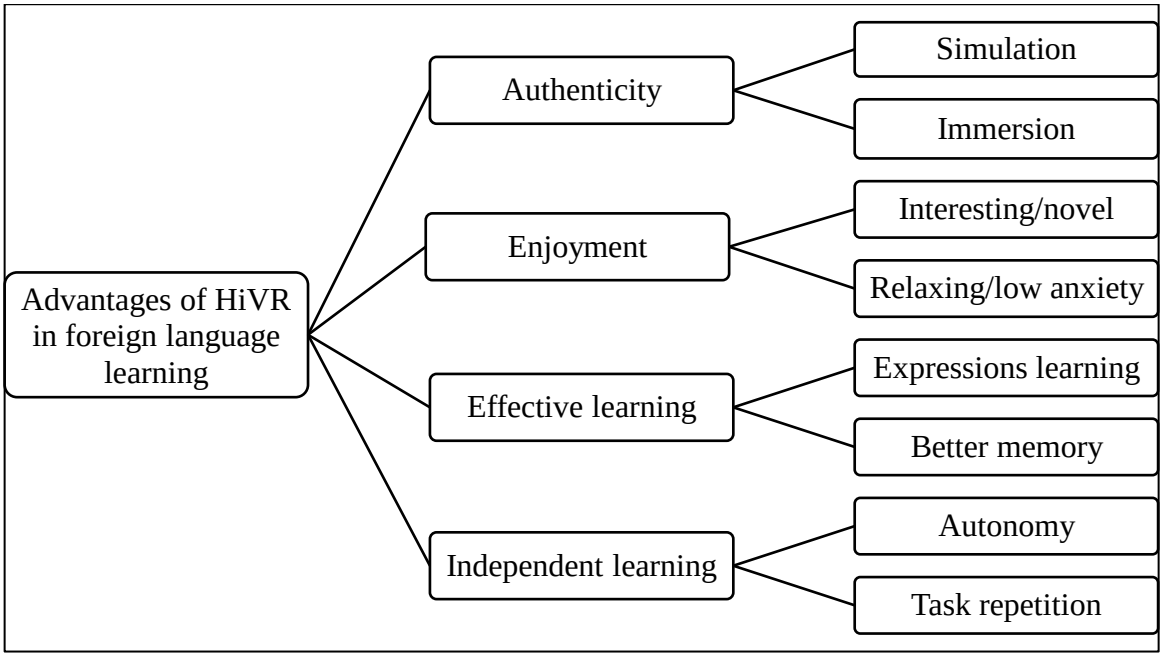
question was designed to answer RQ5 (see below). It was delivered only to students in the HiVR groups (SVR and TVR) rather than the entire sample.

RQ 5: What are students' experiences of using HiVR in oral English learning?

The method used to analyse responses of Q5 was inductive qualitative content analysis (see Section 5.7.1.2). Since Q5 asked about both pros and cons, students gave answers to each aspect separately. Hence, there was no coding scheme for Q5. The responses were grouped and categorised after coding.

With regard to the advantages of using HiVR for oral English learning, the responses could be divided into four categories (see Figure 6-13): authenticity, enjoyment, effective learning, and independent learning.

Figure 6- 12 *Advantages of HiVR in oral English learning based on my data*



Authenticity

About 39 students mentioned *authenticity* in their responses as the advantage of HiVR in oral English learning. The main focus was on simulating real-life situations and immersing participants in them. This is similar to the point that creating an authentic environment using HiVR contributes to alleviate FLSA (see Section 6.3.4.1). Specifically, students stated that

HiVR *simulates real-life scenarios* (e.g. SVR27, SVR26, TVR6, TVR21, TVR34), and *increases the authenticity of the virtual environment* (e.g. SVR11, SVR17, SVR5, TVR5, TVR13, TVR29).

SVR33 gave a detailed response stating: *The virtual environment provides an authentic space for communication, and the design of the scene and the situations increases the authenticity of the conversation.* SVR29 added [*it helped] students to understand the real situation in foreign countries*. TVR34 mentioned the low cost of using HiVR to engage students in an authentic environment, although the environment is virtual. SVR1 compared HiVR course with traditional oral English class: *In comparison with a traditional oral English class, immersive VR provides a more authentic learning environment by simulating real-life situations.*

However, the authenticity of the virtual environment was critiqued by a few students, and this is discussed in Section 6.3.5.2.

Enjoyment

The majority of students (N = 40) found the HiVR lessons enjoyable, and this was viewed as an advantage of HiVR for oral English learning. This point is similar to students' responses for Q4 (see Section 6.3.4). Students found the HiVR learning design *interesting* (e.g. SVR22, SVR15, SVR34, TVR2, TVR31, TVR35), *novel* (e.g. SVR9, SVR24, TVR3, TVR5, TVR8, TVR27), and *relaxing* (e.g. SVR13, SVR27, SVR31, TVR16, TVR30).

Some students' responses added to what discussed in Section 6.3.4. TVR32 stated: *The advantage of an immersive VR lesson is that it is not boring but can intrigue me to learn English.* This is consistent with previous findings that HiVR can enhance foreign language learners' motivation and willingness to communicate (see Section 3.3.1.3). According to SVR9, *the immersive VR course was interesting and close to real-life situations, so I did not take it as a serious task but rather enjoy it*. Noting the advantages of HiVR, SVR5 stated: *immersion VR courses are interesting because the conversation tasks are similar to games*. TVR14 also mentioned that she could learn while playing in HiVR, which she though is impressive. Some students (TVR26, TVR28) enjoyed the HiVR course because it provided vocabulary and expressions, which were helpful for continuing the conversation in the

learning tasks. The vocabulary and expressions provided in HiVR would also increase students' willingness to communicate (SVR15).

Furthermore, as discussed in Section 6.3.4.1, being unable to see the real person eliminates some students' anxiety while speaking English. With HiVR, SVR25 expressed that she did not feel so embarrassed when she was unsure what to say, because she could not see other people's eyes. In a similar response, TVR22 stated that the advantage of using VR was: *I will not be afraid of making mistakes while not facing a real person*. SVR25 gave no fear of being embarrassed and of making mistakes as the advantages of being unable to see the real person in HiVR. However, it is understandable that some students did not find the learning sessions enjoyable or find it helpful to not be able to see the real person. The detailed reasons for these responses are discussed in Section 6.3.5.2.

Effective learning

Students also reflected on using HiVR to learn linguistic knowledge. A few students (N = 15) viewed it as an effective platform for improving oral proficiency. Besides similar points as stated in Section 6.3.4.1, SVR34 and TVR23 both expressed interest in learning more expressions that could be applied in real-life scenarios through HiVR. Some students (SVR20, SVR35, TVR5, TVR18, TVR35) felt their memory of the learning content was enhanced while using HiVR. For example, TVR20 explained:

> *As a result of the situated learning in immersive VR, we are able to recall the learning content more easily, as we are able to connect the English expressions with the scenario more easily. We would benefit from this improvement in our ability to use English in the real world* (TVR20).

TVR31 also thought that learning efficiency was higher in HiVR. She believed the higher learning efficiency would lead to a better learning outcome. This links with high engagement in learning while using HiVR, as discussed in Section 3.3.1.2. Many students (e.g. SVR5, SVR15, SVR20, SVR30, TVR9, TVR6, TVR17) noted that enhanced engagement was an advantage of using HiVR for oral English learning. Some students (SVR6, SVR24 and TVR6) also stated: *I would focus on the task*. However, some students held the opposite opinion about the effectiveness of HiVR on learning a foreign language, and this will be discussed in Section 6.3.5.2.

Independent learning

Some students believed that HiVR had great potential for independent learning, in which students can have control and ownership in learning. This can be linked with the sense of agency as a main affordance of HiVR in learning (see Section 3.2.2), which allows the freedom to control the learning process. SVR32 reflected on this point in detail, stating: *VR courses are easily accessible because we can have a conversation with any partner as long as we have the Internet.* Similarly, TVR24 mentioned:

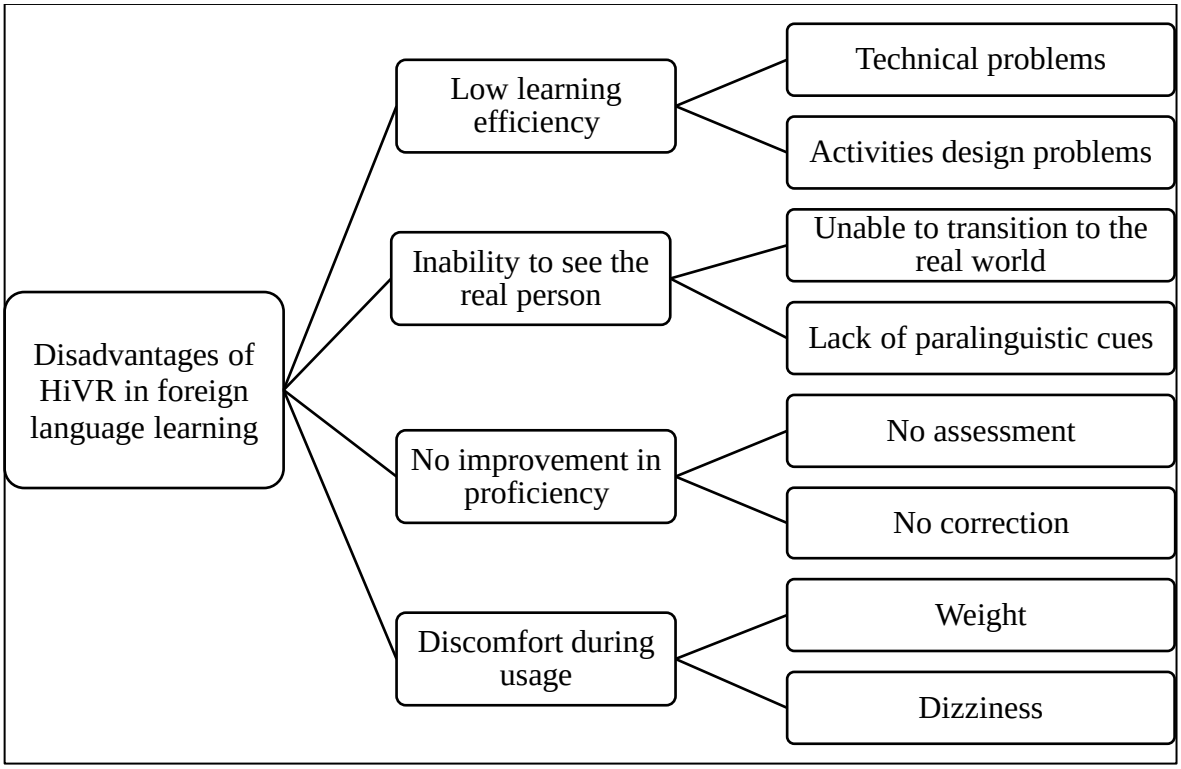
> *Immersive VR is convenient for self-learning of a foreign language. It is difficult to find partners in the real world whenever I want to practise oral English, but it can be easily achieved in VR* (TVR24).

TVR12 felt that the autonomy or sense of agency in learning was highlighted by the fact that *We are in control of what we learn and what we do in the VR course*.

A few students though that HiVR lessons had the advantage of repeating the learning content, which also helped with their independent learning. For example, TVR22 and TVR 20 both responded: *I can practise the learning activities repeatedly using VR*. In contrast, some students found the HiVR course to be less beneficial than the traditional classroom setting. Details can be found in the following section.

According to the students in my study, HiVR has four disadvantages pertaining to oral English learning (see Figure 6-14): low learning efficiency, inability to see the real person, no improvement in proficiency, and discomfort during usage.

Figure 6- 13 *Disadvantages of HiVR in oral English learning based on my data analysis*



Low learning efficiency

Some students (N = 25) reported low learning efficiency as a disadvantage of HiVR-based learning sessions, and this was due primarily to technical problems they encountered. This is similar to findings in previous research (see Section 3.3.2.2). The Internet connection was sometimes not stable (SVRS26), the manipulation of the device was complex (SVR14, TVR21), and the programme sometimes crashed without any foreseen reason (SVR31, SVR7, TVR8). In terms of the functions of the VR learning content, it was difficult for students to return to previous steps of the learning activity (SVR3). In my fieldwork, all technical problems prolonged the duration of the course, which seemed to influence students' learning efficiency as reported by the participants.

SVR29's critique of the learning content noted that the vocabulary and sample dialogues given in the HiVR environment limited the vocabulary and expressions used in her utterance. She expected less structured activities, in which she could have more freedom to produce her own utterance. Some negative perceptions were raised about the design of the activities, which was held to be not flexible (SVR12, TVR8), not explicit (TVR9), and not as rich as in a traditional class (TVR7, TVR21). By not flexible and not explicit, participants meant

that the device and the learning activity design were not smart enough to be used without assistance from the researcher (me). Specifically, SVR12 had this complaint about the learning content in HiVR:

> *The links designed between activities are rigid, sometimes meaningless, but increase the time consumption of learning in VR. Some of the activities require collaboration with the partner, but the device is not smart enough for collaboration between two users* (SVR12).

Therefore, it seems that improvement is needed in both hardware and software of HiVR in order to enhance learning efficiency.

Inability to see the real person

Being unable to see the real person was seen as an advantage of HiVR in alleviating FLSA by some students (see Section 6.3.5.1), but also as a disadvantage. This is consistent with what is summarised in Section 6.3.4.2 that being unable to see the real person makes it difficult to transition the low anxiety in HiVR to real-world situations. Similar points were found in students' responses to this question. For example, *there were no possibilities for eye contact, facial expressions, and body gestures in the HiVR learning sessions due to technical limitations* (SVR17). The same point was made by SVR9 and SVR21, who suggested adding facial expressions and body behaviours to the HiVR lessons.

Some students (N = 11) felt communication in HiVR with an avatar was different from real-world communication. This resulted in the difficulty of transition the low anxiety in using HiVR for communication to real-life conversation. SVR25 specified: *I don't feel embarrassed and anxious while speaking using VR, but when I speak English in the real world without VR, the panic feelings appear again.* TVR33 expressed: *We will not have task guidance in the real world as we do while wearing a VR headset, so it doesn't help with my anxiety.* TVR32 reported, similarly: *I would feel differently when I was facing a real person, even though my level of anxiety is reduced while taking a virtual reality course*.

Thus, the difference between the virtual presentation of users (avatar) and a real person seemed to be an obstacle for some learners' seeking to export the low anxiety experienced in HiVR into real-world situations. Lack of paralinguistic cues (e.g. facial expressions, body behaviours) generated difficulties for some students in the communication using HiVR.

No improvement in proficiency

Proficiency is the most important aspect that Chinese EFL learners take into account in foreign language learning. Although many students perceived an improvement in English proficiency after the intervention (see Section 6.3.5.1), a few students did not. According to their responses, no assessment (SVR34) and no correction (TVR11, TVR5, TVR8) in terms of linguistic knowledge were seen as disadvantages of HiVR-based learning sessions. Corrections were expected in pronunciation (TVR5, TVR8) and grammar (TVR11). Some students (SVR3, TVR21) also thought the learning content was too limited to improve their speaking skills.

Although no assessment or correction of produced language was intentionally designed for the intervention in order to avoid providing with negative experience to students of oral English learning, it may be modified in future learning designs. Improving students' oral proficiency along with offering evaluation or correction to their languages may be taken into consideration in the learning design. Based on my research results, proficiency improvement appears to be important for alleviating Chinese EFL students' FLSA.

Discomfort during usage

A number of problems were encountered during the intervention due to the head-mounted display. Many students (e.g. SVR2, SVR12, SVR33, TVR18, TVR27) complained about its weight. They reported that the headsets were too heavy to be worn for long periods of time, although taking off the head-mounted display for a break after 10 minutes of use was utilised in my fieldwork (see Section 5.3.3). The headsets also cause dizziness, as TVR22 and TVR23 pointed out.

Overall, students' experiences with HiVR for oral English learning were mainly positive, but negative perceptions were also noted. This answers RQ5. According to participants' responses to Q5 in the open-ended questionnaire, HiVR brings authenticity and enjoyment to foreign language speaking and improves the learning effectiveness and independence. However, due to technical and activity design problems, some participants did not find the HiVR learning sessions flexible and efficient. The lack of paralinguistic cues was also seen as a disadvantage. No assessment or corrective feedback in the course seemed to be a problem for some students, who expected improvement in their oral proficiency after the

intervention. Last but not least, some students complained about discomfort while using the head-mounted display, such as its heavy weight and the dizziness it caused.

6.3.6 Summary of Qualitative Results

Qualitative data was collected in order to answer RQ3, RQ4 and RQ5. The integration of quantitative and qualitative data for answering RQ3 can be found in Section 6.3.3, so it is not repeated here.

With regard to RQ4, most students held positive perceptions of HiVR and FLSA. They believed that HiVR might be beneficial for FLSA alleviation in several ways, such as 1) creating an authentic environment, 2) raising concentration on the task, 3) improving perceived oral proficiency, 4) providing more practice, 5) being interesting and relaxing, and 6) being unable to see the real person (see Section 6.3.4.1). The negative perceptions of HiVR and FLSA were based on the short duration of the intervention and the lack of paralinguistic cues (e.g. facial expressions, body behaviours) (see Section 6.3.4.2).

According to students' experiences of using HiVR for oral English learning (RQ5), the advantages of HiVR technology were primarily found from: it adds authenticity to the conversation, and most of the students found it enjoyable. In addition, some students stated that HiVR enabled them to learn more effectively and independently. However, some students noted disadvantages of HiVR in English speaking, such as low learning efficiency as a result of technical problems and activity design issues, a lack of paralinguistic cues and difficulty transitioning to real-world communication, as well as difficulty improving proficiency. Several students reported difficulty with the weight of the head-mounted display, which caused discomfort during the intervention. A few students even felt dizzy while using the head-mounted display.

6.4 Summary

This chapter presented the results for both quantitative and qualitative databases as well as the integration of the two datasets. According to the statistical analysis, HiVR did not affect general FLSA levels significantly after a nine-session intervention. Although students' self-rating of specific FLSA levels were lower in the second speaking test than in the first, the main reason for this, as reported by the participants, was that they were more familiar with

the test after experiencing it once and had gained similar task practice in the intervention. Thus the reported decreases in the self-rating of FLSA levels in the speaking tests might not apply to other types of speaking tests or tasks. It might therefore be concluded that the interventions based on HiVR did not result in significant FLSA changes, which answers RQ1.

In relation to RQ2, the four groups did not show significant differences in their general and specific FLSA levels. However, in terms of self-ratings of FLSA in the speaking tests, the SC group seemed to benefit most from the intervention and the TVR group benefited the least. The difference in FLSA change among four groups may also be attributed to the fact that the type of task (role-play) used in the speaking test was the same as those used by groups utilising situational learning approach (SVR and SC groups). Hence it cannot be concluded that situated learning theory is more effective than traditional teacher-centred teaching in alleviating Chinese EFL learners' FLSA, but more practice of the same task type may make a difference. As a result of a nine-session intervention, HiVR did not prove to be effective in decreasing FLSA levels among Chinese EFL learners.

With regard to RQ3, students' responses in the open-ended questionnaire further supported the quantitative results of RQ1 and RQ2. The main reason for perceiving no change in general FLSA was that participants perceived no improvement in oral proficiency during the intervention. Additionally, the short duration of the intervention was also put forward as a reason why no changes in FLSA were perceived. A majority of students suggested that test familiarity was a key reason for the decrease in self-ratings of FLSA levels in the speaking tests. Perceived proficiency improvement, increase in experience, and self-confidence development also explained the decrease in the self-rating of FLSA levels in the speaking tests.

In terms of students' specific perceptions of HiVR and FLSA (RQ4), their positive views could be categorised as: 1) creating an authentic environment, 2) raising concentration on the task, 3) improving perceived oral proficiency, 4) providing more practice, 5) being interesting and relaxing, 6) being unable to see the real person. A few students stated that the duration of the intervention was short and there were no paralinguistic cues, which resulted in their negative perceptions of the impact of HiVR on FLSA. Therefore, the answer

to RQ4 is that students' perceptions of HiVR and FLSA were mainly positive, although negative positions were also taken.

For RQ5, based on students' experience, the advantages of utilising HiVR for oral English learning included bringing authenticity to the conversation and seeming to be enjoyable to the students. Students also found the HiVR learning sessions efficient and independent. HiVR for oral English learning was, however, found to have a number of disadvantages in my study. These include low learning efficiency, absence of paralinguistic cues, difficulty in improving proficiency, and discomfort while using the device. The discussion of all these results and how they link with previous literature can be found in the following chapter.

