Learning Business English in Virtual Worlds: Effectiveness and Acceptance in a Malaysian Context

A thesis submitted in partial fulfilment of

the requirements for the degree

of

Doctor of Philosophy

in

Management Information Systems

at

Massey University

Palmerston North

Harmi Izzuan Bin Baharum

2013

Abstract

Motivated by the need to provide better opportunities for Malays in Malaysia to improve their oral business English communication skills, the research focused on the use of multiuser virtual environment (MUVE) for learning English and pursued the following research questions: (1) Is MUVE based learning effective in facilitating situated scenario-based learning of oral business English communication skills by Malay learners? (2) Which factors influence the acceptance of MUVE based learning of oral business English communication skills by Malay learners? To address the first research question, a controlled experiment was conducted to compare the learning gains in traditional classroom and in MUVE environment. To address the second research question, an acceptance model based on the social cognitive theory and the technology acceptance model was tested by fitting it to the data obtained by using a questionnaire. The experiment involved 152 Malay tertiary learners, who also filled in the questionnaire. Findings indicated that MUVE was effective in facilitating scenario-based learning of business English by Malay learners. Learners' oral skills showed statistically significant improvement following learning in MUVE. However, the difference between the improvement in the classroom environment and in MUVE was not statistically significant. As for MUVE acceptance, Video Games Affect, English Class Anxiety, and Perceived Usefulness affected the learners' Intention to Use MUVE for e-Learning, although the effect size for Video Games Affect and English Class Anxiety was small. The results of the study suggest that MUVE based learning is an effective environment for learning oral business English communication skills. MUVE is particularly suitable for distance learning, when traditional classroom learning is not available. The study confirmed the claims in the literature that MUVE is particularly suitable for anxious learners and for learners who like to play video games. The study involved Malay university students as participants, and the results are not necessarily generalizable to other types of learners.

Acknowledgement

Studying for a PhD is a challenging journey. It would not be possible to finish without the help, support, guidance, and encouragement from a number of individuals.

First and foremost, I would like to express my deepest gratitude and thanks to my supervisors, Dr Alexei Tretiakov and Dr Barbara Crump from School of Management, Massey University. Moreover, I am also indebted to my other supervisors, Dr Penny Haworth (Massey University), Dr Bill Anderson (Otago University, New Zealand), Assoc Prof Dr Supyan Hussin (National University of Malaysia, Malaysia), and Prof Dr Kinshuk (Athabasca University, Canada). I am grateful for their advice, guidance, and encouragement throughout my PhD journey.

I am grateful to the University of Technology, Malaysia, and to the Public Services Department, Malaysia, for granting me study leave and scholarships for pursuing PhD studies. Further, I thank my colleagues who have taken over my job responsibilities during my absence from work.

Finally, I would like to extend my deepest gratitude and love to my family for being the source of strength and inspiration. Not forgetting my friends and the McQuinlan's family for their continuous support to me.

Table of Contents

List of Figures

List of Tables

Chapter 1: Introduction

1.1 Background of The Study

Recently, electronic learning (e-Learning) practitioners and researchers have become increasingly interested in multi-user virtual environments (MUVEs) learning platforms. However, there is little empirical research devoted to MUVE effectiveness and acceptance, even though both MUVE effectiveness and MUVE acceptance are critical if MUVEs are to eventually succeed as learning environments in mainstream e-Learning practice. The research reported in this thesis relates to MUVE effectiveness and acceptance in the context of promoting the learning of oral business English communication skills of Malay learners.

In the rest of this introduction chapter, the background of the problem (the current state of English teaching in Malaysia and the need to provide better opportunities for Malay learners to acquire oral business English communication skills) is addressed, followed by a statement of the problem and the research questions. The research questions focus on (1) assessing the effectiveness of MUVE in promoting the learning of oral Business English communications skills by Malay English learners and (2) identifying the factors that determine the acceptance of MUVE e-Learning by Malay learners. Then, the chapter introduces the overall theoretical framework and a summary of the methods employed in this study. The study is based on the theory of situated and scenario-based learning, social cognitive theory, and the technology acceptance model. Then, the chapter discusses the significance of the study and its scope. The chapter is concluded with a roadmap for the rest of the thesis.

1.1.1 Virtual Worlds as Learning Environments

Web 2.0 is the term given to describe the second generation of the World Wide Web (O'Reilly, 2007). Web 1.0 was primarily uni-directional. It focused on allowing users to view the contents of sites maintained by large organizations. In other words, users were primarily acting as consumers of content. On the other hand, Web 2.0, is multi-directional. It does not restrict the users to viewing the content but allows users to collaborate and to share content with each other. In other words, Web 2.0 encourages

more interactivity and thus supports group interaction and fosters a greater sense of community among the users (Boulos, Hetherington, & Wheeler, 2007).

A MUVE is a computer simulated environment that users can explore, and where users can interact, collaborate, and share information. MUVEs allow users to interact simultaneously with a rich representation of both a physical and a social environment. The social environment is formed by other users, who are represented in MUVE by their avatars. MUVEs enable social interaction among the users by allowing rich communication via channels such as sound, gesture, text messages, and via sharing of user-generated content (Jin, Wen, & Gough, 2010). MUVEs offer a high degree of flexibility in shaping both the physical and the social environment, as users can set up and change not only the physical environment, but also the appearance of their avatars, while also controlling who has access to the MUVE-based experiences they create. MUVE-based experiences can be shaped as learning experiences, leading to MUVE attracting interest as an e-Learning platform (Bronack, Cheney, Riedl, & Tasher, 2008; Eschenbrenner, Nah, & Siau, 2008; Minocha & Reeves, 2010). MUVEs are virtual reality systems (computer systems that simulate physical presence); however, not all virtual reality systems are MUVEs, because MUVEs offer both physical and social environment, and virtual reality systems may offer just the simulation of a physical environment (Burdea & Coiffet, 2003).

MUVEs have unique capabilities that distinguish them from other virtual environments used for e-Learning. One of the capabilities is to provide opportunities for collaborative learning activities that are highly dangerous to conduct, hard to simulate, or just not possible in the real world (Good, Howland, & Thackray, 2008). For example, one could perform chemistry experiments without coming into contact with hazardous chemicals, experience weightlessness in a virtual outer space laboratory, or take a virtual tour of a no longer existing ancient city. One can also create virtual representations of existing physical scenarios and enact crisis situations (such as pandemics) to explore them in a safe environment (Good et al., 2008). Currently, MUVEs are not regarded as mainstream e-Learning environments, but there is a steady growth in MUVE adoption (Dalgarno, Lee, Carlson, Gregory, & Tynan, 2011; Mikropoulos & Natsis, 2011).

There are many experience reports and conceptual papers about the use of MUVE for learning. However, so far there is relatively little quantitative empirical research of MUVE-based learning (Hew & Cheung, 2010). In particular, the existing studies involving controlled experiments or quasi experiments (reviewed in section 2.11) are limited in scope, and none of the studies addressed the effectiveness of MUVE as an environment for language learning. Most of the existing studies of MUVE acceptance (reviewed in section 2.13) were conducted in contexts other than e-Learning, and the only study conducted in the context of e-Learning (Shen & Eder, 2009; reviewed in section 2.13.4) assessed the factors affecting the acceptance of MUVE in context of learning Management Information Systems, a subject that is very different from second language learning.

1.1.2 The Role of the English Language in Business Communication

The English language is an important language in the world due to the impact of global networks such as the Internet and the need to have a shared linguistic medium in the globalisation era. Globalisation and the advent of e-business and e-commerce have reinforced the status and use of English as the lingua franca in international business communication (Seidlhofer, 2004). McKay (2002) pointed out that the use of English is widely accepted in the global community, especially in economics, politics, and in intellectual areas such as science and technology. Because of the use of English in these areas, countries that wish to participate in the global community are making sure that their citizens are not left behind in terms of proficiency in the English language.

According to Kam (2002) and Butler (2004), numerous Asian countries, including Malaysia, are adopting special measures to promote English language skills among their population. The ruling circles in Malaysia acknowledge the importance of the English language, knowing that employers, particularly the multinational firms, are interested in employing graduates with both professional knowledge and English language skills (David & Govindasamy, 2005).

1.1.3 Malay Learners of English

The present study was motivated by the need to improve business English communication proficiency of Malays in Malaysia. The study, however, was not intended to solve the problem of improving business English communication proficiency of Malays or to suggest a solution directly implementable in practice. Nonetheless, it was hoped that the theoretical insights gained in the present study would contribute to the eventual solution of the problem.