Chapter 7. DISCUSSION AND CONCLUSION

This mixed-method study presents initial research into the impact of High-immersion Virtual Reality (HiVR) on Foreign Language Speaking Anxiety (FLSA) in light of two learning environments (HiVR and classroom) and two learning approaches (situated learning and teacher-centred learning). Firstly, this chapter extends the results presented in Chapter 6 by recapitulating the main findings using existing literature and addressing the research questions. The findings on HiVR and general FLSA are discussed, followed by those on HiVR and specific FLSA in the speaking test. The advantages and disadvantages of using HiVR for oral English learning are then elaborated. Next, the study's theoretical and pedagogical implications are illuminated, as well as its limitations. Following the conclusion of this study, suggestions for future research are provided at the end of the chapter.

The research questions (RQs) investigated in my study are iterated below for understanding this chapter easily.

> **RQ 1:** To what extent (if any) do the FLSA levels of students who used HiVR change after the intervention?
> **RQ 2:** To what extent (if any) do two learning environments (HiVR and classroom) combined with two learning approaches (situated learning and teacher-centred learning) have different degrees of influence on students' FLSA levels?
> **RQ 3:** How do students' perceptions of the difference in their FLSA levels before and after the intervention support, explain or refute the quantitative results?
> **RQ 4:** What are students' perceptions of the impact of HiVR on FLSA?
> **RQ 5:** What are students' experiences of using HiVR in oral English learning?

7.1 HiVR and General FLSA

General FLSA, in my study, refers to learners' anxiety about speaking the foreign language in general situations. Quantitatively, it was measured by the ESAS (English Speaking Anxiety Scale, see Section 5.3.4) and compared among and within four groups in the experiment. Each of these four groups experienced a different combination of learning environment and learning approach. The group categories were as follows:

> SVR group: Situated learning and HiVR

TVR group: Teacher-centred learning and HiVR

SC group: Situated learning and classroom

TC group: Teacher-centred learning and classroom

Qualitative data was gained from open-ended questionnaires after the experiment. The findings of both quantitative and qualitative data sets are illustrated below.

7.1.1 General FLSA Difference (RQ 1, RQ 2, and RQ 3)

According to the statistics of the ESAS scores (see Section 6.2.1) and quantitative content analysis of the open-ended questionnaires (see Section 6.3.1), no evidence was found of significant change in students' general FLSA before and after the intervention (RQ 1 and RQ 3). Linking back to the theoretical framework and the hypotheses (see Section 4.3), no conclusion can be drawn on the positive or negative impact of VR exposure-based therapy and situated learning on general FLSA. Based on the qualitative content analysis of the open-ended questionnaire, the reasons for no change in general FLSA were mainly that students perceived no improvement in their oral proficiency, no change in their fear of failure and the shortness of intervention duration (see Section 6.3.1). This addresses RQ 3.

The findings in my research contradict those of some previous studies on HiVR and FLSA (Thrasher, 2022; York *et al.*, 2021), which found a decrease in FLSA after using HiVR. However, there were some problems with the measurement of FLSA and the research design in these studies (Thrasher, 2022; York *et al.*, 2021), which might have led to invalid or unreliable results. The corresponding problems to each study can be found in Section 3.4.2. For example, the statements in the Likert scale used to measure FLSA sometimes were opposite to each other, but the mean of its scores was used to evaluate FLSA levels in their studies. Further investigation of accurate and unified measurement of FLSA is suggested.

My finding of no statistically significant differences in general FLSA levels among four groups after the intervention (RQ 2) is consistent with that of Chen, Hung, and Yeh (2021). They revealed no statistically significant differences ($p = .18$) between HiVR and non-HiVR groups on their attitudes towards English anxiety after one problem-based learning session. It should be noted that only a single-session intervention was conducted in their study, whereas nine sessions were implemented in my study over a three-month period. It seems the duration of the intervention did not influence the research results. However, different

durations of intervention may be worth considering in future studies. This is because some of my participants gave the short duration of the intervention as a reason for no change in their FLSA levels (see Section 6.3.1.2), but a long duration of HiVR learning may cause fatigue and reduce language learning effects (Qiu *et al.*, 2023)

Although a lower level of general FLSA was found in HiVR group compared to the classroom group in the study of Thrasher (2022) (see Section 3.4.2), she interpreted her results as FLSA levels decreased as time went by, and the decrease could be attributed to the increase practice in speaking rather than HiVR technology *per se*. This is similar with my finding of reasons for specific FLSA change discussed in Section 7.2.

To sum up, my finding of no statistically significant differences of general FLSA between HiVR and non-HiVR groups is inconsistent with most previous research. This leaves space for future research, such as to involve accurate and unified measurement of FLSA and to take durations of intervention into account (see Section 7.7 for future research suggestions).

7.1.2 Perceptions of HiVR and FLSA (RQ 4)

Based on my data, 81.16% of participants (N = 56) in the HiVR groups (SVR and TVR) believed that HiVR technology had a positive effect on alleviating their FLSA, which addresses RQ 4. However, this does not align with the quantitative results that there were no significant differences in their general FLSA after using HiVR. This may be due to the fact that students' perceptions of HiVR and FLSA are reflected on their FLSA while using HiVR to communicate, the FLSA levels measured by the ESAS are about their FLSA while using the foreign language in real-world situations. That is to say, HiVR may be perceived as a useful tool to alleviate students' FLSA while using it for speaking practice, but it seems not to have an influence on general FLSA in real-life settings. Nevertheless, my participants' specific perceptions of HiVR and FLSA are consistent with those of previous studies (see the following for details). A future research focus might be how to export a low-anxiety experience in HiVR to a real-world use of the target language (see Section 7.7).

In line with the findings in the studies of Kaplan-Rakowski and Gruber (2021) and Gruber and Kaplan-Rakowski (2020), students in my study reported that HiVR provided an authentic environment for foreign language interaction. This prepared them for real-life conversations, resulting in a potential reduction in their FLSA levels. When using HiVR, my

participants reported that they were able to concentrate more on the task. Several previous studies (Alfadil, 2020; Enkin, 2022; Tseng & Geng, 2021) also revealed a high level of engagement when using HiVR for learning. It is noted by York *et al.* (2021) that the unfamiliarity of the virtual environment in HiVR increased students' interest in exploring the virtual environment actively and drove them to seek additional content for speech. Participants in my study reported similarly that while using HiVR they were busy observing the virtual environment, attempting to complete the tasks through communication, and searching for appropriate expressions, all of which helped them overcome their nervousness about speaking English.

Perceived oral proficiency seemed to improve after using HiVR according to my participants' responses, and this improvement contributed to the alleviation of their FLSA. Perceived proficiency is not an objective measure of competence but is based on students' own perceptions of their abilities. This is consistent with the findings by Soto *et al.* (2020) and Liaw (2019) that the majority of participants believed HiVR interventions improved their speaking skills. In particular, students in my study believed they had acquired a large amount of vocabulary and expressions, learnt a variety of sentence structures, and improved their speaking skills after the intervention with HiVR. However, the effects of HiVR on objective oral proficiency have been inconsistent in previous studies (see Section 3.3.1). As no objective assessments of oral proficiency were involved in my study, future studies could include both subjective and objective evaluations of oral proficiency in order to examine the impact of HiVR on them.

Aside from this, some students in my study thought the increased practice in oral English while using HiVR helped to alleviate their FLSA. As discussed in Chapter 1 and Chapter 2, English learners in an EFL context (e.g. China) often find it hard to access target-language environments in the real world, and the low frequency of using the language is the main cause of their FLSA. In my study, through using HiVR, Chinese EFL learners have more opportunities to communicate authentically in an English-speaking environment. As a result of the simulation of various scenarios in HiVR, the frequency of speaking a foreign language in a target-language environment increases. Based on the responses from my participants, the increase in experience of speaking English seems to help reduce FLSA levels. This is similar to the finding by Thrasher (2022) that the increase in practice rather than the HiVR technology itself resulted in the decrease in FLSA levels.

Moreover, students in my study found the experience of communicating in HiVR interesting and relaxing. Mirroring the experiences of using HiVR for foreign language learning reported by Enkin (2022) and Xie, Ryder, and Chen (2019), students in my study enjoyed speaking in HiVR, and they were no longer afraid of making mistakes. Similarly, a number of previous studies (Awada & Diab, 2018; Chen *et al.*, 2020; Sally Wu & Alan Hung, 2022) found that participants were more interested in learning English using HiVR. By providing a relaxing and interesting environment, HiVR encouraged students to take more risks and boost their confidence (Ebadi & Ebadijalal, 2020). My research findings support the notion that HiVR provides a low-anxiety environment.

Different from the findings by York *et al.* (2021), many participants in my study stated the benefit of anonymity (being unable to see the real person) in alleviating their anxiety about speaking the foreign language. This is consistent with certain previous studies, which found that students felt less anxious when using HiVR because they were shielded from identifiability by avatars (Rosell-Aguilar, 2005), and their paralinguistics (frowning, raised eyebrows) and social cues (age, gender, race) cannot be observed (Arnold, 2007). Although participants were identifiable in my fieldwork as they were gathered in the same room, students (SVR21, SVR35, TVR20, TVR25) were still benefiting from the use of avatars as they were unable to see the facial expressions of their interlocutors in alleviating their FLSA while using HiVR to practice oral English.

Nevertheless, it was reported that some students could not determine the intentions of their interlocutors without reading facial expressions and body language through the avatars. This is consistent with the lack of paralinguistic cues (e.g. lip reading, facial expressions, and body language) in HiVR, which was identified as a disadvantage of HiVR in foreign language learning by Kaplan-Rakowski and Gruber (2021). It should be noted that the most recent HiVR technology enables eye and facial tracking as well as synchronising body movements, and this could be taken into account in future learning design using HiVR. A future investigation could compare the impact of HiVR on FLSA of using or not using paralinguistic cue tracking.

A further point about the impact of HiVR in alleviating FLSA made by my participants was that it was difficult for them to export the feeling of low anxiety while using HiVR to real-life situations. When they return to the real world, foreign language learners will encounter

real people in circumstances where there are no clues and guidance for conversation as in the virtual environment. This raises their anxiety about speaking the foreign language again. This point may explain why no change in general FLSA were found, but the perceptions of HiVR and FLSA were mainly positive in my study. Students felt less anxious while using HiVR, but this did not influence their actual FLSA levels in real-life situations. This finding is different from claims by Opriş *et al.* (2012) and Kampmann *et al.* (2016) that VR exposure-based therapy has real-life impact that can influence behaviour in the real world. Thus, a further comparison between exposure to the virtual environment in HiVR and real life is suggested in order to identify the difference in their impact on FLSA.

In conclusion, students' positive perceptions of the impact of HiVR on FLSA are mainly consistent with previous research findings. According to my data, the learning environment provided by HiVR is perceived to be relaxing and interesting. This allows students to focus on the learning tasks, increase their experience in such conversations, and improve their perceived language proficiency. In addition, the anonymity of HiVR (e.g. not being able to see the actual person and not being identified by others) contributed to the alleviation of FLSA. It is important to note that the positive perceptions of HiVR and FLSA do not contradict the quantitative results of no significant change in general FLSA after the intervention. The reason for this is that students' perceptions of HiVR and FLSA were mainly based on their experiences of FLSA in the HiVR environment, whereas the FLSA measurement focuses on the FLSA in real-world communications. In other words, students in my study perceived HiVR to be useful for alleviating FLSA while using it, but interventions in HiVR seemed not to have impact on students' FLSA levels in real life. This adds to the current understanding of HiVR and FLSA and leads to several future research directions (see Section 7.7).

7.2 HiVR and FLSA in Specific Tasks

7.2.1 Specific FLSA Difference (RQ 1, RQ 2 and RQ 3)

In my study, FLSA in the speaking tests as measured by 8-item 10-point self-ratings while viewing the video recordings of students' own performance decreased significantly after the intervention in all groups according to the statistical results (SVR: $p = .002$; TVR: $p = .005$; SC: $p < .001$; TC: $p = .004$) (see Section 6.2.2). The quantitative results are supported by qualitative data: 78.57% of students (N = 110) reported feeling more relaxed or a bit more

relaxed in the second speaking test (see Section 6.3.2). This answers RQ 1 and indicates that HiVR seem to have positive influence on the alleviation of FLSA in specific tasks.

Moreover, none of the four groups (SVR, TVR, SC, and TC) showed statistically significant differences in the self-rating of FLSA in the second speaking test (see Section 6.2.2), which addresses RQ 2. This means that FLSA decreased in the speaking tests for all four groups, but the decrease did not differ significantly among groups. This is also supported by the qualitative data that students' descriptions of FLSA in the second speaking test did not differ significantly among the four groups (See Section 6.3.2). Therefore, it can be concluded that interventions with two learning environments (HiVR and classroom) and two learning approaches (situated learning and teacher-centred learning) resulted in a decrease in students' FLSA levels in the speaking tests, but neither the learning environment nor the learning approach led to significant differences in the decreased degree.

The results of students' self-rating of their FLSA levels in speaking tests are different from the results of the ESAS scores. This is because the ESAS and self-rating of FLSA assessed different types of FLSA. Using ESAS, general FLSA was measured without consideration of the type of speaking task, but self-rating of FLSA examined the specific English speaking anxiety in a two-way role-play conversation. This finding can also be interpreted as that it may be the increase in the practice of role-playing tasks resulting in the reduction of FLSA levels rather than the use of technology or the application of various learning theories. It seems that no previous studies have measured or compared FLSA in specific tasks as conducted in my study. For this reason, it may not be possible to compare current findings with existing literature. My pioneering use of the self-rating method to measure FLSA in a two-way conversation may be replicated in future studies. In addition, the reasons for the decrease in FLSA in the speaking test should be noted. They are illustrated in the following section.

7.2.2 Factors Contributing to Specific FLSA Decrease (RQ 3)

Based on students' responses, several factors were attributed to the decrease in their self-ratings of FLSA in the second speaking test, including test familiarity, perceived proficiency improvement, increase in experience, and self-confidence development (see Section 6.3.2). This further answers RQ 3. In accordance with the theoretical framework, I divided the reasons contributing to reducing specific FLSA in speaking tests into two categories: task

repetition and pleasant experience (see Figure 7-1). The following sections provide a detailed explanation to these two reasons. It is surprising to find that no students in the HiVR groups (SVR and TVR) reported reduced FLSA during the speaking test as a result of the HiVR technology *per se*.

Task repetition

According to my participants' responses, test familiarity, including being familiar with the test type, test process and interlocutor, helped to reduce their FLSA in the second speaking test. This can be linked to the fact that test anxiety[34] may decrease when people become more familiar with the testing environment (Reeve, Heggestad, & Lievens, 2009). This is because people acquire test-specific knowledge or increase their readiness, and this reduces the distractions related to unfamiliarity with the testing situation (Reeve, Heggestad, & Lievens, 2009). However, test anxiety is not covered in my study.

The primary reason for the decrease in FLSA in the speaking test was that the interventions provided students with more opportunities to experience similar tasks. This research finding is related to *task repetition* in the field of second language acquisition. *Task* can be defined as "a holistic activity, which engages language use in order to achieve some non-linguistic outcome while meeting a linguistic challenge, with the overall aim of promoting language learning, through process or product or both" (Samuda & Bygate, 2008, p. 69). Generally, task repetition refers to repeating the same task over and over again, but it can also be more specific in certain circumstances, for example, repeating the exact task, repeating the content and procedure of tasks, repeating the task procedure with different content, and repeating the content with different procedures (Kim *et al.*, 2018).

Research on task repetition in the field of second language acquisition has focused primarily on its effects on linguistic performance, such as grammatical structures (De Jong & Tillman, 2018), lexicon use (Kim *et al.*, 2018), and oral production (complexity, accuracy, and fluency) (Hu, 2018). A detailed explanation of how task repetition facilitates oral performance is provided in the study of Bygate (2013), which suggests that students may be able to speak more effectively on subsequent occasions as a result of repeated exposure to

[34] Test anxiety is "the set of cognitive, physiological, and behavioural responses that accompany concern about possible negative consequence or failure on exams or similar evaluative situations" (Zeidner, 2007, p. 166).

the same speaking processes on a particular topic. Little research (Amiryousefi, 2016) has been conducted on task repetition and FLSA, but a positive result was revealed that task repetition helps to reduce FLSA levels. This is consistent with students' comments in my study that encountering various interlocutors and topics in the intervention prepared them for similar speaking tasks in the future, which alleviate their FLSA.

In my study, many students (N = 48) perceived an improvement in their language proficiency after the intervention, in terms of expressions, speaking skills and pronunciation. A similar conclusion was reached in previous studies on task repetition and linguistic proficiency. Nitta and Baba (2018) argue that prior experience with a particular task type but with a different content may be stored in long-term memory, potentially improving language proficiency. According to Baker and MacIntyre (2000), the more opportunities students have to speak the target language, the better their ability to predict and confirm expectations of the conversation, leading to a greater sense of perceived competence or self-esteem. This may explain why the students in my study perceived a decrease in their specific FLSA after repeating the tasks.

In addition, my data indicated that the increase in experience of role-play tasks was the main benefit of the intervention in reducing students' FLSA in the speaking tests. The task type (role-play) in the intervention was the same as that used for the speaking test. This may explain why the self-rating of FLSA levels in the SVR group reduced more than in the TVR group, and those in the SC group reduced more than in the TC group. Speaking tasks in situated learning groups (SVR and SC) involved role-play tasks for the entire session, whereas the teacher-centred groups (TVR and TC) only used a role-play conversation at the end of each session. Nevertheless, it was noted that task repetition may result in students' higher expectation of their language performance. The high expectation may increase their FLSA levels during the speaking test. A question therefore arose of how many sessions of task repetition are appropriate for alleviating FLSA. Future research may further investigate the impact of task repetition on FLSA and compare the impact of different number of tasks on FLSA (see Section 7.7).

Pleasant experience

In spite of the fact that task repetition may be beneficial for improving language performance, foreign language learners' experience during the task is also important for the alleviation of

FLSA, according to my participant's responses. Some students stated that their confidence in speaking English had increased as a result of realising during the intervention that making small mistakes was acceptable. This may be linked to VR exposure-based therapy (see Section 4.1) involved in my learning design, which argues that anxiety levels fall after exposure to the stimulus without receiving negative responses. To put it another way, when students receive positive feedback instead of negative feedback while performing the tasks, their anxiety levels may decrease. In my study, students became less concerned about speaking the foreign language when they were not corrected harshly by the instructor as they usually were in Chinese classrooms (see Section 2.3.3 for factors for Chinese EFL learners' FLSA).

Based on the quantitative and qualitative results of my study, it is unlikely to conclude that HiVR technology itself is responsible for the reduction of specific FLSA, but providing a pleasant experience with positive feedback while involving students in speaking tasks appears to be a potential solution. This can also be linked with positive psychology, which boosts positive emotions in foreign language learning in order to reduce negative emotions. A previous experiment (Jin, Dewaele, & MacIntyre, 2021) indicated that reminiscing about language achievement could significantly reduce FLSA. This is supported by my data, which suggests that experiencing success in foreign language conversations seems to be beneficial for building self-confidence in foreign language speaking. The improvement in self-confidence then alleviates FLSA.

Overall, the in-depth investigation of responses in the open-ended questionnaires in my study demonstrated that multivariate factors were responsible for the decrease in self-rating of FLSA levels in the second speaking test rather than a single factor. This may be due to the fact that FLSA is a multifaceted concept. Based on my data analysis, task repetition and pleasant experience are the main factors contributing to the decrease of FLSA levels in the speaking tests. This expands the understanding of FLSA in specific tasks and proposes a possible way to alleviate it. With regard to the HiVR technology, VR exposure-based therapy seems to be promising for FLSA reduction. Involving positive emotions in traditional classrooms may also be a solution.

7.3 Advantages and Disadvantages of Using HiVR for Oral English Learning (RQ 5)

Students' perceptions of the advantages of HiVR in oral English learning are similar to their reflections on HiVR and FLSA: the authenticity, enjoyment, effective learning and independent learning (see Section 6.3.5.1). The disadvantages of HiVR in oral English learning are mainly to do with the low learning efficiency, being unable to see the real person, experiencing no improvement in proficiency and discomfort during usage (see Section 6.3.5.2). These findings address RQ 5. To avoid repetition with Section 7.1.2, this section excludes the discussion of the first two advantages: authenticity and enjoyment, and focuses instead on other insights into the advantages and disadvantages of using HiVR for oral English learning.

In terms of learning effectiveness, some students considered HiVR to be an effective platform for improving oral proficiency, while others did not. Students who regarded it positively felt they had learnt expressions which could be applied in real-life scenarios, and their memory of the content was enhanced while using HiVR. Participants who saw it negatively attributed the low effectiveness of learning in HiVR primarily to the fact that there was no correction or assessment of linguistic knowledge in the intervention. They felt that the learning content in the intervention was limited to improving their speaking abilities. These positive and negative perceptions of HiVR and oral proficiency can be attributed to Chinese students being often taught in a teacher-centred instructional method, in which they are spoon-fed linguistic knowledge and told that vocabulary and grammar learning is the most important (see Chapter 2). They are used to be corrected with any mistake they made in English language rather than being free to speak for communicative purpose (see Chapter 2). Therefore, providing immediate feedback in the learning design of HiVR can be considered, but the purpose of involving positive emotions should be borne in mind.

As far as learning efficiency is concerned, some students in my study viewed the increase in learning engagement as a benefit of HiVR for oral English learning. While using HiVR, the students were more focused on the learning activities, which improved their learning efficiency. By contrast, some students reported that their learning efficiency was low while using HiVR, primarily because of technical problems. The use of HiVR is greatly dependent on a stable Internet connection, but there were times when the connection was unstable. Consequently, the intervention would be interrupted, and the class would be prolonged in my field work. In addition, some participants had difficulty manipulating the HiVR device and sometimes found the learning activity guidance unclear. These responses were in line

with findings in previous studies that students did not know where to look in HiVR (Xie, Ryder, & Chen, 2019) and the network quality hampered the use of HiVR devices in foreign language learning (Urueta & Ogi, 2019). Additionally, in the study of York *et al.* (2021), participants mentioned that they had difficulty grasping visual saliency due to their unfamiliarity with the HiVR technology. Due to this, although HiVR may engage students more effectively in the learning task than traditional classrooms, these technical issues need to be addressed in order to produce a smooth learning experience.

In terms of autonomy of learning, many students believed that HiVR provided them with greater opportunities for independent learning. It is possible for them to repeat the tasks whenever they desire as long as they have the device to hand. When using HiVR in my study, students felt in control of the learning process, which is related to the second main affordance of HiVR: agency (see Section 3.2.2). Agency refers to the freedom of controlling the virtual environment (Petersen, Petkakis, & Makransky, 2022). Kontra, Goldin-Meadow, and Beilock (2012) noted that active manipulation of the learning content contributed to learning effectiveness. This is supported by my qualitative finding that students perceived learning in HiVR to be effective, and specific investigation of this aspect could be conducted in the future.

Through the use of a head-mounted display , users can experience the virtual environment from a first-person perspective, reducing distractions from the outside world (Freina & Ott, 2015). The use of head-mounted display further enhances the sense of presence for users, which is another affordance of HiVR in learning (see Section 3.2.2). Nevertheless, some students in my study reported negative responses to the use of head-mounted displays. A few students experienced discomfort after wearing the head-mounted display for a prolonged period of time, and it could cause dizziness. This is consistent with findings in the study of Gruber and Kaplan-Rakowski (2020) that participants were overheating while wearing the head-mounted display. The engineers in technology companies have been working for decades to reduce the weight and size of the head-mounted display and make them more comfortable to wear (see Section 3.1.2), but it appears that further technological development of head-mounted display is essential, especially for educational use.

The use of HiVR in oral English learning has both advantages and disadvantages. The majority of the students enjoyed the HiVR oral English learning programme, but some

negative points were raised regarding the hardware and the design of the learning activities. A collaborative effort between educators, engineers and software developers is required in order to make HiVR technology more suitable and useful for educational purposes.

7.4 Theoretical Implications

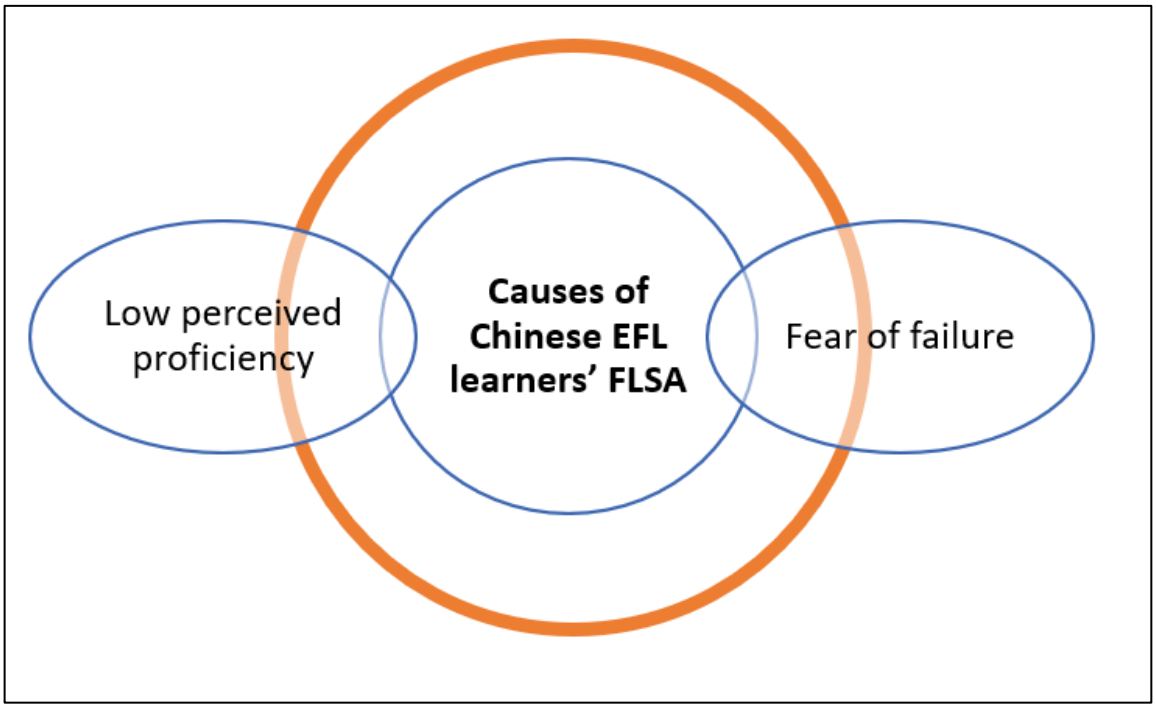
7.4.1 Causes of FLSA Among Chinese EFL *learners*

As Horwitz (2016) argues, it is crucial to understand the components of FLSA in different cultures and learning contexts. In order to better understand the factors related to Chinese EFL learners' FLSA, my study investigated these factors both quantitatively and qualitatively. The construct validity of the adapted Likert scale (English Speaking Anxiety Scale, ESAS) was examined through factor analysis (see Section 6.2.1.2). Although the confirmatory factory analysis of the ESAS did not reveal good model fit for the initial four-factor model (see this model in Section 2.3.3) and the exploratory factor analysis did not reveal any accepted value for factor extraction, the qualitative data from the open-ended questionnaire investigating students' general FLSA contributed to the understanding of factors related to Chinese EFL learners' FLSA. Vocabulary, pronunciation, fluency, and grammar were the biggest worries for students when speaking English with others. This finding is consistent with the findings of Yan and Horwitz (2008) that Chinese EFL learners are worried about their vocabulary range, grammatical knowledge and listening comprehension ability.

Based on participants' responses, I have categorised two main factors related to their FLSA, which are *low perceived proficiency* and *fear of failure* (see Section 6.3.1.1 for details). This is in line with the two factors extracted from the study of Park (2014) as causes of FLSA in Korean English learners. The factors extracted by Park (2014) were communication apprehension and understanding, and communication apprehension and confidence. Unlike findings in previous studies (Amorati & Venturin, 2021; Jiang & Dewaele, 2019; MacIntyre, 2017), the FLSA factors found in my data were mainly related to the learners themselves, which excluded factors in relation to instructors and peers. This may be because the questions I posed required students to describe their FLSA rather than reflect on the causes of their FLSA. Nonetheless, my categorisation of causes of FLSA fits the distinction of individual (internal) and social (situational or external) factors based on previous literature (Gregersen, MacIntyre, & Olsen, 2017; Jiang & Dewaele, 2019). My categorisation of *low*

perceived proficiency is in line with the category of *underestimation of one's own abilities* outlined in the paper of Amorati and Venturin (2021). Above all, the causes of Chinese EFL learners' FLSA are mainly related to but not limited to low perceived proficiency and fear of failure (see Figure 7-1).

Figure 7- 1 *Categorisation of causes of Chinese EFL learner's FLSA based on my data*



The fact that perceived competence and subjective proficiency are more important than objective achievement and measured proficiency was highlighted in early-stage studies (Bailey, 1983; Clément, Dörnyei, & Noels, 1994; Onwuegbuzie, Bailey, & Daley, 1999), because a lower level of perceived proficiency would result in foreign language anxiety. MacIntyre (2007) also argued that learners' self-perceptions of foreign language competence or proficiency can be seen as a predictor for foreign language anxiety. My findings support this point that foreign language learners are anxious because they perceive their language proficiency to be low.

Aside from low perceived proficiency, students are concerned about embarrassment or failure during the target-language conversation. This was categorised as *fear of failure* in my study (see Section 6.3.1.1). This is different from the component of *fear of negative evaluation* for foreign language anxiety defined in the work of Horwitz, Horwitz, and Cope (1986). *Failure*, based on my interpretation of the data, is not only associated with evaluative

situations but all situations involving speaking the target language. Fear of making mistakes, fear of not being understood by others, and fear of being unable to understand others were the main concerns of students in terms of failure. From a psychological perspective, students would fear being laughed at, losing face, and being embarrassed. These are in line with the finding by Ahmed (2016) that anxious students were concerned about failing in English speaking.

7.4.2 FLSA Measures

The measures of FLSA are under-researched in the field of second language acquisition (see Section 2.2). Phillips (1992) argued that the Likert scale used to collect foreign language anxiety levels did not measure anxiety related to specific oral tasks. However, it seems FLSA measurement has not been largely investigated from this perspective in the past three decades. In this thesis, I proposed a distinction between general FLSA and specific FLSA. This is based on my understanding of previous literature, which shows some foreign language learners are generally anxious about speaking the foreign language, but sometimes their anxiety levels may vary according to the situation. This is also based on the view that FLSA is not an invariable concept, but changeable and fluctuating (Gkonou, Daubney, & Dewaele, 2017).

My distinction between general FLSA and specific FLSA serves as a guidance for the measurement of FLSA involved in my study. Two retrospective methods: ESAS and self-rating of FLSA in specific tasks (see Section 5.3 for details), were adapted to examine general FLSA and specific FLSA. These two measurements of FLSA illuminated different results for the change in FLSA before and after the intervention (see Section 6.2). The differences in results for general and specific FLSA change supports the distinction between these two concepts, as proposed in Chapter 2. In other words, when measuring FLSA, it is worth asking whether what is being measured is anxiety in general foreign language speaking situations or anxiety in specific situations.

In addition, self-reporting may not be reliable for some measurements, because it depends on respondents' own evaluations of feelings which are subject to various sources of inaccuracy (Paulhus & Vazire, 2007). Thus, objective measurements should be considered when evaluating FLSA. However, there is a lack of research in this area. The salivary cortisol levels employed by Thrasher (2022) seem to be a possible objective measurement of FLSA,

as the level of salivary cortisol is revealed to be moderately correlated with FLSA. Nevertheless, further research is required to investigate whether salivary cortisol levels can be used to evaluate FLSA levels or not. Furthermore, nonverbal behaviours such as facial expressions seem to offer a potential route for measuring FLSA objectively. However, as discussed in Section 2.2.2, the action units should be identified rather than using separate facial expressions as an indicator for FLSA. Future research on this is also essential.

The ESAS adapted for my study gives an insight into Chinese EFL learners' FLSA and takes the learning context and culture into account, but further exploration of ESAS is essential. Other than the self-ratings method I adapted, future investigation of measuring FLSA in specific tasks is recommended (see Section 7.7). Furthermore, retrospective approaches to measuring FLSA are widely used, but objective measurements should be explored due to their potential for higher accuracy.

7.4.3 Using HiVR in Foreign Language Learning

My study was conducted as a preliminary experiment in using HiVR as a foreign language learning platform underpinning different learning approaches for alleviating foreign language learners' anxiety. Based on my results, HiVR should not just be seen as an instructional technology and medium for learning; it is important to make the most of its main affordances to enhance learning. The sense of presence and agency while using HiVR are suggested to be maximised within the learning design. The simulation of real-life scenarios in HiVR is recommended to be as authentic as possible in order to provide a better experience of *being there* to users. Moreover, a highly immersive virtual environment requires more advanced hardware (e.g. the headset and the controllers) to reduce the distractions of the outside world and increase the sense of presence. The continuous development of hardware is required.

In order to achieve learner agency while using HiVR, interactivity with the environment should be freer, and the activity design is better to be semi-structured with freedom for communication. Utilising HiVR with traditional learning methods, such as memorising words and sentences as Sally Wu and Alan Hung (2022) did in their study, did not offer agency to the users. Unsurprisingly, the results of Sally Wu and Alan Hung (2022) showed that a 10-week intervention of using HiVR in conjunction with traditional teaching methods did not a make difference to students' willingness to communicate. However, the semi-

structured activities applied in my learning design, which require students to participate in the activity but do not limit what they say and do, seem to provide more agency for students. From students' responses, it is clear that they enjoyed the learning activities in HiVR offered by my study. My participants also found the learning activities relaxing. This may be due to the appropriate difficulty of the learning task as well as the situated learning approach underlying its design. Hence, various learning approaches are recommended to be considered when designing learning activities for foreign language learning in HiVR. The selection of learning approach can be determined by the purpose of the learning session.

In addition, the results of my study suggest that reading facial expressions and body behaviours from avatars in HiVR is crucial. This is because enabling reading lips and facial expressions may enhance oral communication (Kaplan-Rakowski & Gruber, 2021). This need is also reflected by my participants, who found that lack of nonverbal cues made it difficult for them to ascertain the intention of their interlocutor. The technology of capturing facial expressions and body behaviours was underdeveloped when I undertook my learning design in 2019. However, avatars in HiVR can now display body behaviours and some facial expressions (e.g. eye movement, lips movement) in real time. That is to say, the movements the user makes can be reflected by their avatars in HiVR. Future studies could be conducted that take facial expressions and body movement into consideration in the learning design within HiVR.

In spite of the fact that HiVR is often considered an effective foreign language learning platform, my findings show that improving proficiency, especially expression acquisition, is necessary to meet the needs of students. Corrections and assessments may be involved in the learning process, but they should not be harsh and direct. The activity flow should be as smooth as possible in order to improve learning efficiency, although some technical problems are difficult to avoid. The learning session should not be lengthy while using HiVR, because the long-time wearing of the head-mounted display may cause discomfort. Based on my observations during the fieldwork and on students' responses, 10 minutes is the maximum recommended time for wearing a head-mounted display in an HiVR-based lesson.

Above all, based on my research findings, in order to make the most of HiVR for foreign language learning, the learning design should provide a sense of presence and agency. Semi-structured activities combined with cooperative learning approaches are recommended. The

learning content should be enjoyable and relaxing, for which the level of task difficulty should be considered. The challenge of using HiVR for foreign language learning is to capture and present nonverbal behaviours in real time. The length of learning sessions should also be taken into account because of the discomfort of wearing the head-mounted display for a long time.

7.5 Pedagogical Implications

As my research results showed that HiVR did not make a statistically significant difference in Chinese EFL learners' FLSA, the pedagogical implications discussed here concern general foreign language instruction, whether or not technology (e.g. HiVR) is used. In order to help students reduce their FLSA levels, their oral English proficiency needs to be improved. Chinese EFL teachers are advised to change their attitudes towards students' mistakes and invoke students' positive emotions in learning.

7.5.1 Improving Students' Oral English Proficiency

Based on my research findings, oral English proficiency is a main concern of Chinese EFL learners. Although China's Ministry of Education emphasised communicative competence in its College English Curriculum Requirements, oral English is still undervalued by EFL practitioners (Han & Yin, 2016). In order to reduce anxiety about speaking English, it is important to improve learners' oral English proficiency.

In Chinese EFL classrooms, the Grammar-Translation Method and Audiolingual Method are still dominant in actual practices (Rui & Chew, 2013). Students taught by these traditional approaches develop a good knowledge of vocabulary and grammar but are incompetent for communication (Han & Yin, 2016). Even if communicative language teaching was introduced to China in the early 1990s, the implementation of this task-based student-centred teaching method has encountered many challenges in Chinese EFL classrooms (Wei, Lin & Litton, 2018). This is mainly because some Chinese EFL instructors are not competent in English speaking (Han & Yin, 2016). Involving technology (e.g. HiVR) may be helpful for implementing communicative language teaching because the designed activities can provide oral communication tasks (Chun, Kern & Smith, 2016). Chun, Kern and Smith (2016) also argued that using technologies may help students to develop particular communicative competence as designed in the learning tasks, which is potential to improve students' oral

proficiency. The improvement of oral English proficiency is seen as an important element in alleviating FLSA, according to the findings of my study.

7.5.2 Changing Instructors' Attitudes to Students' Mistakes

In addition to improving actual oral English proficiency, it is also important to enhance students' perceived language proficiency. Phillips (1999) recommended that teachers should remind students that making mistakes is a natural part of foreign language learning process. Most Chinese EFL instructors, even now, are unaware of emotional or psychological factors in foreign language learning (Cheng & Chen, 2021). Therefore, as Jiang and Dewaele (2019) and Sun (2022) suggested, Chinese EFL instructors should learn about foreign language anxiety and pay attention to students' mental states during foreign language learning.

Chinese EFL teachers have a low tolerance of errors in linguistic aspects and high expectation of accuracy in grammar (Liu, 2007). Chinese EFL instructors correct students' mistakes directly and harshly in class. However, research (Gregersen, 2003) showed that correcting every error and overcorrecting may demotivate students. My study excluded any harsh correction or direct negative evaluation in the intervention; participant comments revealed that this seemed to relax them in the learning process and reduce their FLSA.

Changing teachers' attitudes to students' mistakes may be a starting point for alleviating FLSA in Chinese EFL classrooms. Chinese EFL learners need to be told that making some small mistakes is acceptable in the language learning process. Teachers should encourage students to produce more utterances rather than being afraid of making mistakes and avoiding speaking. This is in line with discussion in Section 2.5.1 that a non-harsh, friendly classroom atmosphere and appropriate techniques to correct students' errors may help to alleviate students' FLSA (Young, 1990; Gregersen, 2003). Even though HiVR is not available in each classroom, a low-anxiety environment can be created through EFL instructors' acceptance of students' mistakes. In line with the suggestions made by Mon (2019) and He (2018), Chinese EFL teachers are expected to be humorous, knowledgeable, patient and encouraging; this would help prevent students from worrying about making mistakes and thus alleviate their FLSA.

7.5.3 Invoking Students' Positive Emotions in Learning

According to my findings, students feel less anxious while using HiVR because it does not provide any negative feedback about their language production. The feeling of lessened anxiety is attributed to the utilisation of VR exposure-based therapy in the learning design for my study. Hence, Chinese EFL teachers should understand the importance of invoking positive emotions rather than negative emotions in foreign language learning, whether or not using technology. Nevertheless, Chinese EFL classrooms are predominantly teacher-centred (Hu, 2002). Teachers in China are often regarded as the authority that students must respect and obey (Jin & Cortazzi, 2006). However, the pedagogical methods, class procedures, and activities need to be adjusted to provide a comfortable environment for students to practise the target language (Li, 2016). The involvement of positive psychology (see Section 2.5.2.2) in foreign language teaching may have the potential to alleviate FLSA, but it needs further exploration in the Chinese EFL learning context (Cheng & Chen, 2021).

To sum up, in order to help alleviate Chinese EFL learner's FLSA, it is recommended to improve students' oral proficiency, change teachers' attitudes to students' mistakes and invoke students' positive emotions in class.

7.6 Limitations

The limitations of my study can be discussed from five aspects: generalisability of findings, bias of self-reporting approach, samples, technical limitations, and duration of intervention.

First, the research is focused on Chinese EFL learners at one university, but China is a big country with 23 provinces and five autonomous regions, four municipalities and two special administrative regions. English teaching is slightly different in schools and universities in different regions in China. For example, big cities, such as the four municipalities, may value oral English more than some underdeveloped provinces. Therefore, my current research findings may not apply to all Chinese EFL learners or a non-Chinese population; these different contexts would require further research. The current research design should be replicated in different contexts. Nevertheless, the main purpose of my study is to investigate whether using HiVR makes a difference to FLSA levels. As long as the measurement of FLSA before and after the intervention is the same, the research findings can provide reliable insights into the impact of HiVR on FLSA.

Second, the instruments (ESAS, self-rating of FLSA levels in specific tasks, and open-ended questionnaire) employed for my data collection were subjective self-reports. The accuracy of these in measuring FLSA may be doubted. With regard to the ESAS, it can be questioned that students may not be able to assess their FLSA accurately before the intervention due to the lack of oral English use in and outside the classroom. Moreover, with the Likert scale, participants may avoid extreme answers, which will produce inexact results (Bertram, 2007). In addition, in the open-ended questionnaire, participants may know what they are expected to respond and choose to please the researcher (Reja *et al.*, 2003). Therefore, objective measurement of FLSA is recommended to be involved in a future study comparing the difference in FLSA levels before and after the intervention. Moreover, a triangulation of methods (e.g. questionnaires, interviews, and observations) is suggested to provide different lenses for the same research problem.