Malaysia is a multi-racial, multi-cultural, and multi-religious country with population consisting of 54.6% Malay, 24.8% Chinese, 7.4% Indians, and 13.2% of other ethnic minorities (Economic Planning Unit, 2012). Formerly a British colony, Malaysia gained independence from the British rule in 1957. After the British colonisation ended, the Malay language (Bahasa Malaysia) became the national language and is used as the language of administration. Schools and universities eventually transitioned from English to Malay as a medium of instruction. Most of the rural schools did not have to make a transition as they used Malay even under colonial rule. English remains as a second language and is taught as a compulsory subject in schools (Kim, 2003). The change from English to Malay as the medium of instruction has offered a broader population of learners access to higher education, but at the same time has also led to a decline in the standard of English (Gill, 1993; Peter, 2005).

In Malaysia, both Malay and English are used for communication among representatives of different ethnic groups. The quality of the English language spoken by Malays is often poor, and often poorer than among the Malaysian Chinese and Indians (Pillay, 1998, Berhannudin, Ahmad, Asri, Hussian, & Khairul, 2010). As a result of that, many Malay graduates find it difficult to gain employment, especially in information technology, engineering, science, and business sectors (Nadzri, 2005; Kassim & Ali, 2010). Although Malays comprise only about 55% of the population in Malaysia, David and Govindasamy (2005) found that 94% of unemployed graduates in the country were Malays. The inability to master the English language, and, in particular, lack of competence in using the English language for oral business communication contribute to the high unemployment rate

among Malay undergraduates (Maarof, Osman, Yamat, & Yunus, 2003; David & Govindasamy, 2005).

A number of factors contribute to the low English proficiency among Malays. These factors include lack of exposure to English language and attitudes that hinder English language learning.

1.1.3.1 Lack of Exposure

In Malaysia, even though English is widely used in the private sector, Malay is the dominant language in the civil service sector (such as in government administration offices) where many Malays are employed (Ting, 2003). As a consequence, Malays lack exposure to English language compared to other ethnicities in Malaysia. By enrolling in English classes conducted face to face, one can get some exposure to communication in English; distance learners, however, remain disadvantaged (White, 2003).

1.1.3.2 Negative Attitude Towards English because of the Colonial Past

The Malay attitude towards learning the English language is sometimes clouded by their resentment towards the colonisation in the past. Hostility towards the language is due to perceptions that they are showing off or are too Westernised if they speak mostly in English (Ming, 2004). This shying-away-from-speaking-English attitude, formed because of negative peer pressure, is preventing Malays from attaining progress in oral English communication even when they are attending English classes. Ming, Ling and Jaafar (2011) found in a survey of Malay university students that the students' motivation to learn English was primarily extrinsic, driven by future career considerations, and that the students lacked confidence and enthusiasm towards learning English.

1.1.3.3 Cultural Features Hindering Engagement.

Using Hofstede's cultural dimensions as a comparison tool (Hofstede, 1991), Malaysia is considerably higher than New Zealand in Power Distance (22 for New Zealand and 104 for Malaysia) and considerably lower than New Zealand in Individualism (79 for New Zealand and 26 for Malaysia). In comparison, the United

States (Power Distance 40 and Individualism 91) and Australia (Power Distance 36 and Individualism 90) have much closer values to New Zealand.

Large power distance suggests that in the classroom Malay learners may expect the teacher to control learning, delegating the learners to following the teacher's instructions, rather than actively contributing, as suggested by state of the art learning paradigms, such as communicative approach to language learning (discussed in section 2.3) or by social constructivist learning (discussed in section 2.9.1). As the result, even though communicative approach to language learning has been promoted by the authorities (Malaysian Education Ministry, 1989; Darus, 2010), it may be difficult to implement in practice. Indeed, in practice schools in Malaysia continue to rely on grammar translation and audio lingual approaches to language teaching (Pandian, 2002; Musa, Koo, & Azman, 2012), known to be ineffective in promoting communication skills (Prator & Celce-Murcia, 1979; Brown 2000). Nunan (2003) explicitly highlighted the gap between the official promotion of student-centred communicative approaches to English teaching and the actual practice that tends to be teacher-centred.

Even though direct numerical comparisons between Malay learners and learners in other countries are difficult to obtain, opinions and qualitative results presented in the literature are overall in agreement with a view that Malay learners prefer to be passive in face to face classrooms. In an article describing her experiences teaching English to Malay students, Bahiyah (1992) emphasized their shyness, commenting on students refraining from speaking out in class to preserve social harmony and to keep distance from the teacher. Yong (2010) described Malaysian students as preferring to receive knowledge directly from the teacher and reluctant to ask questions. To learn oral communication skills, learners need to engage in active interactions; according to Zuwati (2006), such interactions tend to lead to anxiety and discomfort among Malay learners.

1.2 Motivation behind Choosing the Problem to Address in this Study

The low level of oral business English communication competency disadvantages Malays in Malaysia economically. As discussed in section 1.1.3, lack of exposure, personal traits, such as shyness, and peer pressure arising from stigma attached to

English language due to the colonial past impede Malays from achieving proficiency in oral business English communication. Traditional learning materials in the form of textbooks, worksheets, exercise books, or audio and video recordings do not nurture oral communication skills sufficiently, as these resources neither sufficiently support interactivity nor provide a safe environment (Faridah, 2005).

This problem can be addressed by employing an e-learning based approach that is both interactive and safe in terms of intermediating interactions to protect anxious students. In particular, the environment should have the following features:

1. Offer a level of interactivity similar to face to face classroom (or better than face to face classroom) to enable communicative language teaching.
2. Offer intermediation to overcome learners' anxiety resulting from their shyness and from the negative attitudes of the peers with respect to being active in class and with respect to communicating in English.
3. Offer exposure to English communication to distance learners.

As follows from the discussion of MUVE in section 1.1.1, MUVE is an environment that is likely to meet these requirements:

1. MUVE offers a powerful platform for implementing authentic interactive experiences, as it allows learners to engage in rich communication in the context of an authentic (simulated) physical and social setting. When interacting in MUVE, learners can rely on simulated physical artefacts to enrich interactions by providing a suitable context.
2. At the same time, MUVE offers safety because interactions between users are intermediated by avatars. MUVE offers an environment that may be of benefit to students who are anxious about communicating in English, because they do not need to be present in person to communicate. Avatars offer safety through anonymity.
3. Moreover, as an online virtual environment, MUVE offers the same levels of exposure to English communication to distance learners as to learners present on campus.

As seen from the discussion in this section, MUVE e-learning is likely to contribute to addressing the problem of many Malay learners not achieving proficiency in

English needed for successfully entering employment in Malaysian private sector. Nonetheless, before contemplating the introduction of MUVE in real classroom, it is necessary to attain solid empirical evidence suggesting that it is likely to be successful.

At present, such evidence is lacking in two key areas. In particular, even though, as demonstrated in this section, it is possible to provide arguments in favour of MUVE being effective in supporting communicative teaching of English, MUVE effectiveness as an environment for teaching English has never been objectively compared to the effectiveness of face to face teaching (for example, by using a controlled experiment). Even though there are experience reports suggesting that MUVE is effective in promoting language learning (reviewed in section 2.10), the conclusions of such reports are subjective, and because the reports are usually made by innovators who invested themselves in promoting the use of MUVE, such reports may be biased in favour of MUVE.

Moreover, for an information system such as MUVE to be successful, it is not sufficient for it to be effective when used. It is also essential that MUVE is actually used by its intended users (Venkatesh & Davis, 2000; Chang & Tung, 2007), and when MUVE is used in the context of promoting learning, the most important category of users are, clearly, the learners. Therefore, it is essential that factors affecting the adoption of MUVE by Malay learners are understood. Even though several studies have explored the factors affecting MUVE acceptance (reviewed in section 2.13), none of these studies was conducted in the context of using MUVE for second language learning. Moreover, none of the prior studies focused on factors that are of particular relevance to Malay learners, such as anxiety with respect to studying English, and none of the studies was conducted in the Malaysian context.

Therefore, the problem addressed by the present study (formulated in section 1.3) focused on testing MUVE effectiveness by using a rigorous approach ensuring that the outcome of the test for as much as possible is not affected by the subjective point of view of the researcher and on exploring factors affecting MUVE acceptance by learners.

The intent of the present study is not to fully assess the suitability of MUVE at improving communicative English language proficiency of Malays in Malaysia; multiple studies are needed to fully address this, not just in Management Information

Systems field, but, possibly, also in Education, Management, Communications, and Politics. Rather, by addressing research questions pertaining to the field of Management Information System, the present study aims to make a limited, but essential, contribution in key areas.