The third limitation is associated with the samples, including the sampling technique and gender imbalance. Due to the need to implement the intervention in my study, a volunteer sampling was employed for individuals to decide whether to take part in the research or not. This was mainly based on ethical considerations but would result in some biases. Anxious EFL learners with FLSA are more likely to participate in the study than those who are not anxious. This may limit the representative nature of the population being studied. Similarly, the gender imbalance of the participants (132 females, 8 males) may result in the sample not being representative, even though Jiang and Dewaele (2019) claimed that gender may not influence FLSA levels. I therefore suggest replicating my study using a random sampling with a balance of gender.

The fourth limitation is the underdevelopment of the HiVR technology. When I started my PhD in 2018, there had been recent developments in HiVR hardware, such as smaller headsets and lower cost; but aspects of the software, such as interactivity with the virtual environment and capturing and real-time display of facial expressions, needed continuous development. Exploration with HiVR in foreign language learning can be interesting involving more technological functionality in the learning design and flexible interactivity with the virtual environment.

The final limitation is the duration of the intervention. The original 12-session intervention plan was cut down to nine sessions because of the Covid-19 pandemic. Some students in my

study mentioned the short duration of the intervention as a reason for no change in their FLSA. Hence, in future studies, it would be essential to expand the duration length and add more sessions in order to investigate the impact of HiVR on FLSA.

7.7 Conclusion

In conclusion, this mixed-method study of 140 Chinese EFL learners provides an understanding of the impact of HiVR on FLSA. It is a pioneering study, taking both learning environments (HiVR and classroom) and learning approaches (situated learning and teacher-centred learning) into consideration in the experiment design. The statistical results reveal no significant differences in students' general FLSA measured by English Speaking Anxiety Scale (ESAS) before and after the intervention. The decrease in FLSA levels in the speaking test was mainly due to task repetition and a pleasant experience. It can therefore be concluded that HiVR may not result in change to FLSA by itself, but the increase in successful experience of speaking the target language in HiVR may make a difference to FLSA levels.

In terms of using HiVR for alleviating FLSA, most participants perceived it positively. HiVR can provide an authentic environment by simulating real-life situations for communication in the target language. With the head-mounted display and interesting learning design, students are engaged in learning while using HiVR. Although there is a lack of paralinguistic cues, the use of avatars to represent real people is seen as beneficial for lowering FLSA while using HiVR. However, exporting the low anxiety while using HiVR to real-world situations seems to be a problem. According to students' experience, discomfort when wearing the head-mounted display was a common issue that needs to be solved. In addition, learning design with HiVR should be more flexible in order to provide efficient learning.

7.8 Suggestions for Future Research

Based on the findings of my study, some suggestions can be made for future research in aspects of research design, FLSA measures, learning design and other research directions.

In terms of research design, my research was conducted at one university in one location. It should be replicated at other universities in both Chinese and non-Chinese contexts. Given gender imbalance of the participants in my study, a future study could include an equal number of male and female participants to determine whether gender influences the results.

Investigating different durations of the intervention and their effects on FLSA is also suggested to be investigated in future research. Additionally, my study compares only two different learning environments (HiVR and classroom). Further research comparing HiVR with classroom settings as well as with Low-immersion Virtual Reality (LiVR) and other modalities (e.g. video conferencing) needs to be explored. Finally, other qualitative research methods, such as in-depth interviews and focus groups, could be used to gain a deeper understanding of students' experiences when using HiVR for foreign language learning.

With regard to FLSA measures, it is recommended to explore objective measurement of FLSA in order to provide accurate results for the change in FLSA levels. Involving objective assessment of oral proficiency could be considered in future studies, because the improvement in perceived oral proficiency may be related to the reduction of FLSA. It would be worth investigating if the objective oral proficiency developed during the intervention, as this may add value to the research on HiVR and FLSA. Moreover, further investigation into the construct of ESAS is essential, involving a larger data set and obtaining more in-depth data on causes of FLSA among Chinese EFL learners. The larger data set and more in-depth data also apply to the investigation of FLSA construct in other learning contexts. With regard to the specific FLSA examined in my study, future research may replicate it with different languages and different learning contexts.

In terms of the learning design in HiVR, the continuing development of HiVR technology will allow more functions of the technology to be considered in future research. The impact of HiVR on FLSA of the use or non-use of paralinguistic cue tracking are suggested to be compared as well as exposure to the virtual environment in HiVR and in real life. Different designs of the avatars can also be considered as my study only applied to white people look to represent dominant Western groups. Moreover, my study involves situated learning with role-play tasks in the learning design. It may be worthwhile to investigate other learning approaches and learning tasks in order to see if they result in differences in FLSA levels while using HiVR. In addition, mid-tests and follow-up tests could be added in the research design to investigate the change of FLSA during the intervention and the potential maintenance of its effects.

In addition to investigating HiVR and FLSA, other research directions could be proposed. First, identifying how foreign language learning can take the most advantage of HiVR's

affordances may be valuable. It is also worth to investigating how to export the low-anxiety experience in HiVR to real-world use of the target language. In addition, investigation of the impact of task repetition and positive psychology on FLSA should be carried out in future in order to find a potential way to alleviate FLSA.

REFERENCES

Abood, M. H., & Abu-Melhim, A.-R. H. (2015). Examining the effectiveness of group counseling in reducing anxiety for Jordanian EFL learners. *Journal of Language Teaching and Research, 6*(4), 749-757.

Ahmed, N. F. (2016). An Exploration of Speaking Anxiety with Kurdish University EFL Learners. *Journal of Education and Practice, 7*(27), 99-106.

Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The Modern Language Journal, 78*(2), 155-168.

Akkakoson, S. (2016). Speaking Anxiety in English Conversation Classrooms among Thai Students. *Malaysian Journal of Learning and Instruction, 13*(1), 63-82.

Al-Shuaibi, J., Hamdan-Mansour, A. M., & Azzeghaiby, S. N. (2014). Foreign language anxiety among students studying foreign languages. *Life Science Journal, 11*(8), 197-203.

Albudaiwi, D. (2017). Survey: Open-Ended Questions. In M. Allen (Ed.), *The SAGE Encyclopedia of Communication Research Methods* (Vol. 4, pp. 1716-1717). Thousand Oaks, California: SAGE Publications, Inc.

Alfadil, M. (2020). Effectiveness of virtual reality game in foreign language vocabulary acquisition. *Computers & Education, 153*, 103893.

Allen, M. (2017). The SAGE Encyclopedia of Communication Research Methods. In (Vol. 1-4): SAGE Publicationsk, Inc.

Alnahidh, F., & Altalhab, S. (2020). The Level and Sources of Foreign Language Speaking Anxiety among Saudi EFL University Students. *Advances in Language and Literary Studies, 11*(1), 55-64.

Alsina-Jurnet, I., Carvallo-Beciu, C., & Gutiérrez-Maldonado, J. (2007). Validity of virtual reality as a method of exposure in the treatment of test anxiety. *Behavior Research Methods, 39*(4), 844-851.

Amiryousefi, M. (2016). The influence of affective variables on the complexity, accuracy, and fluency in L2 oral production: The contribution of task repetition. *Journal of English Language Teaching and Learning, 8*(17), 25-48.

Amorati, R., & Venturin, B. (2021). "I command thee thou shalt speak": Understanding and counteracting foreign language anxiety in the classroom. *Ricerche di Pedagogia e Didattica. Journal of Theories and Research in Education, 16*(3), 39-57.

Anderson, P. L., Zimand, E., Hodges, L. F., & Rothbaum, B. O. (2005). Cognitive behavioral therapy for public‐speaking anxiety using virtual reality for exposure. *Depression and anxiety, 22*(3), 156-158.

Ansari, M. S. (2015). Speaking anxiety in ESL/EFL Classrooms: A holistic approach and practical study. *International Journal of Education Investigation, 2*(4), 38-46.

Arch, J. J., & Abramowitz, J. S. (2015). Exposure therapy for obsessive–compulsive disorder: An optimizing inhibitory learning approach. *Journal of Obsessive-Compulsive and Related Disorders, 6*, 174-182.

Arnold, N. (2007). Reducing foreign language communication apprehension with computer-mediated communication: A preliminary study. *System, 35*(4), 469-486.

Atkinson, T. M., Rosenfeld, B. D., Sit, L., Mendoza, T. R., Fruscione, M., Lavene, D., Basch, E. (2011). Using confirmatory factor analysis to evaluate construct validity of the Brief Pain Inventory (BPI). *Journal of Pain and Symptom Manage, 41*(3), 558-565.

Awada, G., & Diab, H. B. (2018). The Effect of Google Earth and Wiki Models on Oral Presentation Skills of University EFL Learners. *International Journal of Teaching and Learning in Higher Education, 30*(1), 36-46.

Aydın, S. (2018). Technology and foreign language anxiety: Implications for practice and future research. *Journal of Language and Linguistic Studies, 14*(2), 193-211.

Bailey, K. M. (1983). Competitiveness and anxiety in adult second language learning: Looking at and through the diary studies. In W. Welinger & M. Long (Eds.), *Classroom oriented research in second language acquisition* (pp. 67-102). Rowley, MA: Newvury House.

Baker, S. C., & MacIntyre, P. D. (2000). The role of gender and immersion in communication and second language orientations. *Language learning, 50*(2), 311-341.

Balemir, S. H. (2009). *The sources of foreign language speaking anxiety and the relationship between proficiency level and degree of foreign language speaking anxiety.* (Doctoral Dissertation, Bilkent University, Ankara). Retrieved from http://hdl.handle.net/11693/14917

Barlow, D. H. (2000). Unraveling the mysteries of anxiety and its disorders from the perspective of emotion theory. *American Psychologist, 55*(11), 1247-1263.

Bashori, M., van Hout, R., Strik, H., & Cucchiarini, C. (2022). Web-based language learning and speaking anxiety. *Computer Assisted Language Learning, 35*(5-6), 1058-1089.

Baumeister, R. F. (1999). Self-concept, self-esteem, and identity. In B. A. W. V. J. Derlega, & W. H. Jones (Eds.), *Personality: Contemporary theory and reseach* (pp. 339-375): Nelson-Hall Publishers.

Baus, O., & Bouchard, S. (2010). The sense of olfaction: Its characteristics and its possible applications in virtual environments. *Journal of CyberTherapy & Rehabilitation, 3*(1), 31-50.

Baxter, J. (2009). Content Analysis. In R. Kitchin & N. Thrift (Eds.), *International Encyclopedia of Human Geography* (Vol. 1, pp. 275-280). Oxford:Elsevier.

Benesch, S. (2013). *Considering emotions in critical English language teaching: Theories and praxis*. Routledge.

Berg, L. P., & Vance, J. M. (2017). Industry use of virtual reality in product design and manufacturing: a survey. *Virtual reality, 21*(1), 1-17.

Bertram, D. (2007). *Likert scales* (CPSC 681 – topic report). Retrived from https://www.researchgate.net/profile/Mahdi-Safarpour-2/post/what_is_a_logistic_regression_analysis/attachment/59d622fb79197b8077981515/AS%3A304626539139075%401449640034760/download/Likert+Scale+vs+Likert+Item.pdf.

Bhatt, R. M. (2001). World Englishes. *Annual review of anthropology, 30*(1), 527-550.

Blackstone, A. (2012). Inductive or deductive? Two different approaches. In A. Blackstone (Ed.), *Principles of sociological inquiry: Qualitative and quantitative methods*. Saylor Academy Open Textbooks.

Blöte, A. W., Kint, M. J., Miers, A. C., & Westenberg, P. M. (2009). The relation between public speaking anxiety and social anxiety: A review. *Journal of anxiety disorders, 23*(3), 305-313.

Bobsin, D., Petrini, M., & Pozzebon, M. (2019). The value of technology affordances to improve the management of nonprofit organizations. *RAUSP Management Journal, 54*(1), 14-37.

Botella, C., García-Palacios, A., Villa, H., Baños, R., Quero, S., Alcañiz, M., & Riva, G. (2007). Virtual reality exposure in the treatment of panic disorder and agoraphobia: A controlled study. *Clinical Psychology & Psychotherapy: An International Journal of Theory & Practice, 14*(3), 164-175.

Botella, C., Serrano, B., Baños, R. M., & Garcia-Palacios, A. (2015). Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: a review of its efficacy, the adequacy of the treatment protocol, and its acceptability. *Neuropsychiatric disease and treatment, 11*, 2533-2545.

Botes, E., Dewaele, J.-M., & Greiff, S. (2020). The foreign language classroom anxiety scale and academic achievement: An overview of the prevailing literature and a meta-analysis. *JOURNAL FOR THE PSYCHOLOGY OF LANGUAGE LEARNING, 2*(1), 26-56.

Bouchard, S., Dumoulin, S., Robillard, G., Guitard, T., Klinger, E., Forget, H., Roucaut, F. X. (2017). Virtual reality compared with in vivo exposure in the treatment of social anxiety disorder: a three-arm randomised controlled trial. *The British Journal of Psychiatry, 210*(4), 276-283.

Boudreau, C., MacIntyre, P., & Dewaele, J.-M. (2018). Enjoyment and anxiety in second language communication: An idiodynamic approach. *Studies in second language learning and teaching, 8*(1), 149-170.

Brislin, R. W., & Freimanis, C. (2001). Back-translation: A tool for cross-cultural research. In D. E. Pollard & C. Sin-wai (Eds.), *An Encyclopaedia of Translation: Chinese-English, English-Chinese* (Reprint ed., pp. 22-40). Chinese University Press.

Brown, H. D. (2000). *Principles of language learning and teaching* (Vol. 4). Longman New York.

Brown, T. A., & Moore, M. T. (2012). Confirmatory factor analysis. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 261-276). New York: NY: The Guilford Press.

Bryman, A. (2016). *Social research methods*. Oxford university press.

Burgoon, J. K. (1976). The unwillingness-to-communicate scale: Development and validation. *Communications Monographs, 43*(1), 60-69.

Bygate, M. (2013). Effects of task repetition on the structure and control of oral language. In P. S. M. Bygate & M. Swain (Eds.), *Researching Pedagogic Tasks: Second Language Learning, Teaching, and Testing* (pp. 23-48). Routledge.

Çağatay, S. (2015). Examining EFL students' foreign language speaking anxiety: The case at a Turkish state university. *Procedia-Social and Behavioral Sciences, 199*, 648-656.

Campbell, D. T., & Stanley, J. C. (2015). *Experimental and quasi-experimental designs for research*. Ravenio Books.

Carmines, E. G., & Zeller, R. A. (1979). *Reliability and Validity Assessment*. SAGE Publications, Inc.

Carrier, M., Damerow, R. M., & Bailey, K. M. (2017). *Digital language learning and teaching: Research, theory, and practice*. Taylor & Francis.

Carruth, D. W. (2017). *Virtual reality for education and workforce training.* 2017 15th International Conference on Emerging eLearning Technologies and Applications (ICETA), 1-6. Doi: 10.1109/ICETA.2017.8102472

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate behavioral research, 1*(2), 245-276.

Cervone, D., & Pervin, L. A. (2015). *Personality: Theory and research*. John Wiley & Sons.

Chen, C. H., Hung, H. T., & Yeh, H. C. (2021). Virtual reality in problem-based learning contexts: Effects on the problem-solving performance, vocabulary acquisition and motivation of English language learners. *Journal of Computer Assisted Learning, 37*(3), 851-860.

Chen, H. Y., & Boore, J. R. (2010). Translation and back-translation in qualitative nursing research: methodological review. *Journal of clinical nursing, 19*(1-2), 234-239.

Chen, J. C. (2016). The crossroads of English language learners, task-based instruction, and 3D multi-user virtual learning in Second Life. *Computers & Education, 102*, 152-171.

Chen, T. Y., & Chang, G. B. (2004). The relationship between foreign language anxiety and learning difficulties. *Foreign Language Annals, 37*(2), 279-289.

Chen, Y., Smith, T. J., York, C. S., & Mayall, H. J. (2020). Google Earth Virtual Reality and expository writing for young English Learners from a Funds of Knowledge perspective. *Computer Assisted Language Learning, 33*(1-2), 1-25.

Cheng, A., Yang, L., & Andersen, E. (2017). *Teaching language and culture with a virtual reality game.* Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems, 541-549. Doi: 10.1145/3025453.3025857

Cheng, M., & Chen, P. (2021). Applying PERMA to develop college students' English listening and speaking proficiency in China. *International Journal of English Language and Literature Studies, 10*(4), 333-350.

Cheng, Y.-S. (2004). A measure of second language writing anxiety: Scale development and preliminary validation. *Journal of second language writing, 13*(4), 313-335.

Cheng, Y. s., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language learning, 49*(3), 417-446.

Chou, M. H. (2018). Speaking Anxiety and Strategy Use for Learning English as a Foreign Language in Full and Partial English-Medium Instruction Contexts. *Tesol Quarterly, 52*(3), 611-633.

Chun, D., Kern, R., & Smith, B. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal*, *100*(S1), 64-80.

Christoforou, M., Xerou, E., & Papadima-Sophocleous, S. (2019). Integrating a virtual reality application to simulate situated learning experiences in a foreign language course. In F. Meunier, J. V. d. Vyver, L. Bradley, & S. Thouësny (Eds.), *CALL and complexity–short papers from EUROCALL* (pp. 82-87). Research-publishing.net.

Clément, R., Dörnyei, Z., & Noels, K. A. (1994). Motivation, self‐confidence, and group cohesion in the foreign language classroom. *Language learning, 44*(3), 417-448.

Clément, R., Gardner, R. C., & Smythe, P. C. (1977). Motivational variables in second language acquisition: A study of francophones learning English. *Canadian Journal of Behavioural Science/Revue canadienne des sciences du comportement, 9*(2), 123-133.

Coelho, C. M., Santos, J. A., Silvério, J., & Silva, C. F. (2006). Virtual reality and acrophobia: one-year follow-up and case study. *Cyberpsychology & behavior, 9*(3), 336-341.

Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education*. Routledge.

Cohen, Y., & Norst, M. J. (1989). Fear, dependence and loss of self-esteem: Affective barriers in second language learning among adults. *RELC journal, 20*(2), 61-77.

Couper, M. P. (2001). *Web surveys: The questionnaire design challenge.* Proceedings of the 53rd session of the ISI. Retrived from https://2001.isiproceedings.org/pdf/263.PDF.

Craig, D. A., & Kim, J. (2010). Anxiety and performance in videoconferenced and face-to-face oral interviews. *Multimedia Assisted Language Learning, 13*(3), 9-32.

Craske, M. G., Kircanski, K., Zelikowsky, M., Mystkowski, J., Chowdhury, N., & Baker, A. (2008). Optimizing inhibitory learning during exposure therapy. *Behaviour Research and Therapy, 46*(1), 5-27.

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy, 58*, 10-23.

Creswell, J. W. (2021). *A concise introduction to mixed methods research*. SAGE publications.

Creswell, J. W., & Clark, V. L. P. (2017). *Designing and conducting mixed methods research*. Sage publications.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika, 16*(3), 297-334.

Crookall, D., & Oxford, R. (1991). Dealing with anxiety: Some practical activities for language learners and teacher trainees. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (pp. 141-150): Pearson.

Crozier, W. R. (2013). *Individual learners: Personality differences in education*. Routledge.

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.

Dalkiliç, N. (2001). The role of foreign language classroom anxiety in English speaking courses. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 8*(8), 70-82.

De Andrés, V., & Arnold, J. (2009). *Seeds of confidence: Self-esteem activities for the EFL classroom*. Helbling Innsbruck.

De Jong, N., & Tillman, P. (2018). Grammatical structures and oral fluency in immediate task repetition. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 43-70). John Benjamins.

Deacon, B., Kemp, J. J., Dixon, L. J., Sy, J. T., Farrell, N. R., & Zhang, A. R. (2013). Maximizing the efficacy of interoceptive exposure by optimizing inhibitory learning: A randomized controlled trial. *Behaviour Research and Therapy, 51*(9), 588-596.

Dewaele, J.-M. (2004). The emotional force of swearwords and taboo words in the speech of multilinguals. *Journal of multilingual and multicultural development, 25*(2-3), 204-222.

Dewaele, J.-M. (2008). The emotional weight of I love you in multilinguals' languages. *Journal of Pragmatics, 40*(10), 1753-1780.

Dewaele, J.-M., & Al-Saraj, T. M. (2015). Foreign Language Classroom Anxiety of Arab learners of English: The effect of personality, linguistic and sociobiographical variables. *Studies in second language learning and teaching, 5*(2), 205-228.

Dewaele, J.-M., Chen, X., Padilla, A. M., & Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Frontiers in psychology, 10*, 1-13.

Dewaele, J.-M., & Li, C. (2020). Emotions in second language acquisition: A critical review and research agenda. *Foreign Language World, 196*(1), 34-49.

Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in second language learning and teaching, 4*(2), 237-274.

Dewaele, J.-M., MacIntyre, P. D., Boudreau, C., & Dewaele, L. (2016). Do girls have all the fun? Anxiety and enjoyment in the foreign language classroom. *Theory and practice of second language acquisition, 2*(1), 41-63.

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language teaching research, 22*(6), 676-697.

Dhimolea, T. K., Kaplan-Rakowski, R., & Lin, L. (2022). A systematic review of research on high-immersion virtual reality for language learning. *TechTrends, 66*, 810–824.

Diamantopoulos, A. (1994). Modelling with LISREL: A guide for the uninitiated. *Journal of Marketing Management, 10*(1-3), 105-136.

Dickey, M. D. (2005). Three‐dimensional virtual worlds and distance learning: two case studies of Active Worlds as a medium for distance education. *British journal of educational technology, 36*(3), 439-451.

Dittmann, A. (2018). *Correlation between tolerance of ambiguity, anxiety, and willingness to communicate with risk taking in the second language class.* Proceedings of MAC 2018 in Prague, 291-300.

Dorathy, A. A., & Mahalakshmi, S. (2011). Second language acquisition through task-based approach–role-play in English language teaching. *English for specific purposes world, 11*(33), 1-7.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics Quantitative, Qualitative, and Mixed Methodologies*. Oxford University Press.

Dörnyei, Z., & Murphey, T. (2003). *Group dynamics in the language classroom*. Ernst Klett Sprachen.

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

Douglas, S. P., & Craig, C. S. (2007). Collaborative and iterative translation: An alternative approach to back translation. *Journal of International Marketing, 15*(1), 30-43.

Ebadi, S., & Ebadijalal, M. (2020). The effect of Google Expeditions virtual reality on EFL learners' willingness to communicate and oral proficiency. *Computer Assisted Language Learning, 35*(8), 1-25.

Edwards, H., Nash, R., Sacre, S., Courtney, M., & Abbey, J. (2008). Development of a virtual learning environment to enhance undergraduate nursing students' effectiveness and interest in working with older people. *Nurse Education Today, 28*(6), 672-679.

Eelen, P., Hermans, D., & Baeyens, F. (2001). Learning perspectives on anxiety disorders. In E. J. L. Griez, C. Faravelli, D. Nutt, & D. Zohar (Eds.), *Anxiety disorders* (pp.249-264). John Wiley & Sons Ltd.

El-Sakka, S. M. F. (2016). Self-Regulated Strategy Instruction for Developing Speaking Proficiency and Reducing Speaking Anxiety of Egyptian University Students. *English Language Teaching, 9*(12), 22-33.

Elkhafaifi, H. (2005). Listening comprehension and anxiety in the Arabic language classroom. *The Modern Language Journal, 89*(2), 206-220.

Ellis, G., & Sinclair, B. (1989). *Learning to learn English learner's book: A course in learner training*. Cambridge University Press.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis: A focus on trustworthiness. *SAGE Open, 4*(1), 1-10.

Elo, S., & Kyngäs, H. (2008). The qualitative content analysis process. *Journal of advanced nursing, 62*(1), 107-115.

Ely, C. M. (1986). An analysis of discomfort, risktaking, sociability, and motivation in the L2 classroom. *Language learning, 36*(1), 1-25.

Emmelkamp, P. M., Meyerbröker, K., & Morina, N. (2020). Virtual reality therapy in social anxiety disorder. *Current psychiatry reports, 22*(7), 1-9.

Enkin, E. (2022). Comparing Two Worlds: Spanish Learners' Face-to-face and Immersive Social VR Speaking Experiences. *Computer Assisted Language Learning, 23*(1), 22-42.

Europe, C. o. (2020). *Common European framework of reference for languages: Learning, teaching, assessment*. Council of Europe Publishing, Strasbourg.

Eysenck, M. W. (1979). Anxiety, learning, and memory: A reconceptualization. *Journal of research in personality, 13*(4), 363-385.

Falkoni-Mjehović, D., Tolj, Z. R., & Baranac, S. D. (2022). Research on risk-taking, anxiety and shyness in foreign language learning. *KNOWLEDGE-International Journal, 52*(2), 267-272.

Farrer, C., Bouchereau, M., Jeannerod, M., & Franck, N. (2008). Effect of distorted visual feedback on the sense of agency. *Behavioural neurology, 19*(1-2), 53-57.

Faul, A. C. (2018). Scale. In B. B. Frey (Ed.), *The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation* (pp. 1455-1456). SAGE Publications.

Felix, U. (2002). The web as a vehicle for constructivist approaches in language teaching. *ReCALL, 14*(1), 2-15.

Feng, Z. (2007). *The perceptions of Chinese junior high and senior high students regarding the teaching and learning of English listening and speaking skills.* (Doctoral Dissertation, Texas Tech University, United States). Retrieved from http://hdl.handle.net/2346/20529.

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs—principles and practices. *Health services research, 48*(6pt2), 2134-2156.

Fetters, M. D., & Freshwater, D. (2015). The 1+ 1= 3 Integration Challenge. *Journal of mixed methods research, 9*(2), 115-117.

Fishbach, A., & Woolley, K. (2022). The structure of intrinsic motivation. *Annual Review of Organizational Psychology and Organizational Behavior, 9*, 339-363.

Fisher, R. A. (1936). Design of experiments. *British Medical Journal, 1*(3923), 554.

Foa, E. B., Huppert, J. D., & Cahill, S. P. (2006). Emotional processing theory: An update. In B. O. Rothbaum (Ed.), *Pathological anxiety: Emotional processing in etiology and treatment* (pp. 3–24). The Guilford Press.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological bulletin, 99*(1), 20-35.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*(3), 354-375.

Freina, L., & Ott, M. (2015). *A literature review on immersive virtual reality in education: state of the art and perspectives.* The International Scientific Conference eLearning and Software for Education, 133, 10-1007.

Ganschow, L., & Sparks, R. (1996). Anxiety about foreign language learning among high school women. *The Modern Language Journal, 80*(2), 199-212.

Gardner, R. (1987). *The Role of Anxiety in Second Language Performance of Language Dropouts* (No. 657). Retrieved from: https://files.eric.ed.gov/fulltext/ED323774.pdf

Gardner, R. C. (1985). *The attitude/motivation test battery: Technical report*. University of Western Ontario.

Gardner, R. C., & MacIntyre, P. D. (1993). On the measurement of affective variables in second language learning. *Language learning, 43*(2), 157-194.

Gardner, R. C., Masgoret, A. M., Tennant, J., & Mihic, L. (2004). Integrative motivation: Changes during a year-long intermediate-level language course. *Language learning, 54*(1), 1-34.

Gardner, R. C., & Smythe, P. C. (1975). *Second Language Acquisition: A Social Psychological Approach* (No. 332). Retrieved from: https://files.eric.ed.gov/fulltext/ED163754.pdf

Garrett, P. J., & Young, R. F. (2009). Theorizing Affect in Foreign Language Learning: An analysis of one learner's responses to a communicative-based Portuguese course. *The Modern Language Journal, 93*(2), 209-226.

Gass, S. M., Behney, J., & Plonsky, L. (2020). *Second language acquisition: An introductory course*. Routledge.

Gebara, C. M., Barros-Neto, T. P. d., Gertsenchtein, L., & Lotufo-Neto, F. (2015). Virtual reality exposure using three-dimensional images for the treatment of social phobia. *Brazilian Journal of Psychiatry, 38*(1), 24-29.

Gee, J. P. (2004). *Situated language and learning: A critique of traditional schooling*. Psychology Press.

Geraets, C. N., Veling, W., Witlox, M., Staring, A. B., Matthijssen, S. J., & Cath, D. (2019). Virtual reality-based cognitive behavioural therapy for patients with generalized social anxiety disorder: a pilot study. *Behavioural and cognitive psychotherapy, 47*(6), 745-750.

Girvan, C. (2018). What is a virtual world? Definition and classification. *Educational Technology Research and Development, 66*(5), 1087-1100.

Gkonou, C. (2017). Towards an ecological understanding of language anxiety. In C. Gkonou, M. Daubney, J.-M., Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 135-155). Multilingual Matters.

Gkonou, C., Daubney, M., & Dewaele, J.-M. (2017). *New insights into language anxiety: Theory, research and educational implications*. Multilingual Matters.

Greenbaum, P. (1992). The lawnmower man. *Film and video, 9*(3), 58-62.

Greenberg, J., Solomon, S., Pyszczynski, T., Rosenblatt, A., Burling, J., Lyon, D., Pinel, E. (1992). Why do people need self-esteem? Converging evidence that self-esteem serves an anxiety-buffering function. *Journal of personality and social psychology, 63*(6), 913-922.

Gregersen, T. (2013). Language learning vibes: what, why and how to capitalize for positive affect. In D. gabrys-Barker & J. Bielska (Eds.), *The affective dimension in second language acquisition* (pp. 89-98). Multilingual Matters.

Gregersen, T. (2020). Dynamic properties of language anxiety. *Studies in second language learning and teaching, 10*(1), 67-87.

Gregersen, T., & Horwitz, E. K. (2002). Language learning and perfectionism: Anxious and non-anxious language learners' reactions to their own oral performance. *The Modern Language Journal, 86*(4), 562-570.

Gregersen, T., MacIntyre, P. D., & Olsen, T. (2017). Do you see what i feel? an idiodynamic assessment of expert and peer's reading of nonverbal language anxiety cues. In C. Gkonou, M. Daubney, J.-M., Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 110-134). Multilingual Matters.

Gregersen, T., Meza, M. D., & MacIntyre, P. D. (2014). The Motion of Emotion: Idiodynamic Case Studies of Learners' Foreign Language Anxiety. *Modern Language Journal, 98*(2), 574-588.

Gregersen, T. S. (2003). To err is human: A reminder to teachers of language-anxious students. *Foreign Language Annals, 36*(1), 25-32.

Gregersen, T. S. (2005). Nonverbal cues: Clues to the detection of foreign language anxiety. *Foreign Language Annals, 38*(3), 388-400.

Gregersen, T. S. (2007). Breaking the code of silence: A study of teachers' nonverbal decoding accuracy of foreign language anxiety. *Language teaching research, 11*(2), 209-221.

Gregersen, T. S. (2009). Recognizing visual and auditory cues in the detection of foreign-language anxiety. *TESL Canada Journal, 26*(2), 46-64.

Gruber, A., & Kaplan-Rakowski, R. (2020). User experience of public speaking practice in virtual reality. In R. Z. Zheng (Ed.), *Cognitive and affective perspectives on immersive technology in education* (pp. 235-249). IGI Global.

Gruber, A., & Kaplan-Rakowski, R. (2022). The impact of high-immersion virtual reality on foreign language anxiety when speaking in public. *SSRN*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3882215

Hair, J. F. (2009). *Multivariate data analysis: A Global Perspective* (7th ed.). Upper Saddle River: Prentice Hall.

Hamamorad, A. M. (2020). Foreign Language Anxiety and Communicative Performance: A Structural Equation Modelling. *Koya University Journal of Humanities and Social Sciences, 3*(1), 199-207.

Hamilton, D., McKechnie, J., Edgerton, E., & Wilson, C. (2021). Immersive virtual reality as a pedagogical tool in education: a systematic literature review of quantitative learning outcomes and experimental design. *Journal of Computers in Education, 8*(1), 1-32.

Hammad, E. A., & Ghali, E. M. A. (2015). Speaking anxiety level of Gaza EFL pre-service teachers: Reasons and sources. *World Journal of English Language, 5*(3), 52-64.

Hammick, J. K., & Lee, M. J. (2014). Do shy people feel less communication apprehension online? The effects of virtual reality on the relationship between personality characteristics and communication outcomes. *Computers in Human Behavior, 33*, 302-310.

Han, J., & Yin, H. (2016). College English curriculum reform in Mainland China: Contexts, contents and changes. *Asian Education Studies, 1*(1), 1-10.

Han, L. (2022). Students' Daily English Situational Teaching Based on Virtual Reality Technology. *Mobile Information Systems, 2022*. doi:https://doi.org/10.1155/2022/1222501

Harris, S. R., Kemmerling, R. L., & North, M. M. (2002). Brief virtual reality therapy for public speaking anxiety. *Cyberpsychology & behavior, 5*(6), 543-550.

Hartanto, D., Kampmann, I. L., Morina, N., Emmelkamp, P. G., Neerincx, M. A., & Brinkman, W.-P. (2014). Controlling social stress in virtual reality environments. *PloS one, 9*(3), e92804.

Hartley, C. A., & Phelps, E. A. (2012). Extinction Learning. In N. M. Seel (Ed.), *Encyclopedia of the Sciences of Learning* (pp. 1252-1253). Boston, MA: Springer US.

He, D. (2011). *Foreign language speaking anxiety: An investigation of non-English majors in mainland China.* (Doctoral Dissertation, City University of Hong Kong, Hong Kong). Retrieved from https://scholars.cityu.edu.hk/en/theses/theses(5e9699a2-3f24-4868-b265-bcaedc1e212d).html

He, D. (2013). What makes learners anxious while speaking English: A comparative study of the perceptions held by university students and teachers in China. *Educational Studies, 39*(3), 338-350.

He, D. (2017). How to Cope with Foreign Language Speaking Anxiety Effectively? The Case of University Students in China. *Electronic Journal of Foreign Language Teaching, 14*(2), 159–174.

He, D. (2018). *Foreign Language Learning Anxiety in China: Theories and Applications in English Language Teaching*. Springer.

Herrington, J., & Oliver, R. (2000). *Critical characteristics of situated learning: Implications for the instructional design of multimedia*. Paper presented at the ASCILITE 1995 Conference, University of Melbourne, Melbourne. Retrieved from https://researchportal.murdoch.edu.au/esploro/outputs/conferencePaper/Critical-characteristics-of-situated-learning-Implications/991005540189907891/filesAndLinks?index=0.

Hewitt, E., & Stephenson, J. (2012). Foreign language anxiety and oral exam performance: A replication of Phillips's MLJ study. *The Modern Language Journal, 96*(2), 170-189.

Hilfert, T., & König, M. (2016). Low-cost virtual reality environment for engineering and construction. *Visualization in Engineering, 4*(2), 1-18.

Hofmann, S. G., Asnaani, A., Vonk, I. J., Sawyer, A. T., & Fang, A. (2012). The efficacy of cognitive behavioral therapy: A review of meta-analyses. *Cognitive Therapy and Research, 36*(5), 427-440.

Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching, 43*(2), 154-167.

Horwitz, E. K. (2016). Factor structure of the foreign language classroom anxiety scale: Comment on Park (2014). *Psychological reports, 119*(1), 71-76.

Horwitz, E. K. (2017). On the misreading of Horwitz, Horwitz, and Cope (1986) and the need to balance anxiety research and the experiences of anxious language learners. In C. Gkonou, M. Daubney, J.-M., Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 31-47). Multilingual Matters.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132.

Hossainy, F. N., Zare, H., Hormozi, M., Shaghaghi, F., & Kaveh, M. H. (2012). Designing and Implementing a Situated Learning Program and Determining Its Impact on the Students' Motivation and Learning. *Turkish Online Journal of Distance Education, 13*(2), 36-47.

Hu, G. (2002). Potential cultural resistance to pedagogical imports: The case of communicative language teaching in China. *Language Culture and Curriculum, 15*(2), 93-105.

Hu, L. t., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal, 6*(1), 1-55.

Hu, P. J.-H., & Hui, W. (2012). Examining the role of learning engagement in technology-mediated learning and its effects on learning effectiveness and satisfaction. *Decision support systems, 53*(4), 782-792.

Hu, X. (2018). Effects of task type, task-type repetition, and performance criteria on L2 oral production. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 143-169). John Benjamins.

Huang, H. (2004). *The relationship between learning motivation and speaking anxiety among EFL non-English major freshmen in Taiwan.* (Master's Dissertation, Chaoyang University of Technology, Taiwan). Retrieved from http://ethesys.lib.cyut.edu.tw/ETD-db/ETD-search-c/getfile?URN=etd-0706105-031009&filename=etd-0706105-031009.pdf

Ikart, E. M. (2018). Questionnaire Pretesting Methods: A Comparison of Cognitive Interviewing and Respondent Debriefing Vis-à-vis the Study of the Adoption of Decision Support Systems by Knowledge Workers. *International Journal of Business & Information, 13*(2), 119-154.

Jacoby, R. J., & Abramowitz, J. S. (2016). Inhibitory learning approaches to exposure therapy: A critical review and translation to obsessive-compulsive disorder. *Clinical psychology review, 49*, 28-40.

Jee, M. J. (2018). Foreign language anxiety in relation to affective variables: Learners of Korean as a foreign language in Australia. *Australian Review of Applied Linguistics, 41*(3), 328-348.

Jeong, J.-Y., & Jeong, H. (2021). Effects of immersive virtual reality English conversations on language anxiety and learning achievement. *The Journal of the Korea Contents Association, 21*(1), 321-332.

Jiang, Y., & Dewaele, J.-M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners? *System, 82*, 13-25.

Jin, L., & Cortazzi, M. (2006). Changing practices in Chinese cultures of learning. *Language, Culture and Curriculum, 19*(1), 5-20.

Jin, Y., Dewaele, J.-M., & MacIntyre, P. D. (2021). Reducing anxiety in the foreign language classroom: A positive psychology approach. *System, 101*, 102604.

Johnson-Glenberg, M. C. (2019). The necessary nine: Design principles for embodied VR and active stem education. In P. Díaz, A. Ioannou, K. K. bhagat, & J. M. Spector (Eds.), *Learning in a digital world: Perspective on interactive technologies for formal and informal education* (pp. 83-112). Springer.

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal of mixed methods research, 1*(2), 112-133.

Jung, H.-J. (2002). Virtual reality for ESL students. *The internet TESL journal, 8*(10), 1-7.

Jupp, V. (2006). *The SAGE Dictionary of Social Research Methods*. London: SAGE Publications, Ltd.

Kachru, B. (1996). Norms, models, and identities. *JALT, 20*(10), 25-32.

Kaiser, H. F. (1970). A second generation little jiffy. *psychometrika, 35*, 401–415.

Kamisznikow, A. (2003). Towards a psycholinguistic understanding of non-verbal behaviour. *Scripta Neophilologica Posnaniensia, 5*, 67-84.

Kampmann, I. L., Emmelkamp, P. M., Hartanto, D., Brinkman, W.-P., Zijlstra, B. J., & Morina, N. (2016). Exposure to virtual social interactions in the treatment of social anxiety disorder: A randomized controlled trial. *Behaviour Research and Therapy, 77*, 147-156.

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

Kaplan-Rakowski, R., & Gruber, A. (2021). One-on-one foreign language speaking practice in high-immersion virtual reality. In Y.-J. Lan & S. Grant (Eds.), *Contextual Language Learning: Real Language Learning on the Continuum from Virtuality to Reality* (pp. 187-202). Springer.

Kaplan-Rakowski, R., & Wojdynski, T. (2018). *Students' attitudes toward high-immersion virtual reality assisted language learning.* Future-proof CALL: Language learning as exploration and encounters—short papers from EUROCALL, 124-129

Karatas, H., Alci, B., Bademcioglu, M., & Ergin, A. (2016). An investigation into university students' foreign language speaking anxiety. *Procedia-Social and Behavioral Sciences, 232*, 382-388.

Kasbi, S., & Shirvan, M. E. (2017). Ecological understanding of foreign language speaking anxiety: emerging patterns and dynamic systems. *Asian-Pacific Journal of Second and Foreign Language Education, 2*(2), 1-20.

Keith, T. Z. (2019). *Multiple regression and beyond: An introduction to multiple regression and structural equation modeling*. Routledge.

Kendall, J. (2003). Designing a research project: randomised controlled trials and their principles. *Emergency medicine journal: EMJ, 20*, 164-168.

Ketelhut, D., Dede, C., Clarke, J., Nelson, B., & Bowman, C. (2007). Studying situated learning in a multi-user environment. In B. D. W. O'NeilH (Ed.), *Assessment of problem solving using simulations. New York: Lawrence Erlbaum Associates*. Routledge.

Kim, Y., Crossley, S., Jung, Y., Kyle, K., & Kang, S. (2018). The effects of task repetition and task complexity on L2 lexicon use. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 75-96). John Benjamins.

King, J., & Smith, L. (2017). Social anxiety and silence in Japan's tertiary foreign language classrooms. In C. Gkonou, M. Daubney, J.-M., Dewaele (Eds.), *New insights into language anxiety: Theory, research and educational implications* (pp. 91-109). Multilingual Matters.

Kirkpatrick, A., & Zhichang, X. (2002). Chinese pragmatic norms and 'China English'. *World Englishes, 21*(2), 269-279.

Kitano, K. (2001). Anxiety in the college Japanese language classroom. *The Modern Language Journal, 85*(4), 549-566.

Kneebone, R. L., Scott, W., Darzi, A., & Horrocks, M. (2004). Simulation and clinical practice: strengthening the relationship. *Medical education, 38*(10), 1095-1102.

Koch, A. S., & Terrell, T. D. (1991). Affective reactions of foreign language students to natural approach activities and teaching techniques. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom implications* (Vol. 3, pp. 109-126). Englewood Cliffs, NJ: Prentice Hall.