1.3 Statement of the Problem

Although MUVE is claimed in the literature to provide benefits as an e-learning platform, there is little research-based evidence to back such claims, other than experience reports and qualitative studies relying on practitioner and student opinion, rather than on measurement (see section 1.1.1). In particular, empirical studies directly and objectively comparing the effectiveness of MUVE with traditional face to learning are limited—there is no evidence based on true experiments with random assignment of subjects (or even based on quasi experiments) of the effectiveness of MUVE in supporting scenario-based learning of oral business English communication skills.

More broadly, there are very few studies of MUVE effectiveness as an e-Learning environment employing methods not relying on subjective judgements (methods such as true experiments with random assignment of the participants or quasi-experiments). Only two studies of this kind were identified in the literature review of the present study (for a discussion of these studies, refer to section 2.11). One of the studies (Cobb, Heaney, Corcoran, & Henderson-Begg, 2009) was devoted to teaching microbiology lab skills, and the other one (Wrzesien & Raya, 2010) was devoted to teaching science and ecology. None of the studies addressed the effectiveness of MUVE at facilitating communicative language teaching.

Even though there are a number of studies of factors affecting MUVE acceptance in general (reviewed in 2.13), only one of these studies (the study by Shen & Eder, 2009) was devoted to MUVE acceptance as an e-Learning environment, in the context of teaching Management Information Systems, a topic that is clearly very different from language study. (English as a second language, unlike Management Information Systems, is a non-technical subject; and unlike Management Information Systems, English language mastery depends on the development of skills, rather than on conceptual knowledge.) Moreover, the study by Shen and Eder

(2009) did not explore the possible effect of subject matter anxiety. Further, the study by Shen and Eder (2009) did not explore the consequences of the exposure of learners to video games—a technology that is similar to MUVE. In view of the popularity of computer games (Hussein, Wahid, & Saad, 2009; Mohamed, Jan, & Daud, 2010), it is likely that students have had positive or negative experiences with them, which may affect their acceptance of MUVE as an e-learning environment.

Therefore, the problem to be addressed in the present study is the lack of objective evidence of MUVE effectiveness as an environment supporting the teaching of English and the lack of research on MUVE acceptance conducted in the context of using MUVE for learning English.

In view of the intent of the present study to contribute to better understanding of the viability of MUVE as an environment for language study used to improve the performance of Malays in Malaysia at communicating in English in business contexts (see the discussion in section 1.2), the problem is addressed in a particular context: in the context of Malays in Malaysia using MUVE to improve business English communication skills.

Communicative language teaching is commonly seen as the most effective paradigm for improving second language communication performance (as discussed in section 2.3). Moreover, task based language learning focuses on particular tasks and thus is particularly suitable for teaching business English communications skills pertaining to a particular task set. Therefore, the problem is addressed in the context of using MUVE as an environment for communicative English teaching relying on task based English teaching design. Further, to fully leverage the benefits offered by the MUVE environment (described in section 1.1.1), the tasks are formulated as a scenario making use of both social and physical environment; therefore, the learning experience can be described as scenario based, situated learning (introduced in section 2.9.3).

1.4 Research Questions

The research questions for this study focus on MUVE effectiveness (ability to promote learning) and MUVE acceptance by the learners:

1. Is MUVE based learning effective in facilitating situated scenario-based learning of oral business English communication skills by Malay learners?
2. Which factors influence the acceptance of MUVE based learning of oral business English communication skills by Malay learners?

1.5 Theoretical Foundations

To address the research questions of this study, Malay tertiary students were involved as participants.

To address research question one (related to MUVE effectiveness), a learning experience in MUVE was designed. The learning experience was based on the theory of situated and scenario based learning (theories that are rooted in constructivism). The learning experience was tested for effectiveness in promoting the learning of oral business English communication skills in a true experiment with random allocation of participants to two groups: an experimental (or treatment) group (learning in MUVE) and a control group (learning in a traditional classroom).

To address research question two (related to MUVE acceptance), a research model of MUVE acceptance relying on social cognitive theory (Bandura, 1997) and on the technology acceptance model (Davis, 1989) was developed. The model was tested by fitting it to the data collected in a cross-sectional survey.

A detailed introduction of the theoretical issues underpinning this research is provided in Chapter 2, the research hypotheses are introduced in Chapter 3, and the details of the research design are provided in Chapter 4.

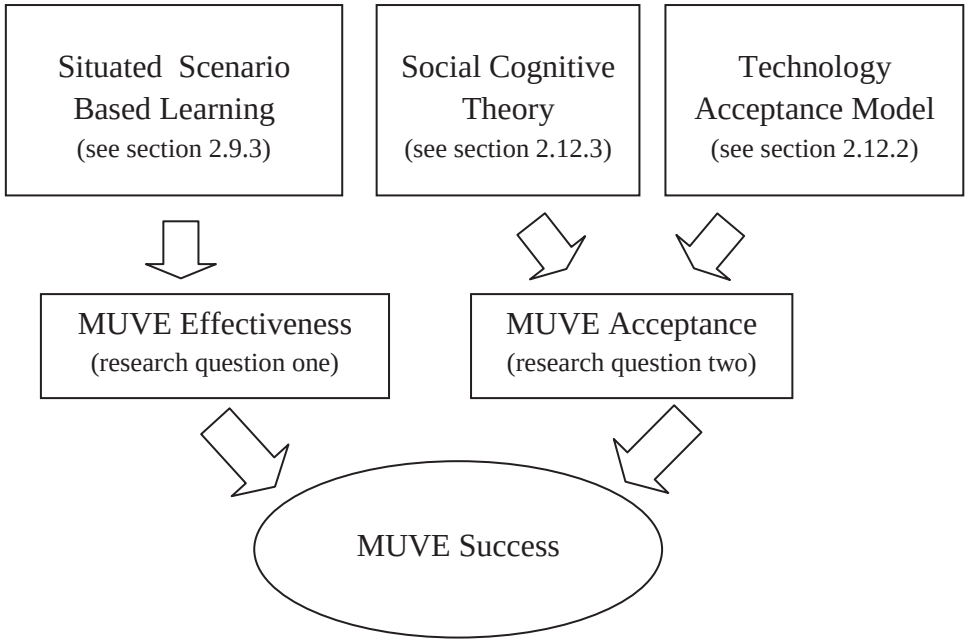

Figure 1. The theoretical foundations of this research.

1.6 Summary of Methods: How the Research Methodology Addressed the Research Problem

As indicated in section 1.4, to address the research question 1, "Is MUVE based learning effective in facilitating situated scenario-based learning of oral business English communication skills by Malay learners?", a controlled experiment with random assignment of participants was conducted. To address the research question 2, "Which factors influence the acceptance of MUVE based learning of oral business English communication skills by Malay learners?", a survey was conducted. The experiment and the survey were combined, and thus involved the same participants (for the details of the research procedure, refer to Table 4).

The participants were university students at a public institution of higher learning in Malaysia involved in an English for Professional Communication course. The instructional intervention was not a part of the course, but was conducted in a controlled environment, with the participants volunteering to take part.

The instructional intervention involved scenario based learning using a scenario of selling a house. The participants learned in pairs, adopting in turn the roles of the buyer and of the seller, with an instructor providing scaffolding when necessary. The instructional intervention continued for 40 minutes.

In the experiment (addressing the research question 1), the participants were assigned at random in pairs to the experimental group (learning in MUVE using a highly realistic model of a house in the MUVE environment) and the control group (learning face to face using a plastic model of a house). Learning was assessed via pre and post tests involving audio recording the participants executing the scenario and rating the communication performance at performing the scenario roles. Statistical analysis was used to determine if the improvement in communication performance in the house selling scenario from pre test to post test for participants who learned in MUVE (the experimental group) differed from the improvement for the participants who learned in traditional classroom (the control group).

In the survey (addressing the research question 2), the participants filled in a questionnaire measuring the participants' anxiety and affect (with respect to studying English and with respect to video games) and other factors hypothesised to affect MUVE acceptance (for the details of the research model, refer to Figure 16), along with the participants' Intention to Use MUVE e-Learning. Statistical analysis was conducted to test the hypotheses suggested by the research model.

1.7 Significance of the Study

Providing an empirical confirmation of the effectiveness of MUVE as an e-Learning platform. The study provides an objective empirical confirmation of the effectiveness of MUVE as an e-Learning platform for situated scenario-based learning of oral business English communication skills by Malay learners. This contributes to the overall understanding of MUVE effectiveness as an e-Learning platform.

Validating technology acceptance theories in the context of MUVE. The study employs a research model based on existing technology acceptance theories, and thus validates these theories in the context of MUVE. In particular, the study is the first to employ the concepts of social cognitive theory (Bandura, 1986; Compeau, Higgins, & Huff, 1999) to account for learners' feelings with respect to the subject matter and with respect to playing video games as factors possibly affecting the acceptance of MUVE for e-Learning.

Informing decision makers. By exploring the effectiveness and the acceptance of MUVE based e-Learning in the Malaysian context, the study is of value for educators and e-Learning managers in Malaysia. The study contributes to their ability to make informed decisions on adopting MUVE for promoting oral business English communication skills. In particular, it aids managers and instructors in deciding if the introduction of MUVE is going to bring the expected benefits in terms of the learning effectiveness and the engagement by learners. (The study makes a meaningful contribution in key areas; it does not purport, however, to provide all the necessary information needed to make such decisions. As highlighted at the end of section 1.2, multiple studies in a range of disciplines are needed to fully support managerial decision making.)

1.8 Scope of the Study and its Limitations

The study objectively compared the effectiveness of MUVE in promoting situated scenario-based learning of oral business English communication skills by Malay learners with the effectiveness of face to face learning using the same instructional approach. The objectivity of the comparison was ensured by using a controlled experiment with random assignment of the participants.

To make a controlled experiment feasible, the instructional intervention was brief (about 40 min) and highly focused, involving a single, well-defined scenario (a scenario involving selling a house). A longer instructional intervention in the experiment, covering a broad range of scenarios, may result in stronger external validity because it would closer resemble a course of study at a real educational institution and would assess the effectiveness of MUVE for a broader area of content. Nonetheless, a longer instructional intervention was not possible in the present study because of the difficulty of ensuring the involvement of the participants over a long period. In terms of the length of the instructional intervention, the present study was similar to the studies comparing the effectiveness of web based chat to face to face teaching by Abrams (2003) and Sykes (2005) and to both of the prior experimental studies (studies relying on controlled experiments or quasi experiments) assessing the effectiveness of MUVE as an e-Learning environment, by Cobb et al. (2009) and by Wrzesien and Raya (2010). Indeed, it appears that the practicalities of conducting controlled experiments involving

MUVE based instructional interventions make it difficult to conduct studies involving instructional interventions of long duration. Cobb et al. did not report the duration of the instructional intervention, but from the nature of the intervention it appears that it was a single session continuing for an hour or less (the duration of a typical lecture) and, therefore, was similar to the present study (possibly, somewhat longer). Wrzesien and Raya used an instructional intervention of 25 min duration, which, thus, was considerably shorter than the instructional intervention used in the present study.

The study tested the effects of anxiety and affect with respect to studying English and with respect to video games (along with other factors suggested by the TAM theory and by the SCT theory—see the research model in Figure 16 for details) on MUVE acceptance by learners by using a cross sectional survey.

Cross sectional surveys do not allow to test hypotheses regarding the direction of causation (Kline, 2005). For example, it is not possible to demonstrate that higher Perceived Ease of Use of MUVE results in greater Intention to Use MUVE, rather than Intention to Use results in higher Perceived Ease of Use, solely on the basis on the statistical analysis of cross sectional survey data. Moreover, cross sectional surveys are vulnerable to common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003). The respondents' answers may reflect their own theories about relationships between constructs, rather than reflect the true values of the latent variables. A longitudinal study, with independent variables (factors) assessed at a different time from the dependent variables (such as Intention to Use MUVE e-Learning in the present study) would, to an extent, test the direction of causation. Moreover, a study using different methods to measure the independent variables and the dependent variables (for example, measuring the actual use of MUVE e-Learning via direct observation, rather than via a survey questionnaire) would better address the danger of common method bias. A longitudinal study, however, was not feasible because it was not possible to ensure the involvement of the participants over a long period of time. Moreover, it was not possible to measure the use of MUVE to assess MUVE acceptance, because an organizational practice to use MUVE for communicative English learning is not established. By using a cross sectional survey to test a research model involving multiple constructs, the present study followed a well-established practice in Management Information Systems and e-Learning

research; studies by Bourgonjon, Smet, Van Looy, Soetaert, & Valcke (2013), Wojciechowski & Cellary (2013), and Campbell, Wells, & Valacich (2012) are example studies relying on cross-sectional surveys published recently in top e-Learning and Management Information Systems (MIS) journals.

The study's participants were Malay students enrolled in an English language course at a university in Malaysia, between 18 and 21 years old.

The participants were not drawn at random from all Malay learners of English in Malaysia. Therefore, automatic generalizability of the results to all Malay learners of English in Malaysia is not assured (generalizability to a population requires random sampling from the population, Zikmund, Carr, & Griffin, 2012). The course in which the participants were enrolled, however, was a typical English language course—the course did not have any special characteristics that would distinguish it from other, similar courses. Therefore, even though the specific numbers (such as the learning gains between pretest and posttest) may not accurately represent the averages for the population of all learners in Malaysia, the overall conclusions of the study are likely to be valid for a broader population (and even for populations in different cultural contexts), particularly for the learners in a similar age range and in similar circumstances. Ultimately, the readers of the account of the present research will compare the characteristics of the cohort the readers are interested in to the characteristics of the participants of the present study and judge the generalizability of the results of the present study to the readers' particular situation (in other words, analytical generalization, rather than automatic generalization to a well defined population, applies, see Yin, 2009). It is very common in MIS and e-Learning research to rely on participants drawn from a particular organization even though generalizability to a broader population is desirable, because drawing at random from the full target population is very rarely feasible. Published recently in top e-Learning and MIS journals studies by Abdulla (2012), by Daniel (2012), and by Polites and Karahanna (2012) are examples of studies using participants from a single organization even though generalizability to a broader population is desirable.

1.9 Thesis Structure

The thesis has six chapters structured as follows.

Chapter 1 gives a general overview of the motivation and rationale for conducting the research. It includes the statement of the research problem, the research questions, a summary of methods, and the statements of the significance of the research and of the scope of the research.

Chapter 2 begins by discussing the issues around teaching business English in Malaysia. Then, the chapter discusses second language learning paradigms and technology used to promote language learning. The use of MUVE as an e-learning platform and the applications of MUVE in promoting language learning are then discussed. Then, prior studies involving objective comparisons of the effectiveness of MUVE based learning with face to face learning are discussed in detail. Then, an overview of theories relevant to technology acceptance is given. Finally, the chapter discusses prior MUVE acceptance studies.

Chapter 3 poses the research hypotheses of this study, including hypothesis E1 and E2, which address research question one (related to MUVE effectiveness), and hypotheses A1 to A14, which address research question two (related to MUVE acceptance).

Chapter 4 presents the methodology of the research, including research setting, participants of the study, procedures used in data collection, operationalisation of constructs, and approaches to data analysis.

Chapter 5 describes the results of the data analysis for MUVE effectiveness and acceptance.

Chapter 6 presents the summary of the study followed by the discussion of the findings. It also presents the implications of the study for theory and for practice and discusses the limitations of the study, followed by presenting suggestions for future research. The chapter ends by presenting the conclusions.

Chapter 2: Review of Literature

2.1 Introduction

The problem addressed by the present study relates to the use of MUVE as an e-learning environment for learning business English communication skills, and the problem is motivated by the desire to contribute to improving business English communication skills of Malays in Malaysia. Therefore, this chapter discusses the state of English and English learning in Malaysia, second language learning paradigms and instructional designs, and Business English as a subject of study. The chapter then focuses on the technology used to support language learning and on prior studies involving comparisons between technology-supported language learning and traditional face to face language learning.

Then, this chapter discusses MUVE affordances for e-Learning and the current practice of MUVE e-Learning, along with broader e-Learning technology trends. As the first of the research questions of this study relates to the effectiveness of MUVE as an e-learning system for language learning, the existing studies of the uses of MUVE to promote language learning are reviewed.

The first research question of this thesis (see section 1.4) relates to MUVE effectiveness in supporting situated scenario-based learning. Therefore, this chapter discusses the theory underpinning situated scenario-based learning, and constructivism—the paradigm underlying scenario-based learning. The chapter also presents a detailed discussion of the existing studies of MUVE effectiveness as a learning environment.

Factors that are likely to influence MUVE acceptance as an e-learning system by Malay learners are the focus of the second research question (see section 1.4). Therefore, this chapter also discusses various theories and models of technology acceptance. This includes the use of social cognitive theory constructs in explaining and predicting technology acceptance. The constructs of social cognitive theory are particularly important for this study because they capture the relevant characteristics of Malay learners. The chapter discusses in particular detail prior studies of the acceptance of MUVE e-Learning.

The chapter concludes by stating the research gaps identified in the literature.