Kontra, C., Goldin‐Meadow, S., & Beilock, S. L. (2012). Embodied learning across the life span. *Topics in cognitive science, 4*(4), 731-739.

Kozak, M. J., Foa, E. B., & Steketee, G. (1988). Process and outcome of exposure treatment with obsessive-compulsives: Psychophysiological indicators of emotional processing. *Behavior therapy, 19*(2), 157-169.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon Press Inc.

Krashen, S. D. (1985). *The input hypothesis: Issues and implications*. London: Longman.

Krashen, S. D., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom*. Padstow, Cornwall: T.J. International Ltd.

Krippendorff, K. (2018). *Content analysis: An introduction to its methodology*. Sage publications.

Kruk, M. (2016). Variations in motivation, anxiety and boredom in learning English in Second Life. *The EuroCALL Review, 24*(1), 25-39.

Kruk, M. (2019). Dynamicity of perceived willingness to communicate, motivation, boredom and anxiety in Second Life: the case of two advanced learners of English. *Computer Assisted Language Learning*. *35*(1-2), 190-216.

Lang, A. J., & Craske, M. G. (2000). Manipulations of exposure-based therapy to reduce return of fear: A replication. *Behaviour Research and Therapy, 38*(1), 1-12.

Lang, A. J., Craske, M. G., & Bjork, R. A. (1999). Implications of a new theory of disuse for the treatment of emotional disorders. *Clinical Psychology: Science and Practice, 6*(1), 80-94.

Lang, P. J. (1977). Imagery in therapy: An information processing analysis of fear. *Behavior therapy, 8*(5), 862-886.

Lang, P. J. (1984). Cognition in emotion: Concept and action. In C. E. Izard, J. Kagan, R. B. Zajonc (Eds.) *Emotions, cognition, and behavior* (pp. 192-226). Cambridge Unviersity Press.

Larrivee, B. (2005). *Authentic classroom management: Creating a learning community and building reflective practice*. Pearson Allyn and Bacon.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.

Leary, M. R., & Kowalski, R. M. (1997). *Social anxiety*. Guilford Press.

Lee, E. A.-L., & Wong, K. W. (2014). Learning with desktop virtual reality: Low spatial ability learners are more positively affected. *Computers & Education, 79*, 49-58.

Legault, J., Zhao, J., Chi, Y.-A., Chen, W., Klippel, A., & Li, P. (2019). Immersive virtual reality as an effective tool for second language vocabulary learning. *Languages, 4*(1), 13.

Li, Y. (2016). *A challenging path: Seeking improvement for English speaking anxiety in Chinese university students through an intervention-based action research with drama.* (Doctoral Dissertation, University of Southampton, United Kingdom). Retrieved from https://eprints.soton.ac.uk/398054/.

Liao, C.-w., & Wei, Y. (2010). *Statistical Analyses for the TOEIC speaking and writing pilot study.* Research Foundation for TOEIC: A Compendium of Studies, Educational Testing Service, 9.1-9.25

Liaw, M.-L. (2019). EFL learners' intercultural communication in an open social virtual environment. *Journal of Educational Technology & Society, 22*(2), 38-55.

Lim, C. S. (2007). Characteristics of mathematics teaching in Shanghai, China: Through the lens of a Malaysian. *Mathematics Education Research Journal, 19*(1), 77-88.

Lin, Y., & Lin, Y. (2020). The enlightenment of affective filter hypothesis and risk-taking on English learning. *Studies in Literature and Language, 20*(2), 51-57.

Lincoln, Y. S., & Guba, E. G. (1985). Naturalistic inquiry. In N. Salkind (Ed.), *Encyclopedia of research design* (pp. 880-885): sage.

Lindridge, A. M., & Wang, C. (2008). Saving "face" in China: modernization, parental pressure, and plastic surgery. *Journal of Consumer Behaviour: An International Research Review, 7*(6), 496-508.

Liu, M. (2006). Anxiety in EFL classrooms: Causes and consequences. *TESL Reporter, 39*(1), 13-32.

Liu, M. (2007). Anxiety in oral English classrooms: A case study in China. *Indonesian JELT, 3*(1), 119-137.

Liu, M. (2018). Interactive effects of English-speaking anxiety and strategy use on oral English test performance of high-and low-proficient Chinese university EFL learners. *Cogent Education, 5*, 1562410.

Liu, M., & Huang, W. (2011). An exploration of foreign language anxiety and English learning motivation. *Education Research International, 2011*, 493167.

Liu, M., & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *The Modern Language Journal, 92*(1), 71-86.

Liu, M., & Jackson, J. (2011). Reticence and anxiety in oral English lessons: a case study in China. In L. Jin & M. Cortazzi (Eds.), *Researching Chinese learners: Skills, perceptions and intercultural adaptations* (pp. 119-137). Palgrave Macmillan.

Long, J. S. (1983). *Confirmatory Factor Analysis*. SAGE Publications, Inc.

Long, M. H. (1980). *Input, interaction, and second language acquisition*. University of California, Los Angeles.

Luoma, S. (2004). *Assessing speaking*. Ernst Klett Sprachen.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological methods, 1*(2), 130–149.

MacIntyre, P. D. (1995). How does anxiety affect second language learning? A reply to Sparks and Ganschow. *The Modern Language Journal, 79*(1), 90-99.

MacIntyre, P. D. (2002). Motivation, anxiety and emotion in second language acquisition. In P. Robinso (Ed.), *Individual differences and instructed language learning* (pp. 45-68). John Benjamins.

MacIntyre, P. D. (2007). Willingness to communicate in the second language: Understanding the decision to speak as a volitional process. *The Modern Language Journal, 91*(4), 564-576.

MacIntyre, P. D. (2017). An overview of language anxiety research and trends in its development. In C. Gkonou, M. Daubney, & J.-M. Dewaele (Ed.), *New insights into language anxiety: Theory, research and educational implications* (pp. 11-30). Mutilingual Matters.

MacIntyre, P. D., Clément, R., Dörnyei, Z., & Noels, K. A. (1998). Conceptualizing willingness to communicate in a L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal, 82*(4), 545-562.

MacIntyre, P. D., & Gardner, R. C. (1991a). Investigating language class anxiety using the focused essay technique. *The Modern Language Journal, 75*(3), 296-304.

MacIntyre, P. D., & Gardner, R. C. (1991b). Language anxiety: Its relationship to other anxieties and to processing in native and second languages. *Language learning, 41*(4), 513-534.

MacIntyre, P. D., & Gardner, R. C. (1991c). Methods and results in the study of anxiety and language learning: A review of the literature. *Language learning, 41*(1), 85-117.

MacIntyre, P. D., & Gardner, R. C. (1994). The subtle effects of language anxiety on cognitive processing in the second language. *Language learning, 44*(2), 283-305.

MacIntyre, P. D., & Legatto, J. J. (2011). A dynamic system approach to willingness to communicate: Developing an idiodynamic method to capture rapidly changing affect. *Applied Linguistics, 32*(2), 149-171.

Mahmoodzadeh, M. (2012). Investigating foreign language speaking anxiety within the EFL learner's interlanguage system: The case of Iranian learners. *Journal of Language Teaching and Research, 3*(3), 466-476.

Mahmoodzadeh, M. (2015, August). A Study on the Dynamic Peer Orientation of Foreign lLnguage Anxiety: A Classroom Perspective. *Humanising Language Teaching*. *17(4),* 1-7.

Mak, B. (2011). An exploration of speaking-in-class anxiety with Chinese ESL learners. *System, 39*, 202-214.

Mak, B., & White, C. (1997). Communication apprehension of Chinese ESL students. *Hong Kong Journal of Applied Linguistics, 2*(1), 81-96.

Makransky, G., & Lilleholt, L. (2018). A structural equation modeling investigation of the emotional value of immersive virtual reality in education. *Educational Technology Research and Development, 66*(5), 1141-1164.

Makransky, G., & Petersen, G. B. (2021). The cognitive affective model of immersive learning (CAMIL): A theoretical research-based model of learning in immersive virtual reality. *Educational Psychology Review, 33*(3), 937-958.

Makransky, G., Terkildsen, T. S., & Mayer, R. E. (2019). Adding immersive virtual reality to a science lab simulation causes more presence but less learning. *Learning and instruction, 60*, 225-236.

Maltby, N., Kirsch, I., Mayers, M., & Allen, G. J. (2002). Virtual reality exposure therapy for the treatment of fear of flying: a controlled investigation. *Journal of consulting and clinical psychology, 70*(5), 1112–1118.

Martin, S., & Valdivia, I. M. A. (2017). Students' feedback beliefs and anxiety in online foreign language oral tasks. *International Journal of Educational Technology in Higher Education, 14*(18). Retrieved from https://doi.org/10.1186/s41239-017-0056-z.

Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System, 32*(1), 21-36.

Mayo-Wilson, E., Dias, S., Mavranezouli, I., Kew, K., Clark, D. M., Ades, A., & Pilling, S. (2014). Psychological and pharmacological interventions for social anxiety disorder in adults: a systematic review and network meta-analysis. *The Lancet Psychiatry, 1*(5), 368-376.

McCroskey, J. C. (1970). Measures of communication‐bound anxiety. *Communication Monographs, 37*(4), 269-277.

McCroskey, J. C. (1977). Oral communication apprehension: A summary of recent theory and research. *Human communication research, 4*(1), 78-96.

McCroskey, J. C. (1984). The communication apprehension perspective. In J. A. Daly & J. C. McCroskey (Eds.), *Avoiding communication: Shyness, reticence, and communication apprehension* (pp. 13-38). SAGE Publications.

McLellan, H. (1991). Virtual environments and situated learning. *Multimedia Review, 2*(3), 30-37.

McLellan, H. (1996). *Situated learning perspectives*. Englewood Cliffs, New Jersey: Educational Technology Publications.

Mede, E., & Karaırmak, Ö. (2017). The predictor roles of speaking anxiety and English self efficacy on foreign language speaking anxiety. *Journal of Teacher Education and Educators, 6*(1), 117-131.

Mei, H. H., & Sheng, L. S. (2011). Applying situated learning in a virtual reality system to enhance learning motivation. *International journal of information and education technology, 1*(4), 298-302.

Melouah, A. (2013). Foreign Language Anxiety in EFL Speaking Classrooms: A Case Study of First-year LMD Students of English at Saad Dahlab University of Blida, Algeria. *Arab World English Journal, 4*(1), 64-76.

Mon, P. (2019). Alleviating Students' Foreign Language Speaking Anxiety through Cooperative Learning. *International Journal of Trend in Scientific Research and Development, 3*(5), 1368-1375.

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Consciousness and cognition, 21*(1), 59-68.

Morina, N., Brinkman, W.-P., Hartanto, D., Kampmann, I. L., & Emmelkamp, P. M. (2015). Social interactions in virtual reality exposure therapy: A proof-of-concept pilot study. *Technology and Health Care, 23*(5), 581-589.

Morina, N., Kampmann, I., Emmelkamp, P., Barbui, C., & Hoppen, T. H. (2021). Meta-analysis of virtual reality exposure therapy for social anxiety disorder. *Psychological Medicine, 53*, 2176–2178.

Nagahashi, T. L. (2007). Techniques for reducing foreign language anxiety: Results of a successful intervention study. *Annual research report on general education, Akita University*. 56-70.

Neuendorf, K. A. (2017). *The content analysis guidebook* (2nd ed.). SAGE Publications.

Newbutt, N., Bradley, R., & Conley, I. (2019). Using virtual reality head-mounted displays with autistic children: views, experiences and future directions. *Cyberpsychology, behavior and social networking, 23*(1), 23–33.

Nitta, R., & Baba, K. (2018). Understanding benefits of repetition from a complex dynamic systems perspective. In M. Bygate (Ed.), *Learning language through task repetition* (pp. 279-309). John Benjamins.

Ockey, G. J. (2013). Exploratory factor analysis and structural equation modeling. In A. J. Kunnan (Ed.), *The Companion to Language Assessment Evaluation, Methodology, and Interdisciplinary Themes Quantitative Analysis* (pp. 1224-1244). Wiley-Blackwell.

OECD, O. (2015). *Students, computers and learning: Making the connection*. PISA, OECD Publishing.

Ohata, K. (2005). Potential sources of anxiety for Japanese learners of English: Preliminary case interviews with five Japanese college students in the US. *TESL-EJ, 9*(3), 1-21.

Oing, T., & Prescott, J. (2018). Implementations of virtual reality for anxiety-related disorders: systematic review. *JMIR serious games, 6*(4), e10965.

Oleinik, A., Popova, I., Kirdina, S., & Shatalova, T. (2014). On the choice of measures of reliability and validity in the content-analysis of texts. *Quality & Quantity, 48*(5), 2703-2718.

Omaggio, A. (1992). Thoughts on language anxiety by Alice Omaggio, Stephen Krashen, Tracy Terrell, and Jennybelle Rardin. *Foreign Language Annals, 25*, 157-172.

Onwuegbuzie, A. J., Bailey, P., & Daley, C. E. (1999). Factors associated with foreign language anxiety. *Applied Psycholinguistics, 20*(2), 217-239.

Opriş, D., Pintea, S., García‐Palacios, A., Botella, C., Szamosközi, Ş., & David, D. (2012). Virtual reality exposure therapy in anxiety disorders: a quantitative meta‐analysis. *Depression and anxiety, 29*(2), 85-93.

Orth, U., & Robins, R. W. (2014). The development of self-esteem. *Current directions in psychological science, 23*(5), 381-387.

Ozawa, Y. (2006). *Authentic communication for Japanese language learning: A single case study of Midwestern university students.* (Doctoral Dissertation, The University of Nebraska-Lincoln, Retrieved from https://www.proquest.com/docview/305274694?pq-origsite=gscholar&fromopenview=true.

Öztürk, G., & Gürbüz, N. (2014). Speaking anxiety among Turkish EFL learners: The case at a state university. *Journal of Language and Linguistic Studies, 10*(1), 1-17.

Pakdaman, A., Alibakhshi, G., & Baradaran, A. (2022). The Impact of Negotiated Syllabus on Foreign Language Learners' Language Anxiety and Learning Motivation. *Teaching English Language, 16*(1), 35-63.

Palacios, L. M. (1999). *Foreign language anxiety and classroom environment: A study of Spanish university students.* (Doctoral Dissertation, The University of Texas at Austin, United States). Retrieved from https://www.proquest.com/dissertations-theses/foreign-language-anxiety-classroom-environment/docview/304457490/se-2.

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS*. Routledge.

Park, G.-P. (2014). Factor analysis of the foreign language classroom anxiety scale in Korean learners of English as a foreign language. *Psychological reports, 115*(1), 261-275.

Parvaneh, H., Zoghi, M., & Asadi, N. (2022). Effect of the Flipped Classroom Approach and Language Proficiency on Learner Autonomy and Foreign Language Anxiety. *International Journal of Foreign Language Teaching and Research, 10*(41), 77-88.

Paulhus, D. L., & Vazire, S. (2007). The self-report method. In R. W. Robins, R. C. Fraley, & R. F. Krueger (Eds.), *Handbook of research methods in personality psychology* (Vol. 1, pp. 224-239). The Guilford Press.

Peixoto, B., Pinto, D., Krassmann, A., Melo, M., Cabral, L., & Bessa, M. (2019). *Using virtual reality tools for teaching foreign languages.* In A. Rocha, H. Adeli, L. Reis, & S. Costanzo (Eds.), *New knowledge in inforamtion systems and technologies. WorldCIST19 2019. Advances in intelligent systems and computing* (pp. 581-588). Springer.

Peixoto, B., Pinto, R., Melo, M., Cabral, L., & Bessa, M. (2021). Immersive virtual reality for foreign language education: A prisma systematic review. *IEEE Access, 9*, 48952-48962.

Pérez Castillejo, S. (2019). The role of foreign language anxiety on L2 utterance fluency during a final exam. *Language Testing, 36*(3), 327-345.

Petersen, G. B., Petkakis, G., & Makransky, G. (2022). A study of how immersion and interactivity drive VR learning. *Computers & Education, 179*, 104429.

Phillips, E. M. (1992). The effects of language anxiety on students' oral test performance and attitudes. *The Modern Language Journal, 76*(1), 14-26.

Phillips, E. M. (1999). Decreasing language anxiety: Practical techniques for oral activities. In D. J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere* (pp. 124-143). McGraw-Hills, Boston.

Pica, T., Kang, H.-S., & Sauro, S. (2006). Information gap tasks: Their multiple roles and contributions to interaction research methodology. *Studies in second language acquisition, 28*(2), 301-338.

Price, M. (1988). *Anxiety and the foreign language learner: Correlates of foreign language anxiety* (Doctoral Dissertation, The University of Texas at Austin, United States). Retrieved from https://www.proquest.com/docview/303711646?pq-origsite=gscholar&fromopenview=true.

Pyun, D. O., Kim, J. S., Cho, H. Y., & Lee, J. H. (2014). Impact of affective variables on Korean as a foreign language learners' oral achievement. *System, 47*, 53-63.

Qiu, X. B., Shan, C., Yao, J., & Fu, Q. K. (2023). The effects of virtual reality on EFL learning: A meta-analysis. *Education and Information Technologies*, 1-27.

Radianti, J., Majchrzak, T. A., Fromm, J., & Wohlgenannt, I. (2020). A systematic review of immersive virtual reality applications for higher education: Design elements, lessons learned, and research agenda. *Computers & Education, 147*, 103778.

Rafieyan, V., & Yamanashi, K. (2016). Discovering factors of foreign language speaking anxiety and coping strategies. *Journal for the Study of English Linguistics, 4*(1), 111-125.

Rassaei, E. (2015). The effects of foreign language anxiety on EFL learners' perceptions of oral corrective feedback. *Innovation in Language Learning and Teaching, 9*(2), 87-101.

Reeve, C. L., Heggestad, E. D., & Lievens, F. (2009). Modeling the impact of test anxiety and test familiarity on the criterion-related validity of cognitive ability tests. *Intelligence, 37*(1), 34-41.

Reinders, H., & Wattana, S. (2015). Affect and willingness to communicate in digital game-based learning. *ReCALL, 27*(1), 38-57.

Reja, U., Manfreda, K. L., Hlebec, V., & Vehovar, V. (2003). Open-ended vs. close-ended questions in web questionnaires. *Developments in applied statistics, 19*(1), 159-177.

Riff, D., Lacy, S., & Fico, F. (2014). *Analyzing media messages: Using quantitative content analysis in research*. Routledge.

Riva, G. (2005). Virtual reality in psychotherapy. *Cyberpsychology & behavior, 8*(3), 220-230.

Riva, G.,& IJsselsteijn, W. A. (2003). Being there: The experience of presence in mediated environments. In G. Riva, F. Davide, & W.A IJsselsteijn (Eds.), *Being there: Concepts, effects and measurement of user presence in synthetic environments* (pp. 3-14). IOS Press.

Rosell-Aguilar, F. (2005). Task design for audiographic conferencing: Promoting beginner oral interaction in distance language learning. *Computer Assisted Language Learning, 18*(5), 417-442.

Ross, S. M., & Morrison, G. R. (2004). Experimental research methods. In J. M. Spector, M. D. Merrill, J. Elen & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 1021-1043). Springer.

Rourke, L., & Anderson, T. (2004). Validity in quantitative content analysis. *Educational Technology Research and Development, 52*(1), 5-18.

Rubio, F. D. (2014). Self-esteem and self-concept in foreign language learning. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA* (pp. 41-58). Multilingual Matters.

Rui, T., & Chew, P. G.-L. (2013). Pedagogical use of two languages in a Chinese elementary school. *Language, Culture and Curriculum, 26*(3), 317-331.

Ryan, W. S., Cornick, J., Blascovich, J., & Bailenson, J. N. (2019). Virtual reality: Whence, how and what for. In A. Rizzo & S. Bouchard (Eds.), *Virtual Reality for Psychological and Neurocognitive Interventions* (pp. 15-46). Springer.

Safir, M. P., Wallach, H. S., & Bar-Zvi, M. (2012). Virtual reality cognitive-behavior therapy for public speaking anxiety: one-year follow-up. *Behavior modification, 36*(2), 235-246.

Saito, Y., Garza, T. J., & Horwitz, E. K. (1999). Foreign language reading anxiety. *The Modern Language Journal, 83*(2), 202-218.

Saito, Y., & Samimy, K. K. (1996). Foreign language anxiety and language performance: A study of learner anxiety in beginning, intermediate, and advanced - level college students of Japanese. *Foreign Language Annals, 29*(2), 239-249.

Sakamoto, Y., & Sakata, N. (2018). A pilot study of medical English language learning materials using virtual reality and a communication robot. *Journal of Medical English Education, 17*(3), 117-120.

Salem, A. A. M., & Al Dyiar, M. A. (2014). The relationship between speaking anxiety and oral fluency of special education Arab learners of English. *Asian Social Science, 10*(12), 170-176.

Sally Wu, Y.-H., & Alan Hung, S.-T. (2022). The Effects of Virtual Reality Infused Instruction on Elementary School Students' English-Speaking Performance, Willingness to Communicate, and Learning Autonomy. *Journal of Educational Computing Research, 60*(6), 1558-1587.

Saltan, F. (2003). *EFL speaking anxiety: How do students and teachers perceive it.* (Master's Dissertation, Middle East Technical University, Turkey). Retrieved from https://hdl.handle.net/11511/13325

Samuda, V., & Bygate, M. (2008). *Tasks in second language learning*. Springer.

Sanaei, O., Zafarghandi, A. M., & Sabet, M. K. (2015). The effect of classroom anxiety on EFL learner's oral narratives fluency: the case of intermediate level students. *Theory and Practice in Language Studies, 5*(7), 1390-1440.

Schell, J. W., & Black, R. S. (1997). Situated learning: An inductive case study of a collaborative learning experience. *Journal of Industrial Teacher Education, 34*(3), 5-28.

Schlenker, B. R., & Leary, M. R. (1982). Social anxiety and self-presentation: A conceptualization model. *Psychological bulletin, 92*(3), 641–669.

Schuemie, M. J., Van Der Straaten, P., Krijn, M., & Van Der Mast, C. A. (2001). Research on presence in virtual reality: A survey. *Cyberpsychology & behavior, 4*(2), 183-201.

Segalowitz, N. (2010). *Cognitive bases of second language fluency*. Routledge.

Serrano, B., Botella, C., Baños, R. M., & Alcañiz, M. (2013). Using virtual reality and mood-induction procedures to test products with consumers of ceramic tiles. *Computers in Human Behavior, 29*(3), 648-653.

Sharma, Y. (2016). Alleviating Mathematics Anxiety of Elementary School Students: A Situated Perspective. *International Journal of Research in Education and Science, 2*(2), 509-517.

Shi, X., Brinthaupt, T. M., & McCree, M. (2015). The relationship of self-talk frequency to communication apprehension and public speaking anxiety. *Personality and Individual Differences, 75*, 125-129.

Shih, Y.-C., & Yang, M.-T. (2008). A collaborative virtual environment for situated language learning using VEC3D. *Journal of Educational Technology & Society, 11*(1), 56-68.

Slater, M. (2003). A note on presence terminology. *Presence connect, 3*(3), 1-5.

Slater, M., & Wilbur, S. (1997). A framework for immersive virtual environments (FIVE): Speculations on the role of presence in virtual environments. *Presence: Teleoperators & Virtual Environments, 6*(6), 603-616.

Smelser, N. J., & Baltes, P. B. (Eds.). (2001). *International encyclopedia of the social & behavioral sciences* (Vol. 11). Amsterdam: Elsevier.

Smith, O. (2017). Mapped: Where to go if you can't be bothered to learn the language. *The Telegraph*. Retrieved from https://www.telegraph.co.uk/travel/maps-and-graphics/mapped-english-speaking-countries/.

Soto, J. B., Ocampo, D. T., Colon, L. B., & Oropesa, A. V. (2020). Perceptions of ImmerseMe virtual reality platform to improve English communicative skills in higher education. *International Journal of Interactive Mobile Technologies, 14*(7), 4-19.

Sparks, R. L., & Ganschow, L. (1991). Foreign language learning differences: Affective or native language aptitude differences? *The Modern Language Journal, 75*(1), 3-16.

Spence, J. T. (1984). *Achievement and Achievement Motives: Psychological and Sociological Approaches*. W.H. Freeman.

Spielberg, S. (Director). (2018). Ready player one [Film]. Warner Brothers.

Spielberger, C. D. (1971). Notes and comments trait-state anxiety and motor behavior. *Journal of motor behavior, 3*(3), 265-279.

Stein, M. B., & Stein, D. J. (2008). Social anxiety disorder. *The lancet, 371*, 1115-1125.

Steinberg, F. S., & Horwitz, E. K. (1986). The effect of induced anxiety on the denotative and interpretive content of second language speech. *Tesol Quarterly, 20*(1), 131-136.

Stemler, S. (2000). An overview of content analysis. *Practical assessment, research, and evaluation, 7*(17), 1-6.

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

Stevens, J. P. (2012). *Applied multivariate statistics for the social sciences*. Routledge.

Su, C. (2019). Characteristics and limitations of English language teaching in China: Autoethnography of a Mainland-born English learner and teacher. *Changing English, 26*(2), 188-197.

Subekti, A. S. (2018). Investigating the Relationship between Foreign Language Anxiety and Oral Performance of Non-English Major University Students in Indonesia. *Dinamika Ilmu, 18*(1), 15-35.

Suleimenova, Z. (2013). Speaking anxiety in a foreign language classroom in Kazakhstan. *Procedia-Social and Behavioral Sciences, 93*, 1860-1868.

Sun, R. (2022). *Influencing Factors and Coping Strategies of English Learning Anxiety among Chinese Non-English Major Students.* Paper presented at the 8th International Conference on Humanities and Social Science Research (ICHSSR 2022), 2843-2847. Retrieved from https://www.atlantis-press.com/proceedings/ichssr-22/125974736.

Sutherland, I. E. (1968). *A head-mounted three dimensional display.* Proceedings of the December 9-11, 1968, Fall Joint Computer Conference, part I, 757-764. Retrieved from https://doi.org/10.1145/1476589.1476686.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistics* (Vol. 5). Pearson Boston, MA.

Taguchi, N. (2022). Immersive virtual reality for pragmatics task development. *Tesol Quarterly, 56*(1), 308-335.

Tai, T.-Y., Chen, H. H.-J., & Todd, G. (2022). The impact of a virtual reality app on adolescent EFL learners' vocabulary learning. *Computer Assisted Language Learning, 35*(4), 892-917.

Tajeddin, Z., & Pakzadian, M. (2020). Representation of inner, outer and expanding circle varieties and cultures in global ELT textbooks. *Asian-Pacific Journal of Second and Foreign Language Education, 5*(1), 1-15.

Tanner, K. (2018). Experimental research. In K. Williamson & G. Johanson (Eds.), *Research Methods* (pp. 337-356). Elsevier.

Tanveer, M. (2007). *Investigation of the factors that cause language anxiety for ESL/EFL learners in learning speaking skills and the influence it casts on communication in the target language.* (Master's Dissertation, University of Glasgow, Scotland). Retrieved from https://www.researchgate.net/publication/293263722_Investigation_of_the_factors_that_cause_language_anxiety_for_ESLEFL_learners_in_learning_speaking_skills_and_the_influence_it_casts_on_communication_in_the_target_language#fullTextFileContent.

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in second language acquisition, 41*, 363–387.

Thrasher, T. (2022). The impact of virtual reality on L2 French learners' language anxiety and oral comprehensibility: An exploratory study. *CALICO journal, 39*(2), 219-238.

Tien, C.-Y. (2018). English speaking anxiety in EFL university classrooms in Taiwan. *European Journal of English Language Teaching, 4*(2), 21-34.

Timina, S. (2015). *Causes of English speaking anxiety among Taiwanese university students.* Proceedings of INTCESS15-2nd international conference on education and social sciences, 1305-1309

Tobias, S. (1979). Anxiety research in educational psychology. *Journal of Educational Psychology, 71*(5), 573–582.

Tolin, D. F. (2019). Inhibitory learning for anxiety-related disorders. *Cognitive and Behavioral Practice, 26*(1), 225-236.

Toyama, M., & Yamazaki, Y. (2021). Anxiety reduction sessions in foreign language classrooms. *The Language Learning Journal, 49*(3), 330-342.

Trickett, E. J., & Moos, R. H. (1995). *Classroom environment scale manual: Development, applications, research*. Consulting Psychologists Press.

Trylong, V. L. (1988). *Aptitude, attitudes, and anxiety: A study of their relationships to achievement in the foreign language classroom.* (Doctoral Dissertation, Purdue University, United States). Retrieved from https://www.proquest.com/docview/303617466?pq-origsite=gscholar&fromopenview=true.

Tsao, J. C., & Craske, M. G. (2000). Timing of treatment and return of fear: Effects of massed, uniform-, and expanding-spaced exposure schedules. *Behavior therapy, 31*(3), 479-497.

Tschirner, E., Bärenfänger, O., & Wanner, I. (2012). *Assessing evidence of validity of assigning CEFR ratings to the ACTFL oral proficiency interview (OPI) and the oral proficiency interview by computer (OPIc)*. (Technical Report 2012-US-PUB-1). University of Leipzig. Retrieved from: https://www.languagetesting.com/pub/media/wysiwyg/research/OPIc-CEFR-Study-Final-Report.pdf.

Tseng, M.-f., & Geng, Z. (2021). Implementing Virtual Reality–Enhanced Tasks in Chinese Language Teaching. In Y.-J. Lan & S. Grant (Eds.), *Contextual Language Learning: Real Language Learning on the Continuum from Virtuality to Reality* (pp. 91-118). Springer.

Tsiplakides, I., & Keramida, A. (2009). Helping Students Overcome Foreign Language Speaking Anxiety in the English Classroom: Theoretical Issues and Practical Recommendations. *International Education Studies, 2*(4), 39-44.

Turner, A. (2015). Generation Z: Technology and social interest. *The journal of individual Psychology*, *71*(2), 103-113.

Urueta, S. H., & Ogi, T. (2019). A TEFL virtual reality system for high-presence distance learning. In L. Barolli, H. Nishino, T. Enokido, & M. Takizawa (Eds.), *International Conference on Network-Based Information Systems* (vol. 1036, pp. 359-368). Springer.

Uştuk, Ö., & Aydın, S. (2018). The effects of the use of paralinguistic cues on foreign language anxiety among English as a foreign language speakers. *Innovation in Language Learning and Teaching, 12*(3), 289-302.

Van Der Meij, H., & Zhao, X. (2010). Codeswitching in English courses in Chinese universities. *The Modern Language Journal, 94*(3), 396-411.

Vázquez, C., Xia, L., Aikawa, T., & Maes, P. (2018). Words in Motion: Kinesthetic Language Learning in Virtual Reality. *2018 IEEE 18th International Conference on Advanced Learning Technologies (ICALT)*, 272-276. Retrieved from https://aikawa.mit.edu/sites/default/files/documents/WordsInMotion.pdf.

Velev, D., & Zlateva, P. (2017). Virtual reality challenges in education and training. *International Journal of Learning and Teaching, 3*(1), 33-37.

Waite, S., & Pratt, N. (2015). Situated learning (learning in situ). In J. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences* (2nd ed., Vol. 22, pp. 5-12). Elsevier Ltd.

Walk, R. D. (1956). Self ratings of fear in a fear-invoking situation. *The Journal of abnormal and social psychology, 52*(2), 171-178.

Walz, J. C. (1982). Error Correction Techniques for the Foreign Language Classroom. In *Language in Education: Theory and Practice, No. 50*. Eric. Center for Applied Linguistics.

Wang, C., Lian, X., Zhuang, C., Kwok, P. K., & Yan, M. (2021). *A virtual reality-based spoken English learning platform.* 2021 IEEE 24th International Conference on Computer Supported Cooperative Work in Design (CSCWD), 867-872

Wang, Y. F., Petrina, S., & Feng, F. (2017). VILLAGE—Virtual Immersive Language Learning and Gaming Environment: Immersion and presence. *British journal of educational technology, 48*(2), 431-450.

Wang, Z. (2014). Developing Accuracy and Fluency in Spoken English of Chinese EFL Learners. *English Language Teaching, 7*(2), 110-118.

Wang, Z., Guo, Y., Wang, Y., Tu, Y.-F., & Liu, C. (2021). Technological Solutions for Sustainable Development: Effects of a Visual Prompt Scaffolding-Based Virtual Reality Approach on EFL Learners' Reading Comprehension, Learning Attitude, Motivation, and Anxiety. *Sustainability, 13*(24), 13977.

Wasserman, L. (2006). *All of nonparametric statistics*. Springer Science & Business Media.

Wehner, A. K., Gump, A. W., & Downey, S. (2011). The effects of Second Life on the motivation of undergraduate students learning a foreign language. *Computer Assisted Language Learning, 24*(3), 277-289.

Wei, L., Lin, H. H., & Litton, F. (2018). Communicative language teaching (CLT) in EFL context in Asia. *Asian Culture and History*, *10*(2), 1-9.

Wenger, E. (1999). *Communities of practice: Learning, meaning, and identity*. Cambridge university press.

West, R. F., Meserve, R. J., & Stanovich, K. E. (2012). Cognitive sophistication does not attenuate the bias blind spot. *Journal of personality and social psychology, 103*(3), 506 –519.

White, M. D., & Marsh, E. E. (2006). Content analysis: A flexible methodology. *Library trends, 55*(1), 22-45.

Wiedemann, K. (2001). Anxiety and Anxiety Disorders. In N. J. Smelser & P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp. 560-567). Pergamon.

Wilang, J. D., & Soermphongsuwat, A. (2018). Virtual Reality for Undergraduate English Language Learners: A Formative Study. *Proceedings of ISAGA2018: International Conference*, 191-199. Retrieved from https://files.eric.ed.gov/fulltext/ED585191.pdf.

Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC journal, 37*(3), 308-328.

Woodrow, L. (2014). *Writing about quantitative research in applied linguistics*. Springer.

Xie, Y., Chen, Y., & Ryder, L. H. (2021). Effects of using mobile-based virtual reality on Chinese L2 students' oral proficiency. *Computer Assisted Language Learning, 34*(3), 225-245.

Xie, Y., Ryder, L., & Chen, Y. (2019). Using Interactive Virtual Reality Tools in an Advanced Chinese Language Class: a Case Study. *TechTrends, 63*(3), 251-259.

Yalçın, Ö., & İnceçay, V. (2014). Foreign language speaking anxiety: The case of spontaneous speaking activities. *Procedia-Social and Behavioral Sciences, 116*, 2620-2624.

Yan, J. X., & Horwitz, E. K. (2008). Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. *Language learning, 58*(1), 151-183.

Yang, F.-C. O., Lo, F.-Y. R., Hsieh, J. C., & Wu, W.-C. V. (2020). Facilitating communicative ability of EFL learners via high-immersion virtual reality. *Journal of Educational Technology & Society, 23*(1), 30-49.

Yao, L., Wan, Y., Ni, H., & Xu, B. (2021). Action unit classification for facial expression recognition using active learning and SVM. *Multimedia Tools and Applications, 80*, 24287-24301.

York, J., Shibata, K., Tokutake, H., & Nakayama, H. (2021). Effect of SCMC on foreign language anxiety and learning experience: A comparison of voice, video, and VR-based oral interaction. *ReCALL, 33*(1), 49-70.

Young, D. J. (1990). An investigation of students' perspectives on anxiety and speaking. *Foreign Language Annals, 23*(6), 539-553.

Young, D. J. (1991). Creating a low‐anxiety classroom environment: What does language anxiety research suggest? *The Modern Language Journal, 75*(4), 426-437.

Young, D. J. (1992). Language anxiety from the foreign language specialist's perspective: Interviews with Krashen, Omaggio Hadley, Terrell, and Rardin. *Foreign Language Annals, 25*(2), 157-172.

Yuen, E. K., Goetter, E. M., Stasio, M. J., Ash, P., Mansour, B., McNally, E., Zulaica, K. (2019). A pilot of acceptance and commitment therapy for public speaking anxiety delivered with group videoconferencing and virtual reality exposure. *Journal of Contextual Behavioral Science, 12*, 47-54.

Zeidner, M. (2007). Test anxiety in educational contexts: Concepts, findings, and future directions. In P. A. Schutz & R. Pekrun (Eds.), *Emotion in education* (pp. 165-184). Elsevier.

Zhang, C. & Zhang, W. (2022). The impact of social phobia on willingness to communicate in a second language: The mediation effect of ideal L2 self. *Front Psychol*, *13*, 1018633.

APPENDICES

Appendix 1. English Version of the Questionnaire

Questionnaire

This questionnaire is designed for the present research and aims to help better understand Chinese students' experience in speaking English. All responses will be treated in the way as stated in Participant Information Sheet and Consent Form. The questionnaire may take you 15 minutes. Thank you in advance.

Please identify your gender by ticking appropriate box

☐ A. Female ☐ B. Male ☐ C. Other

Please identify your age by ticking appropriate box

☐ A. 18-19 ☐ B. 20-21 ☐ C. 22-23 ☐ D. Above 23

Section A English Learning and Using Experience

Please tick the box or write down the answer applied to you in the following statements. Only one answer could be selected for each question.

1. When did you start to learn English?

☐ A. The first year at primary school ☐ B. The third year at primary school

☐ C. Middle school ☐ D. High school

☐ E. Higher Education (university) ☐ F. Other (please specify: _____)

2. Have you ever been to foreign countries (English speaking countries)?

☐ A. Yes ☐ B. No

If so, which countries have you been and how long have you been there? Please write down below.

Section B English Speaking Proficiency Self-assessment

Please read the statements carefully and tick the box applied to you. Only one answer could be selected for each question.

1. How do you assess yourself in interacting in English?

☐ A1: I can interact in a simple way provided the other person is prepared to repeat or rephrase things at a slower rate of speech and help me formulate what I'm trying to say. I can ask and answer simple questions in areas of immediate need or on very familiar topics.

☐ A2: I can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar topics and activities. I can handle very short social exchanges, even though I can't usually understand enough to keep the conversation going myself.

☐ B1: I can deal with most situations likely to arise whilst travelling in an area where the language is spoken. I can enter unprepared into conversation on topics that are familiar, of personal interest or pertinent to everyday life (e.g. family, hobbies, work, travel, and current events).

☐ B2: I can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible. I can take an active part in discussion in familiar contexts, accounting for and sustaining my views.

☐ C1: I can express myself fluently and spontaneously without much obvious searching for expressions. I can use language flexibly and effectively for social and professional purposes. I can formulate ideas and opinions with precision and relate my contribution skilfully to those of other speakers.

☐ C2: I can take part effortlessly in any conversation or discussion and have a good familiarity with idiomatic expressions and colloquialisms. I can express myself fluently and convey finer shades of meaning precisely. If I do have a problem I can backtrack and restructure around the difficulty so smoothly that other people are hardly aware of it.

2. How do you assess yourself in expressing in English?

☐ A1: I can use simple phrases and sentences to describe where I live and people I know.

☐ A2: I can use a series of phrases and sentences to describe in simple terms my family and other people, living conditions, my educational background and my present or most recent job.

☐ B1: I can connect phrases in a simple way in order to describe experiences and events, my dreams, hopes and ambitions. I can briefly give reasons and explanations for opinions

and plans. I can narrate a story or relate the plot of a book or film and describe my reactions.

☐ B2: I can present clear, detailed descriptions on a wide range of subjects related to my field of interest. I can explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.

☐ C1: I can present clear, detailed descriptions of complex subjects integrating sub-themes, developing particular points and rounding off with an appropriate conclusion.

☐ C2: I can present a clear, smoothly-flowing description or argument in a style appropriate to the context and with an effective logical structure which helps the recipient to notice and remember significant points.

Section C Technical Experience on Virtual Reality

Please tick the box or write down the answer applied to you in the following statements. Only one answer could be selected for each question.

1. Have you used Virtual Reality technology?

☐ A. No ☐ B. Yes

If your answer is A, please skip the following questions, and go to Section D directly.

2. What do you usually use VR for? (You can select more than one options here)

☐ A. Playing games ☐ B. Watching videos ☐ C. Learning purpose

☐ D. Other purposes (please specify: _____)

Section D English Speaking Anxiety Scale for Chinese English Learners

Please tick the option that best matches your feelings about speaking English. There are no right or wrong options, all depending on your first reaction. The options stand for: (1) Strongly Disagree, (2) Disagree, (3) Neither Agree nor Disagree, (4) Agree, or (5) Strongly Agree.