2.2 English Language in Malaysia

2.2.1 The Role of English in Malaysian Society

Malaysia achieved independence from the British Empire in 1957. According to the estimates of the Economic Planning Unit of the Ministry of Finance and Bank Negara Malaysia (Economic Planning Unit, 2012), in 2012 out of the total population of Malaysia of 28.9 million, 91.8% were Malaysian citizens; 66.4% of the Malaysian citizens were Bumiputera ("sons of the soil"—Malays and other indigenous people), 24.8% were Malaysian Chinese, and 7.4% were Malaysian Indians. 82.3 % of the Bumiputera (54.6% of the Malaysian citizens and 50.1% of the total population in Malaysia) were Malay. Malay are native speakers of Bahasa Malaysia, Malaysian Chinese speak a dialect of Chinese, and Malaysian Indians are native speakers of Tamil (Lazaro & Medalla, 2004). Minority languages spoken by non-Malay Bumiputera are not officially supported as strongly as Bahasa Malaysia, and non-Malay Bumiputera are gradually shifting to using Malay; at the same time, Malaysian Indians are shifting to using English, rather than Tamil, and Malaysian Chinese are shifting to using standard Mandarin Chinese, rather than local dialects (David, Cavallaro, & Coluzzi , 2009).

After the British colonization ended, with the purpose of fostering national unity, Malay became the national language (following the National Language Act of 1967) and has been used as the language of administration and as the main medium of instruction at schools and universities, which, before independence, had mainly used English as a medium of instruction (Lazaro & Medalla, 2004; Gill, 2005). English became the official second language and has been taught as a compulsory subject at primary and secondary schools, as well as at the tertiary level (with the status of English as a language taught, but not the language of instruction, set out in the Third Malaysia Plan, Government of Malaysia, 1976, p. 386). Private universities are an exception, though, and mostly are using English as the medium of instruction (David et al., 2009). Both Malay and English are used for communication between representatives of different ethnic groups (Kim, 2003).

With globalization, the role of English as the international language for business communication among Asian countries has increased (Ismail, 2008). In Malaysia, the awareness of the importance of Business English has increased after 1993, when the Malaysian political establishment set an aim for Malaysia to become a fully developed economy by 2020. There is a growing realization in Malaysia that in order to stay competitive in the global market and to attract foreign investments, the country needs a workforce capable to use English, which is the language of communication in the domain of international business and industry (Kim 2003). Ismail (2008), based on an elegant analysis using the gravity model from transaction cost economics to fit gross national product (GDP) and geographical data, came to a conclusion that the most important international language for trade, finance, and technology in Malaysia is English, followed by Chinese.

Along with globalization, Gill (2005) cited lack of language legislation in the domains of business and industry as a reason for the continued importance of English in Malaysia (even though the use of Bahasa Malaysia is prescribed in civil service and at government funded educational institutions, in corporate business and industry, banking, and finance in Malaysia, the use of English remained widespread). Gill also pointed out that the employment base for graduates from public universities weakened when opportunities for employment in the civil service greatly reduced in the 1990s; private sector, where the most important language was English, become the main employer of graduates. Pang and Heng (1996), in a survey of private companies throughout Malaysia, found that the work at most of the companies required a broad range of English language skills, with speaking in English rated as the most important skill (among listening, speaking, reading, and writing). In a more recent, similar survey of business organizations, Lan, Khaun, and Singh (2011) confirmed that English is seen as important for success, with speaking and writing considered as the most important skills. Further, the importance of oral business English for the success in the chemical engineering industry in Malaysia was highlighted by Kassim and Ali (2010) based on a survey of engineers at multinational chemical companies in Malaysia.

2.2.2 Gap Between Industry Needs and Graduates' Proficiency in English

In Malaysia, in spite of the English language training that students receive, a significant gap remains between the English language skills of the bulk of the graduates and the industries' requirements. Schools and universities have been criticized for failing to meet the industry requirements in terms of preparing graduates with sufficient English language skills. For example, in 2003 the Education Minister at the time, Tan Sri Musa Mohamad, called upon public universities to re-examine their programs to improve English language communication proficiency of the graduates (Kassim & Ali, 2010).

Malay graduates may be disadvantaged compared with Malaysian Indian and Malaysian Chinese graduates, because Malaysian Chinese and Malaysian Indian students are more likely to use English at home and because the Malay community is less accepting of the use of English (Thang 2004; Manan & Shamsudin, 2012). Manan and Shamsudin (2012) conducted content analysis of picture descriptions written in English by Malay and by Chinese students and found that Chinese students performed better than Malay students in terms of language productivity, vocabulary range, and sophistication. Kim (2003), in a critical ethnography study of Malaysians from different ethnic backgrounds, found that Malays who were proficient in English felt that using English could be interpreted by other Malays as renouncing their Malay cultural identity and religious identity, and therefore refrained from using English in front of other Malays.

A survey carried out by the Malaysian Employers Federation (MEF), involving the 205 member companies, uncovered major problems among local graduates in their inability to communicate well in English (MEF, 2004). Kassim and Ali (2010) attributed this finding to a decline in English language proficiency in the generation of Malaysians that received education after the independence and thus did not have English as a medium of instruction and grew up in an environment that did not require them to use English.

Over the years, the political establishment adopted policies to improve the English language proficiency of the graduates entering the workforce. Tan (2005) analysed the content of press and news agency reports to study the debate around a partial re-introduction of English as a medium of instruction in primary and secondary

schools. He noted that the debate was centred on the status of English as an international language, and tended to ignore the established use of English inside Malaysia for communication between the representatives of different ethnic groups. Tan attributed this one-sided view of English in Malaysia to discomfort due to the association of English with the Malaysia's colonial past. The success (or, rather, the lack of success) of an initiative to partially re-introduce English as a medium of instruction of Science and Mathematics is discussed by Darus (2010), based on the analysis of literature and media sources. Darus noted that the initiative has been abandoned, and the teaching of Science and Mathematics in English was to be phased out starting from 2012. Darus (2010) attributed the failure of the initiative to the lack of competent English teachers and to the inability of mathematics and science teachers to teach their subjects effectively in English. Gill (2012) conducted a similar analysis with similar conclusions. Darus (2010) also noted the need to find a balance between using communicative English teaching and audio-lingual and grammar translation that would work in practice, to improve the current practice of over-reliance on behaviourist approaches. Further, Zaaba (2011) and Gill (2012) pointed out the top-down nature of decisions on language policy, with the federal government playing a critical role.

2.2.3 Practice of Teaching English in Malaysia

The communicative approach to teaching English has been introduced in the Malaysian school syllabus with the aim of improving English communication skills of the graduates to meet the manpower needs in the business and commercial sectors (Foo & Richards, 2004). As early as in 1989, the Ministry of Education handbook for English Language Teaching (ELT) promoted the communicative teaching approach to English (Malaysian Education Ministry, 1989). However, the implementation of the communicative teaching approach in the classroom has remained insufficient. Classroom teaching often relies on the chalk-and-talk drill method, with the focus on reading comprehension, writing, grammar, and vocabulary skills, which are the skills tested in the national school examination, thus largely neglecting listening and speaking skills, which are not covered by the examination (Pandian, 2002; Musa, Koo, & Azman, 2012).

At university level in Malaysia, efforts are made to encourage students to use the English language particularly for speaking at presentations, but the speaking component by far is still the least emphasized skill taught compared to other language skills like reading and writing (Hamidah, 2003; Yong, 2010). Furthermore, many students (particularly, Malays) lack confidence and fluency in using the language, which forces them to memorize their lines or content when doing their presentations. As a result, students tend to feel incompetent in situations requiring real face-to-face interaction using English, and are reluctant to engage (Hamidah, 2003; Yong, 2010). Some of the difficulties of introducing communicative language teaching may be because of the cultural background of Malaysian learners, who tend to be closure-oriented and concrete-sequential (prefer clear directions, certainty, and detailed information) and introverted (prefer to receive knowledge from the teacher, quiet, shy, and reticent in class) (Yong, 2010).

Thus, even though English plays an important role in Malaysian society in general and in the economy in particular, and even though because of globalization the importance and the relevance of English is likely to increase, graduates entering the workforce (particularly, Malay graduates) do not have sufficient English communication skills (in particular, Business English communication skills). The current practice of relying on audio-lingual and grammar translation in English teaching appears to be one of the causes of the problem.

2.3 Language Teaching Paradigms and Instructional Designs

This section introduces the three main language teaching paradigms: grammar translation method, audio lingual method (based on behaviourist psychology), and communicative language teaching. Then, the section presents two approaches to instructional design used in language teaching that are particularly relevant to the problem addressed in the present thesis: content based language teaching and task based language teaching.