No.	Statement	1	2	3	4	5
1	I never feel quite sure of myself when I am pronouncing English words while speaking with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

2	It bothers me when I cannot speak English with others because of grammar.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
3	I start to panic when I am not sure of saying something that makes sense in English, while speaking English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
4	It bothers me when I cannot speak English with others with a standard accent.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
5	I become anxious when I get stuck on one or two words in speaking English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
6	It frightens me when I do not understand what others are saying in English.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
7	While speaking English with others, I feel intimidated when I translate word by word the expressions from my first language into English.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
8	When I want to use correct grammar to speak English with others, I get so nervous that I forget what to say.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
9	I always feel that the other students speak English better than I do.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
10	I am worried about making mistakes when I am speaking English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
11	I am afraid that the others will laugh at me when I speak English.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
12	I feel afraid when others do not understand what I mean in English.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
13	Even if I am well prepared for the speaking tasks I am about to do, I feel anxious about it.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
14	I feel very self-conscious about speaking English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
15	I am afraid that my spoken English is not good enough to complete the interaction tasks.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

16	I worry about losing face when I am speaking English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
17	I feel embarrassed to speak English with others because I think I have poor pronunciation and intonation.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
18	I get upset when I have to express complicated ideas while speaking English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
19	I feel more tense and nervous in speaking English rather than reading and writing English.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
20	I feel pressured when I use complex grammatical structures while speaking English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
21	I am nervous if I have to speak English with others when I am not familiar with the topic.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
22	I would be nervous speaking English with native speakers.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
23	While speaking English with others, I can generally get so nervous that I forget things I know.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
24	I start to panic when I have to speak English with others without preparation for the interaction task.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
25	I get worried when I have little time to think about what I have to speak in English with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
26	I get anxious when I find I cannot speak English fluently with others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
27	Generally, I feel nervous when I have to speak to others.	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

Appendix 2. Chinese Version of the Questionnaire

调查问卷

该调查问卷为本次研究而设计，旨在帮助更好地理解中国学生的英语口语情况。您所有的回答都将按照参与者信息表和同意书中描述的方式处理。完成该问卷大约需要 15 分钟。非常感谢您的参与！

您的性别

☐ 女　　☐ 男　　☐其他

您的年龄

☐ 18-19　　☐ 20-21　　☐ 22-23　　☐ 23 岁以上

第一部分：英语学习和使用情况

请根据您自己的实际情况勾选或作答 。每题只能选一个答案。

1. 您什么时候开始学英语？

☐ A. 小学一年级　☐ B. 小学三年级　☐ C. 初中

☐ D. 高中　☐ E. 大学　☐ F. 其他（请详细说明：_____）

2. 您去过需要用英语交流的国家吗？

☐ A. 去过　☐ B. 没去过

如果去过，您去的是哪些国家，在那里呆了多久？请写在下面横线上。

第二部分：英语口语水平的自我评估

请仔细阅读并根据自己的情况勾选。每题只能选一个答案。

3. 您如何评价自己的英语交流水平？

☐ A1: 如果对方随时可以重复或者说得更慢并且帮助我组织语言，我可以用英语进行简单的交际（如互致问候、相互介绍等）或谈论非常熟悉的话题。

☐ A2: 我可以就所熟悉的话题或活动进行简单直接的信息交流。即使无法理解全部话语，我仍能应付非常简单的交际任务。

☐ B1: 我可以应对大部分可能在说英语的地区旅行时出现的情况。我可以在毫无准备的情况下，就熟悉的，个人感兴趣的，或者日常的话题（如家庭，爱好，工作，旅行和时事）进行交流。

☐ B2: 我可以在一定程度上随时流利地对话，使得与英语为母语者进行交流成为可能。我可以主动加入熟悉情境的讨论，阐述并支撑我的观点。

☐ C1: 我可以流利自如地表达自己并且灵活有效地运用英语进行社交和工作。同时，自然地与人交流我的观点和想法。

☐ C2: 我可以毫不费力地加入任何英语对话或者讨论，并且熟知各类地道的英语表达。同时，准确地表达自己的观点。如果出现什么问题，我可以及时调整而不让他人察觉。

4. 您如何评价自己的英文表达能力？

☐ A1: 我可以用简单的短语和句子描述我生活的地方和认识的人。

☐ A2: 我可以用一系列短语和句子简单地描述我的家人和其他人，生活条件，教育背景和我当下的工作。

☐ B1: 我可以简单地使用短语来描述经历和事件，谈论我的梦想和愿望，给出意见和计划的原因和解释。我可以描述故事，讲述书上或者电影中的情节，并谈论我的感受。

☐ B2: 我可以清晰详细地阐述大量我感兴趣的话题， 并且就某个话题阐明自己的观点同时分析其利弊。

☐ C1: 我可以清晰详细地阐述复杂的话题，整合分论点，进行阐释，得出合适的结论。

☐ C2: 我可以用适合当前情境的方式清晰流畅且符合逻辑地描述或者论证，以帮助听者注意和记住重要的部分。

第三部分 虚拟现实技术的使用情况

请根据自己的实际情况勾选或作答。

1. 您使用过虚拟现实技术嘛？

☐ A. 没用过　　☐ B.用过（请说明几次：_____）

如果您的回答是 A 没有，请跳过下面的问题，直接回答第四部分。

2. 您通常用虚拟现实技术做什么？（可多选）

☐ A. 玩游戏 ☐ B. 看视频 ☐ C. 学习 ☐ D. 其他目的（请详细说明：————）

第四部分：中国英语学习者英语口语焦虑量表

请勾选最符合您和别人说英语时的感受的选项。这里没有对错之分，请根据您的第一感觉进行选择。选项代表的是：1-非常不符合，2-不符合，3-不确定，4-符合，5-非常符合。

序号	情况描述	1	2	3	4	5
1	当我和别人说英语时，单词的发音问题会让我很没把握。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
2	当我和别人说英语时，语法问题会让我感到困扰。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
3	当我和别人说英语的时候，不确定我说的是否符合英语表达习惯时，会让我感到慌张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
4	当我和别人说英语时，口音问题会让我感到困扰。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
5	当我和别人说英语时，被一两个单词卡住，会让我感到焦虑。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
6	当我不能理解别人说的英语时，会让我感到慌张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
7	当我通过逐字翻译母语的方式来和别人说英语时，会让我感到慌张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
8	当我想用正确的语法和别人说英语时，我会紧张得忘了自己要说什么。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
9	我总觉得其他同学的英语口语比我好。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
10	我和别人说英语的时候很担心犯错。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
11	我很害怕说英语的时候其他人会嘲笑我。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

12	当别人不理解我说的英语的意思时，我很担心。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
13	就算我充分准备了将要进行的口语任务，和别人说英语依然让我感到焦虑。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
14	和别人说英语会让我感到很不自在。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
15	我担心我的英语口语不好，不足以和别人完成口语交流任务。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
16	我担心和别人说英语的时候会丢面子。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
17	我会不好意思和别人说英语，因为我觉得我的发音和语调很差。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
18	当我需要用英语向别人表达复杂的想法时，我感到很不安。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
19	我在说英语的时候比在读写英语的时候更容易感到紧张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
20	当我和别人说英语时，使用复杂的语法结构会让我感到压力很大。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
21	如果对话题不熟悉，我和别人说英语的时候会感到紧张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
22	和以英语为母语的人说英语的时候我会紧张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
23	当我和别人说英语时，我一般会紧张得忘记我明明知道的东西。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
24	当我对交流任务毫无准备时，需要说英语会让我很慌张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
25	当几乎没有时间思考的情况下开口和别人说英语会让我很担心。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
26	当我发现我不能流利地和别人说英语时，我感觉很焦虑。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐
27	一般来说，我和别人说话就会感到紧张。	1 ☐	2 ☐	3 ☐	4 ☐	5 ☐

Appendix 3. Role-play Prompts for the Speaking Test

1 You will play the role of a ticket seller in a British town. I will play the role of a tourist who wants to buy tickets for a special event. I will call you and ask questions about the event and the tickets.	2 You will play the role of my friend. While in a British town, I need to do some shopping, so I go to the city's largest mall. At the entrance I run into you. You have been to this mall many times. I will ask several questions to get information to help the shopping trip.
3 You will play the role of a customer representative. I want to order a number of clothing items from the online catalogue you work for, but my internet is not working. I will call your company to find out the information I need to order the merchandise on the phone.	4 You will play the role of my friend. While in a British town, I attend an annual music festival. At the festival, I run into you. You have been to this festival many times before. I will ask you several questions to find out everything I need to know to make my day at the festival fun and enjoyable.
5 You will play the role of an estate agent. I want to move to a British town. I will call you to get information about possible properties to rent.	6 You will play the role of my son's teacher in a British school. I am going to meet you to discuss my son's future at the school. I will ask you questions about the school, and you will ask me questions to find out about my son.
7 You will play the role of an employee in a visitor information centre in a British town. I am visiting the town for the first time. I will ask you several questions to get all of the information I need to make my visit a success.	8 You will play the role of my friend. You have just got a new cell phone. I also need a new phone and like the look of yours. I will ask you several questions to get more information about your new phone.
9	10

You are taking a train, and you will play the role of the person sitting next to me on the train. I notice you are reading a new book. The title is one that I am also interested in. We are going to introduce ourselves, and then I will ask you several questions to find out more about the book.	You will play the role of my neighbour. I meet you in the elevator, and we introduce yourselves. Then I will ask you several questions about living in the apartment building, the neighbourhood, and the city in general.

Appendix 4. English Version of the Scale of Self-rating of FLSA Levels in Specific Tasks

Self-rating scale of the FLSA levels in the speaking test

Please watch the video recordings of your performance in the speaking test. Stop and rate your English speaking anxiety levels from 1 to 10 by each 30 seconds with the first four minutes of the role-play conversation. 1 refers to not anxious at all, 10 refers to very anxious. The larger the number, the more anxious it means. Please rate your English speaking anxiety levels depending on your own feelings during the speaking test. Thank you!

Time duration	Anxiety level
0:00-0:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
0:30-1:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
1:00-1:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
1:30-2:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
2:00-2:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
2:30-3:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
3:00-3:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
3:30-4:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐

Appendix 5. Chinese Version of the Scale of Self-rating of FLSA Levels in Specific Tasks

口语测试焦虑程度自我评分表

请认真观看自己的口语测试视频，就角色扮演对话的前 4 分钟，每半分钟在 1-10 区间对自己的焦虑程度进行一次评分。1 为不焦虑，10 为非常焦虑。数字越大，焦虑程度越高。请每次评分的时候暂停视频，并且按照你自己的感受如实评分。谢谢！

对话时间段	焦虑程度
0:00-0:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
0:30-1:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
1:00-1:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
1:30-2:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
2:00-2:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
2:30-3:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
3:00-3:30	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐
3:30-4:00	1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6 ☐ 7 ☐ 8 ☐ 9 ☐ 10 ☐

Appendix 6. Chinese Version of the Open-ended Questionnaire

第一部分：口语焦虑回顾

请回答以下问题：

1. 请描述你在参加课程之前和别人说英语的焦虑情况。
2. 请描述你在参加课程之后和别人说英语的焦虑情况。
3. 实际上，在第二次口语测试中，你是否比第一次感到更轻松？是什么让你有这样的感受呢？

第二部分：虚拟现实与口语焦虑

请回答以下问题：

1. 你觉得沉浸式 VR 对于你现在和别人说英语的焦虑情况有什么影响？
2. 请你谈谈沉浸式 VR 用于英语口语课程的利与弊。

Appendix 7. English Version of Participant Information Sheet

Participant Information Sheet

The Impact of Immersive Virtual Reality on Foreign Language Speaking Anxiety: A Study of EFL Learners in China

Miaomiao Ding School of Education

You are being invited to take part in a research study on English Language Learning in Higher Education. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Thank you for reading this.

The aim of the study is to investigate the effects of immersive Virtual Reality (VR) on foreign language speaking anxiety in English learning in China. Immersive VR (see Figure 1) is a new technology, in which you can experience virtual 'real-life' situations. For the research, you will be asked to attend twelve sessions of an English course, and it will be twice a week. You will be asked to complete a questionnaire twice during the English course. That is before and after the English course. Each questionnaire will take you five minutes. English speaking tests will be conducted after questionnaire completion each time. Your responses in the speaking tests will be video recorded. You will be asked to complete another scale by watching the video recordings after each speaking test, which will take you five minutes. At last, you will be asked to fill in an open-ended questionnaire, which will take you ten minutes.

Figure 1. Immersive VR technology (HTC Vive Focus Plus) used for this study

You will be divided into four groups in the experiment, in which some of you may use VR, some may not. You can experience VR courses after the experiment if you are divided into non-VR group.

The courses, questionnaires, speaking and scales will all be taken place in classrooms/offices at your institute. All the recordings will be taken under your permission. Your participation

is voluntary, and you can withdraw at any time without giving a reason. Your performance in the study will not influence your grade in the university.

Your personal details will be kept confidential and only myself and my supervisors will have access to them. I will anonymise everything you tell me when I present the data in my thesis and possible publications. Please note that assurances on confidentiality will be strictly adhered to unless evidence of wrongdoing or potential harm is uncovered. In such cases, the University may be obliged to contact relevant statutory bodies/agencies.

The data I collect will only be used for the researcher's thesis and related publications. The questionnaires and electronic recordings will be stored securely on a password-protected and encrypted computer and mobile hard disk for backup. All the collected data will be destroyed after the submission of the thesis.

The project has been considered and approved by the College Research Ethics Committee

If you are looking for further information, please contact the researcher through email: m.ding.1@research.gla.ac.uk. You can also contact the supervisor Professor Ellen Boeren through email: ellen.boeren@glasgow.ac.uk and Dr Gabriella Rodolico through email: Gabriella.Rodolico@glasgow.ac.uk **. If you would like to pursue any complaint, please contact the College of Social Sciences Ethics Officer, Dr Muir Houston, email: Muir.Houston@glasgow.ac.uk**

Appendix 8. Chinese Version of Participant Information Sheet

邀 请 信

沉浸式虚拟现实技术对于外语口语交流中语言焦虑的影响：中国英语学习者的案例研究

英国格拉斯哥大学教育学院教育学博士在读学生　丁苗苗

亲爱的同学：

你好！

你被邀请参与一个关于大学生英语学习的研究。在做出决定之前，理解开展研究的原因以及其中包含什么项目对你来说很重要。请你仔细阅读以下信息，如果需要，也可以和他人讨论。如果有任何不清楚的地方，也可以直接询问我。谢谢！

这项研究的目的是调查沉浸式虚拟现实技术对于中国的英语学习者口语交流中语言焦虑的影响。沉浸式虚拟现实技术（图 1 是本研究将使用的设备）是一项新兴的技术，你可以在其中体验虚拟的和现实世界很像的场景。在本次研究中，你预计将被邀请参加 12 节英语口语课程，每周 2 次。你预计将被邀请填写调查问卷 2 次，填写时间分别是第 1 次课前以及第 12 次课后。每次填写问卷需要大概 **5 分钟**。同时，你会被邀请进行 2 次口语测试（每次 **8 分钟**左右），在每次填写完调查问卷后立即进行。你在口语测试中的表现会被视频录制，并且在口语测试结束后，你需要通过看自己的视频为自己的交流程度打分。在研究的最后，你需要填写一份开放性问题的问卷，大概需要花费你 10 分钟的时间。

图 1. 本研究中使用的沉浸式虚拟现实头显

在实验过程中，你们将被分为 4 个小组。有虚拟现实创新课程组也有传统教学组。无论你被分到什么组，在实验结束后，你都有机会体验虚拟现实技术提供的创新课程！

所有的课程，问卷调查，采访都会在校内的教室或者办公室进行。所有的录音都会在你的允许下进行。你的参与完全是自愿的，而且中途你可以无理由退出。你在本研究中的所有表现不会影响你的在校成绩.

你的个人信息会被保密，只有我和我的英国的导师可以查看。在我的毕业论文和可能的文章发表中，所有的信息都会以匿名的形式发布。请注意，保密性原则将被严格执行。一旦有任何问题出现，格拉斯哥大学会采取相应措施追责。

我所收集到的数据将只用于我的毕业论文和可能的文章发表中。所有的数据将被密码保护在电脑和移动硬盘里，并且能够追踪到个人信息的数据将在我完成论文后销毁。

本项研究是得到格拉斯哥大学社会科学学院道德伦理委员会批准的。

如果你想要了解更多的信息，请通过 QQ（微信同号）：908794528 ，或者邮箱：m.ding.1@research.gla.ac.uk 联系我。 你也可以联系我的导师 Ellen Boeren 教授或者 Gabriella Rodolico 博士，她们的邮箱分别是：ellen.boeren@glasgow.ac.uk 和 Gabriella.Rodolico@glasgow.ac.uk。如果你有任何的抱怨，你也可以联系社会科学学院道德伦理办公室的 Muir Houston 博士，他的邮箱的 Muir.Houston@glasgow.ac.uk。

Appendix 9. English Version of the Consent Form

Consent Form

Title of Project: The Effects of Immersive Virtual Reality on Foreign Language Speaking Anxiety: A Case Study of EFL Learners in China

Name of Researcher: Miaomiao Ding

I have read and understood the Participant Information Sheet for the above study and have had the opportunity to ask questions.	Yes□	No □
I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason.	Yes□	No □
I understand that I will attend a twelve-session English course.	Yes□	No □
I understand that I will complete a questionnaire twice, which is before and after the English course.	Yes□	No □
I understand that I will join English speaking tests (video-recorded) twice, which is before and after the English course.		
I understand that I will complete a scale through watching the English speaking tests recordings after each test.	Yes□	No □
I understand that I need to complete an open-ended questionnaire at the end of the study.		
I acknowledge that participants will be referred to by pseudonym.	Yes□	No □
I acknowledge that there will be no effect on my grades arising from my participation or non-participation in this research.	Yes□	No □
I acknowledge that all names and other material likely to identify individuals will be anonymised.	Yes□	No □
I acknowledge that the materials the researched collected will be treated as confidential and kept in secure storage at all times.	Yes□	No□
I acknowledge that the materials will be used in publications, both print and online, and they will be destroyed once the project is complete.	Yes□	No □
I agree to waive my copyright to any data collected as part of this project.	Yes□	No □
I agree to take part in this research study.	Yes□	No □

Name of Participant …… Signature ……

Date ……

Name of Researcher …….. Signature ……

Date ……

Appendix 10. Chinese Version of the Consent Form

同 意 书

研究课题：

沉浸式虚拟现实技术对于外语口语交流中语言焦虑的影响：中国英语学习者的案例研究

研究人员：

英国格拉斯哥大学教育学院教育学博士在读学生　丁苗苗

我通读了邀请信并且知道我有机会询问任何问题。	是 □	否 □
我知道我的参与是自愿的并且任何时候都可以无理由退出。	是 □	否 □
我知道我预计将参加 12 节英语口语课程。	是 □	否 □
我知道我预计将在口语课程前后填写 2 次调查问卷。	是 □	否 □
我知道我预计将在口语课程前后参与 2 次口语测试，并且会被视频录制。		
我知道我将在每次口语测试后通过看我的表现视频完成一份量表填写。	是 □	否 □
我知道在研究的最后我需要填写一份开放性的调查问卷。		
我知道所有的参与者都将被编号。	是 □	否 □
我知道我在本次研究中的表现不会对我的在校成绩有任何影响。	是 □	否 □
我知道所有的个人信息和所以其他有可能追踪到我个人的材料都将被保密，并且匿名化。	是 □	否 □
我知道研究人员收集的所有信息都将被保密并且储存在安全的地方。	是 □	否 □
我知道研究人员收集的所有个人信息在项目完成后会被销毁。	是 □	否 □
我同意放弃我对任何本研究收集到的数据的所有权。	是 □	否 □
我同意参加本项研究。	是 □	否 □

学生姓名 …… 日期 ……

研究者姓名 …… 日期 ……

Appendix 11. Factors Extracted from the ESAS Data

Items	F1[35]	F2[36]	F3[37]
21. I am nervous if I have to speak English with others when I am not familiar with the topic	.81		
26. I get anxious when I find I cannot speak English fluently with others.	.73		
25. I get worried when I have little time to think about what I have to speak in English with others.	.72		
18. I get upset when I have to express complicated ideas while speaking English with others.	.71		
24. I start to panic when I have to speak English with others without preparation for the interaction task.	.67		
12. I feel afraid when others do not understand what I mean in English.	.60		
5. I become anxious when I get stuck on one or two words in speaking English with others.	.60		
19. I feel more tense and nervous in speaking English rather than reading and writing English.	.59		
22. I would be nervous speaking English with native speakers.	.59		
7. While speaking English with others, I feel intimidated when I translate word by word the expressions from my first language into English.	.51		
3. I start to panic when I am not sure of saying something that makes sense in English, while speaking English with others.	.32		
23. While speaking English with others, I can generally get so nervous that I forget things I know.	.34		
9. I always feel that the other students speak English better than I do.	.32		
15. I am afraid that my spoken English is not good enough to complete the interaction tasks.		.35	

[35] F1: low proficiency
[36] F2: low self-confidence
[37] F3: pronunciation difficulty

20. I feel pressured when I use complex grammatical structures while speaking English with others.	.33	
10. I am worried about making mistakes when I am speaking English with others.	.45	
16. I worry about losing face when I am speaking English with others.	.70	
2. It bothers me when I cannot speak English with others because of grammar.	.69	
8. When I want to use correct grammar to speak English with others, I get so nervous that I forget what to say.	.67	
13. Even if I am well prepared for the speaking tasks I am about to do, I feel anxious about it.	.43	
14. I feel very self-conscious about speaking English with others.	.45	
11. I am afraid that the others will laugh at me when I speak English.	.47	
1. I never feel quite sure of myself when I am pronouncing English words while speaking with others.		.77
17. I feel embarrassed to speak English with others because I think I have poor pronunciation and intonation.		.76
4. It bothers me when I cannot speak English with others with a standard accent.		.71