The grammar translation method is based on a premise that languages follow formal, deterministic grammar rules, suggesting that a language can be learned by memorizing formal grammar rules and vocabulary (Prator & Celce-Murcia, 1979; Brown, 2000). In language teaching based on the grammar translation method

classes are taught in the mother tongue; vocabulary is taught as word lists; detailed grammar rules are taught (including long lists of exceptions); and the focus is on grammatical analysis of text rather than on content and on disconnected sentences rather than on meaningful texts. Grammar translation method has been found to lead to low student motivation and to do little to develop communicative ability. Nonetheless, grammar translation method is still widely practiced because it is easy to implement and it is effective in preparing learners for formal language proficiency tests. The premise of the grammar translation method that languages follow formal, deterministic grammar rules has been challenged in recent research, because based on corpus based analysis it was found that the actual language use is highly contextual and is better described in terms of probabilistic statements than in terms of deterministic rules (McCarthy, 1998).

The audio lingual method is based on behaviourist psychology and thus on the premise that linguistic behaviour can be taught via conditioning and habit formation (Prator & Celce-Murcia, 1979; Brown 2000). In language teaching based on audio lingual method the emphasis is on memorization of patterns, with successful utterances reinforced; patterns are taught via repetitive drills, often with the use of tapes and visual aids; unlike in grammar translation method, there is emphasis on pronunciation, and grammar is taught by analogy rather than by memorizing formal rules; and similar to grammar translation method, the emphasis is on form, rather than on content. Audio lingual method had little success in developing communicative ability beyond very simple scenarios. Nonetheless, the audio lingual method is still widely practiced for the same reasons as the grammar translation method—it is easy to implement and it is effective in preparing learners for formal language tests focusing on the use of specific language patterns.

Communicative language teaching is based on multi-component views of communicative competence (such as the theory by Canale & Swain, 1980, distinguishing grammatical, discourse, sociolinguistic, and strategic competences). In communicative language teaching the second language learner engages in communicative interaction with competent speakers of the target language, in unrehearsed contexts; at the beginning, maximal use is made of the communicative competencies the learner already has (e.g., developed when learning the native language); the learners are also taught about the target language culture, to enable

them to understand the social meaning of utterances in the target language; and unlike in grammar translation and in audio lingual methods, the focus is more on fluency (on the ability to communicate in real world situations) rather than on accuracy (on the ability to produce grammatically correct utterances), with accuracy eventually developed as the result of communication practice enabled by fluency (Canale & Swain, 1980; Brown, 2000). Communicative language teaching is believed to result in considerably better communicative competency than grammar translation or audio lingual methods; however, it may be difficult or impossible to implement with large classes, and it requires highly competent teachers.

Content based language teaching is a way to implement the communicative teaching paradigm. In content based language teaching, learners engage in communicative interaction focused on a particular area of content (Met, 1998; Pessoa, 2007). Thus, the focus is on understanding the content and on communicating about the content rather than on the target language, with an important constraint—the content is presented in the target language and, for as much as possible, all communications in the classroom should be in the target language. The learners are expected to attain fluency in communicating about the content, with the mastery of the language achieved almost as a side effect. Content based language learning is likely to be particularly effective when the learners are motivated both in learning about the content and in mastering the target language. To implement content based teaching, teachers both knowledgeable about the content and fluent in the target language are needed. The unsuccessful initiative of the Malaysian government to introduce teaching of Mathematics and Science in English in secondary schools (discussed in section 2.2.2) demonstrates that an effective implementation of content based teaching may be difficult because sufficiently qualified and motivated teachers may not be available.

Task based language teaching focuses on real world tasks in which the learners are likely to use the target language in the future (Nunan, 2005; Oxford, 2006), and involves learners practising such tasks using the target language. Designing a course according to the principles of task based language teaching involves conducting a needs analysis, choosing the most important tasks (in terms of task importance in the real world, rather than in terms of linguistic consideration), creating materials that would both address the tasks and fit student background and level of proficiency,

and organizing the tasks into a sequence to form a task syllabus, according to nonlinguistic criteria such as increasing task complexity. An important principle of task based language teaching is the focus on form—when learners attempting to complete a task make linguistic mistakes, such as using incorrect grammar or using wrong vocabulary, the teacher is supposed to intervene and to teach the correct form. In task-based language learning, tasks are used to create interaction and then to build language awareness and language development around the learners' task performance. Ideally, assessment in task based learning is based on learners' competency at performing the tasks, rather than on generic measures of second language proficiency.

Task based learning relies on all the three major language teaching paradigms: communicative language teaching, audio lingual (behaviourist) teaching, and the grammar translation method (Nunan, 2005). Focusing on learners' execution of real world tasks is clearly in the spirit of communicative language teaching. The primary focus is on achieving fluency in communicative interaction; moreover, learners are taught how to use the language appropriately in the task context (and thus acquire the background cultural knowledge). Nonetheless, the clear secondary focus on form suggests explicitly learning vocabulary and correct (in terms of following the native speaker model) language patterns (in the spirit of the audio-lingual method), or even learning explicitly the relevant grammar rules (in the spirit of the grammar translation method).

2.4 Business English as English for Specific Purposes

Business English is English used in business contexts (Richards, 1989; Boyd, 1991). Business English is a particular type of English for Specific Purposes (ESP). Unlike general English courses, ESP courses focus on specific application areas and emphasise context-specific features of the language, rather than generic skills. Students who are studying ESP are not intended to become fluent speakers of the language. Rather, the intention is for learners to become able to use the language to perform a well defined set of tasks in the target language (Mackey & Mountford, 1978; Graham & Beardsley, 1986; Hutchinson & Waters, 1987).

Language specificity of ESP courses refers to specific lexical, semantic, and syntactic characteristics of the technical language in the target field. In ESP, language specificity is relevant to both teaching and testing (Hutchinson & Waters, 1987; Mumby, 1987; Douglas, 2000).

Language teachers and instructors have long recognised the importance of Business English, especially the importance of developing the oral skills required in real-life business contexts. Various commercially-produced teaching materials intended to promote Business English communication in the form of published texts, audio-visual, multimedia, and online lessons on the Internet are available. However, the teaching methods supported by most of these commercially produced materials do not reflect the state of art in e-learning and educational research, as the materials tend to focus on drills and practice, rather than on interaction in authentic contexts (Faridah, 2005; Lainema & Nurmi, 2006). Therefore, learners end up memorising factual knowledge and do not sufficiently develop the ability to apply the obtained knowledge in the target real-life contexts. To improve the transference of learning to target real-life contexts, learning should happen in physical and social contexts similar to the target contexts (Nair-Venugopal, 2006)

2.5 Language Learning and Technology

Section 2.3 reviewed the main language learning and teaching paradigms and instructional designs. Different language teaching paradigms are supported by different technology. The early CALL systems mainly provided drill-and-practice exercises like text reconstruction, gap-filling, speed-reading, simple simulation, and vocabulary games (Levy, 1997). Thus, the early computer assisted language learning (CALL) systems relied on grammar-translation and audio-lingual methods—the two language learning paradigms that pre-dated communicative language teaching.

Some of the recent CALL systems continue to support the audio-lingual method, with perhaps a higher level of technical sophistication. It is possible, for example, to use automatic speech recognition technology to allow learners to have a simulated conversation with a computer (using a pre-determined and fixed script with rigid decision points) (Hincks, 2002). The software can provide individual feedback on pronunciation skills at phoneme, word, or sentence level. In the pronunciation

training, learner pronunciation is compared to a model speaker. Visual representation of waveforms and pitch curves aids learners to understand how their pronunciation deviates from the model speaker. Such systems, however, limit learners to responding to pre-determined texts, lack context, and do not allow learners to practise free-form conversation, which is necessary for improving communicative skills.

The spread of the Internet enabled computer-mediated interactions between users in different locations; consequently, the focus of English as a Second Language (ESL) in CALL shifted from drill-and-practice to computer-mediated communication (Liu, Moore, Graham, & Lee, 2002). Computer-mediated communication is defined as the application of computer and Internet technology in human communication (Thurlow, Lengeh, & Tomic, 2004). In synchronous communication, users interact via technology in real time (e.g., instant messaging and chats), whereas in asynchronous communications, users rely on technology to transmit messages that can be accessed at a later time (e.g., electronic discussion boards and e-mail).

The communicative approach can be supported by using computer mediated communication tools such as e-mail, chat, and bulletin board systems by employing these tools to enable communication in the target language (Abrams, 2003). Some examples of the uses of synchronous communication tools to promote second language learning are the studies by Abrams (2003) and Blake (2009) (describing the use of text chat) and the study by Sykes (2005) (describing the use of voice chat). These studies are reviewed in detail in section 2.6. However, communication tools such as text chat and voice chat, while enabling communication, do not simulate the physical context in which the target language is to be used.

The approach to teaching oral business English communication skills evaluated in this thesis combines the use of technology for synchronous communication with the use of technology to provide a hi-fidelity simulation of the target physical context. In particular, the approach relies on MUVE technology, which offers both the capability to provide a highly realistic representation of an authentic physical environment (via virtual reality), and the capability to provide an authentic social environment (via multi-user participation using avatars). This approach promises to support communicative language teaching better than the generic communication technology by being more faithful to the principles of constructivist learning that

underpin communicative language teaching (constructivist learning is discussed in section 2.9).

2.6 Effectiveness of Using e-Learning for Second Language Learning: Comparisons with Face-to-face Learning

This section introduces prior research comparing the effectiveness of using e-learning for second language learning with traditional, face to face learning. The section focuses on technologies that are particularly relevant to understanding the potential of, the limitations of, and the alternatives to using MUVE for second language learning.

The search for studies comparing the effectiveness of using e-Learning for second language learning with traditional, face to face learning was conducted as follows. First, review articles covering the research on the effectiveness on the use of technology to promote language learning were identified. The review articles were Golonka, Bowles, Frank, Richardson, & Freynik, 2012; Vanderplank, 2010; Tang, Mu-shang, & Pan-ju, 2009; Felix, 2008; Stockwell, 2007; Blake, 2007; Johnson & Johnson, 2006; Felix, 2005a; and Felix, 2005b. Then, articles cited in these reviews were considered for inclusion.

The following inclusion criteria were applied. First, only the articles that involved a systematic comparison between technology based learning and face to face learning were included. In particular, an article was included if it involved a true experiment (with random assignment of the participants) or a quasi-experiment comparing the effectiveness of technology based language learning with face to face learning. Moreover, multiple case studies (in the sense of Yin, 2009) involving at least one case of language teaching using technology and one case of language teaching face to face and involving purposeful collection of data and systematic cross case comparison were also intended for inclusion.

The following kinds of studies were not included as not involving a systematic comparison between technology based language learning and face to face learning. Single case studies and experience reports limited to describing the use of technology for language learning were not included. Moreover, studies of the effectiveness of technology based language learning involving the use of pre and

post measurements of language proficiency, but with no control groups with face to face learning delivered as an alternative to technology based leaning (thus enabling comparison of the effectiveness of the alternative approaches), were not included.

Second, only the studies of the uses of the following technologies to promote language teaching were intended for inclusion: video conferencing, video games, intelligent tutoring systems, avatars, text chat, social networking, blogs, discussion forums, and wiki. The reason for including these technologies is that all of them enable high levels of interactivity, and thus are similar to MUVE. In particular, studies of non-interactive uses of video to promote language learning were not included.

In addition to considering for inclusion the articles covered in the review articles, Google Scholar (http://scholar.google.com) was searched for keywords such as technology, e-learning, student, learner, and effectiveness, and the resulting articles were considered for inclusion. Moreover, articles citing the articles included and articles cited in the articles included were also considered for inclusion. The same inclusion criteria (described in the preceding two paragraphs) were applied throughout.

Of the technologies that were intended for inclusion, video conferencing, video games, intelligent tutoring systems, avatars, text chat, social networking, blogs, discussion forums, and wiki, studies involving systematic comparison between technology based leaning and face to face learning were found only for video games, text chat, blogs, discussion forum, and wiki. Studies of the effectiveness of text chat as a way to teach a second language were particularly well represented (four studies out of the seven studies found), and are summarized in Table 1. The studies involving blogs, wiki, and video games are summarized separately, in Table 2. One of the studies covering the use of text chat (Abrams, 2003) also covered the use of discussion forums, and it is summarized in Table 1, along with other studies of text chat. Notably, none of the studies involved the use of avatars or MUVE (thus confirming that the problem addressed in the present study was not addressed in prior research).

Table 1: Studies Comparing Text Chat to Face to Face Learning

Study	Abrams (2003)	Sykes (2005)	Lu (2008)	Blake (2009)
Technology	Web based text chat and discussion forum tools	Web based text chat and voice chat tools	SMS text messaging	Web based text chat tool
Method	Quasi-experimental	Experiment	Experiment	Experiment
Participants	Third semester university students	Third semester university students	High school students	University students
Native language	English	English	Mandarin Chinese	Multiple
Target language	German	Spanish	English	English
Learning topic	Comparing German and US culture	Refusing invitation	Building vocabulary	Acquiring basic vocabulary
N in treatment	32 (Group A) and 31 (Group B)	3 teams of 3 students	15	10
Treatment group	Two treatment groups. Group A discussed synchronously for 50 minutes using text chat and Group B discussed asynchronously for one week using a discussion forum.	Two treatment groups. Group A practiced dialogues via text chat (for 30 min), and Group B - via voice chat (via a computer) (for 30 min).	Participants received English words (over two weeks, two times per day) using SMS text messages and wrote down Chinese translations	12 text-chat sessions over six weeks
N in control	33	3 teams of 3 students	15	14
Control group	Face to face: small group and pair work	Practise dialogues in a traditional classroom (for 30 min).	Participants received lists of English words (over two weeks, once at the beginning of each week) on paper and wrote down Chinese translations	12 face to face discussion sessions over six weeks
Measure	Content analysis of oral output in discussions conducted before and after the instructional intervention	Pre and post tests involving recording role play. The content of the recordings was analysed to count head acts and supporting moves and as qualitative data.	Pre-test, post-test, and delayed post test (after three weeks) involving writing Chinese translations of English words, with the the number of correct translations counted.	Content analysis of oral output in participants' answers to open ended questions before and after the instructional intervention, to assess five fluency measures
Treatment / control	Group (A) - students using text chat - outperformed control (face to face) and Group (B) in terms of the amount of speech. There was no statistically significant difference in terms of lexical richness or syntactic complexity.	According to qualitative analysis, the text chat group outperformed voice chat and face to face groups in terms of complexity and variety of speech.	Students in the group using SMS learned statistically significantly more words, according to both the immediate post test and the delayed post test.	On two of the fluency measures, phonation time ratio and mean length of run, the students who participated in text-chat sessions outperformed the students who participated in face to face discussions

Table 2: Studies Comparing Video Games, Wikis, and Blogs to Face to Face Learning

Study	Suh, Kim, & Kim (2010)	Alshumaimeri (2011)	Fageeh (2011)
Technology	Video game	Wiki	Blog
Method	Quasi-experiment	Quasi-experiment	Experiment
Participants	Fifth and sixth year elementary students	First year university students	Forth year university students
Native language	Korean	Arabic	Arabic
Target language	English	English	English
Learning topic	Discussing, reading, and writing in English	Descriptive writing	Developing argumentative and descriptive writing skills
N in treatment	118	22	25
Treatment group	Group based activities in a game environment, structured around completing tasks built into the game over two months (two 40 min lessons per week)	Participants performed writing activities in groups, with each group sharing a Wiki (over 20 weeks, one hour per week)	Participants posted blog entries on set topics, with the blog entries later discussed in face-to-face classroom and revised based on the discussions (over six weeks)
N in control	102	20	25
Control group	Face to face group based activities in a conventional classroom over two months (two 40 min lessons per week)	Writing activities in the classroom (over 20 weeks, one hour per week)	Face to face learning (over six weeks)
Measure	A broad English learning achievement test executed immediately before and immediately after the instructional intervention	Pre and post tests, involving students writing on a given topic. Scored on lexico-grammatical accuracy and quality of writing	Pre-test and post-test involving writing an essay, which was rated on fluency in conveying ideas, grammar, and vocabulary, using a single score covering all three dimensions
Treatment / control	No statistically significant difference	The treatment group statistically significantly outperformed the control group in terms of the improvements in both accuracy and quality of writing	The treatment group statistically significantly outperformed the control group

As to the study design, experiments and quasi experiments were equally well represented. Most of the studies involved non-native speakers studying English; in two of the studies (Abrams, 2003; Sykes, 2005), native English speakers studied other European languages (German and Spanish, respectively).

The focus of the learning interventions ranged from acquiring basic vocabulary to developing sophisticated communicative skills. All of the three main paradigms of second language learning (grammar translation, audio-lingual, and communicative) were represented, with learning experiences ranging from rote memorization of vocabulary items to collaborative writing of complex texts. Nonetheless, all of the studies (with the exception of Sykes, 2005, which focused on a very narrowly defined communicative competency) relied on general measures of language performance, rather than on learners' communicative performance in situations close to real life, as suggested by communicative language teaching (introduced in section 2.3).

Studies with quasi-experimental design tended to have larger numbers of participants than true experimental studies with random assignment of participants to groups. This was probably because large existing classes could be used in quasi experimental studies, but it may be more difficult to recruit large numbers of participants to execute a true experimental design (as discussed in section 4.3.2).

The duration and the specificity of instructional interventions varied considerably from study to study, ranging from highly targeted brief instructional interventions to instructional interventions lasting over several months. The difficulties associated with interpreting the results of studies involving long instructional interventions are highlighted by the outcome of the study by Suh, Kim, & Kim (2010). Even though the learning experiences of the treatment group and of the control group were very different (a video game versus face to face) and the numbers of the participants were rather large, there was no significant difference, unlike in most of the rest of the studies. It is quite possible that any improvement in language proficiency was caused by factors outside of the researchers' control and common to both of the groups, such as the learners watching television programs in English outside the class (and thus, the result is due to a maturation effect, see Weathington, Cunningham, & Pittenger, 2010). In a study conducted

over a long period of time the researchers are unlikely to be able to control all aspects of the learners' experiences, making such studies vulnerable to maturation effects.

In terms of the effectiveness of the use of technology to promote second language learning, most of the studies (five out of seven) reported statistically significant differences, with the participants learning with the use of technology performing better in terms of improving language proficiency than the participants learning face to face.

In the study by Sykes (2005), which compared the use of text chat and voice chat with face to face teaching to teach the participants speech acts used to refuse invitation, no significant difference was found; this, however, is not surprising because the study employed a very brief (even though, targeted) instructional intervention with a very small number of participants. The no significant difference result is more surprising in the study of language learning by using a video game environment by Suh, Kim, and Kim (2010) because the study involved a very large number of participants, with the instructional intervention continuing over a long period of time. The possible explanation that any learning was because of factors not under control of the researchers and common to both the treatment and to the control group was suggested earlier in this section. It is, however, somewhat difficult to interpret the result by Suh et al. (2010) because the nature of the pre and post tests used to assess effectiveness is not reported in the study in any detail. In contrast, in the study by Blake, 2009, which also involved a relatively long instructional intervention, a statistically significant difference (in favour of technology based learning) was found. However, only 71% of the participants who were initially recruited completed the course and were involved in the comparison, possibly biasing the study results. This highlights another problem associated with conducting studies of long duration—participant attrition presents a threat to the validity of the results (Weathington et al., 2010). Participant attrition is not an issue if participants are required to participate; this, however, raises ethical issues (see, for example, Massey University Code of Ethical Conduct for Research, Teaching and Evaluations Involving Human Participants, http://www.massey.ac.nz/massey/fms/Human Ethics/Documents/MUHEC Code 2 02013.pdf).

It is possible that studies that found a statistically significant difference in favour of using technology were more likely to be submitted and, ultimately, accepted for publication (according to Sternberg, Dietz-Uhler, & Leach, 1993, studies reporting statistically significant differences are more likely to be accepted for publication), resulting in an overall bias in favour of technology based learning. Assuming, however, that the statistically significant differences in favour of using technology reported in most of the studies are not because of such bias but because the use of technology does promote second language leaning, they can be seen as a further confirmation of the importance of the present study. Because most of the prior studies comparing the use e-learning technologies promoting interactivity for second language learning to face to face learning (summarized in Table 1 and Table 2) found statistically significant differences in favour of using technology, it is likely that MUVE, which enables high levels of interactivity (as discussed in section 1.1.1), is likely to promote second language learning.

Both synchronous interactivity using technology (text chat and instant messaging, in the studies by Abrams, 2003, Lu, 2008, and Blake, 2009) and asynchronous interactivity using technology (wikis and blog, in the studies by Alshumaimeri , 2011 and Fageeh, 2011) have been demonstrated to lead to better improvements in second language proficiency than face to face learning. MUVE enables both synchronous and asynchronous interactivity. For example, in Second Life MUVE one can interact synchronously by using voice, text messages, and gestures, and asynchronously, by sending notecards (http://wiki.secondlife.com). Therefore, both the results for synchronous interactivity using technology and the results for asynchronous interactivity using technology suggest that MUVE is likely to be effective in promoting second language learning.

The rest of this section discusses the main features of each study, one by one.

In the study by Abrams (2003) (see Table 1 for the study details), the effects of synchronous text chat discussions on the learners' language proficiency were compared with the effects of face-to-face, classroom study and of asynchronous discussions using an online discussion forum tool. The study used a quasi experiment design—existing classes were assigned to different treatments. Pre and post tests were used to assess language proficiency and involved the measurements of the quantity of output, lexical richness, and syntactic complexity. The instructional intervention

involved discussions of German culture (by native speakers of US English in the target language, German); the discussions were conducted face to face, in small groups (in the control group), using synchronous text chat (treatment group A), and over an asynchronous discussion forum (treatment group B). From what is reported in the article, it is not entirely clear if IT based treatments were in addition to face to face learning or were instead of face to face learning. If the IT based treatments were in addition to face to face learning, the study by Abrams does not constitute a comparison of effectiveness between technology based learning and face to face learning and therefore is out of the scope of the review conducted in this section. The rest of the discussion here is based on the assumption that the treatment groups were involved in technology based language learning and were not involved in face to face learning, so that the study does involve a comparison of between technology based learning and face to face learning.

It was found that the use of text chat resulted in a greater improvement in terms in the quantity of output compared with face to face study, and face to face study resulted in a greater improvement in terms of the quantity of output than discussion forum discussions. There were no statistically significant differences in terms of lexical richness or syntactic complexity. Even though the authors suggest that an improvement in the quantity of output reflects an improvement in communicative ability, one may argue that an improvement in the quantity of output was merely the result of the students getting to know each other and, therefore, shy students contributing more to the discussions. If fact, it would not be unreasonable to expect a similar effect even if all of the participants were highly proficient native speakers of the target language. The treatment involving a text chat tool lasted for 50 minutes—a length of time that is hardly sufficient to improve ones vocabulary or grammar in a broad sense.

Nonetheless, according to the principles of communicative language teaching (discussed in section 2.3), higher quantity of output, and thus, higher involvement of learners in communication involving the target language, is likely to eventually result in better improvements in the quality of output. Moreover, from the point of communicative language teaching, the ability to communicate (suggested by the quantity of output) is more important that the quality of output. Therefore, from the perspective of the principles of communicative language teaching, the result by

Abrams (2003) demonstrates that even a brief instructional intervention involving using technology in context of language learning may result in positive effects.

Sykes (2005) conducted a study to compare the effects of text chat, voice chat, and face to face discussions on the acquisition of a speech act (refusal of an invitation) in Spanish (see Table 1 for the study details). The study used a true experimental design, with participants assigned at random. The participants were university students, native speakers of English. There were two treatment groups. Group A practiced dialogues via text chat, and Group B - via voice chat (via a computer). The control group practiced dialogues in traditional classroom. There were three teams of three students in each of the groups (with the teams assigned at random). In pre-and post-tests the participants role-played refusing an invitation, with the performance recorded on video. The content of the recordings was analysed to count head acts and supporting moves. Moreover, the content of the recordings was also analysed as qualitative data.

Statistical analysis was inconclusive, which is not surprising in view of the small number of the participants and in view of the short duration of the instructional intervention. Nonetheless, according to qualitative analysis, the text chat group outperformed the voice chat and the face to face groups in terms of complexity and variety of speech. Sykes (2005) explained the superior performance of the text chat group by suggesting that not being able to rely on communication cues such as gestures and intonation, the participants in the text chat group had to be more expressive in their communications, leading to better practice in the text chat environment.

Lu (2008) compared the effectiveness of vocabulary lessons using SMS text messaging with the effectiveness of vocabulary lessons relying on paper-based materials (see Table 1 for the study details). The study used a true experimental design—the participants were allocated to the treatment (using SMS) and to the control (using paper based materials) groups at random. The participants were 30 high school students, native speakers of Mandarin Chinese. The participants were given lists of English words either on paper (for the control group) at the beginning of the instructional intervention or via SMS messages (for the treatment group) over the duration of two weeks, two times per day every day. The participants were instructed that they were required to learn the vocabulary to prepare for a test. Pre and post tests testing the knowledge of the vocabulary were conducted immediately before and

immediately after the instructional intervention. In addition, three weeks after the end of the instructional intervention, a delayed post test was conducted to test for knowledge retention. Semi-structured interviews of the participants in the treatment group were also conducted, to assess the participants' attitudes with respect to the instructional use of SMS messaging.

The treatment group (taught by using SMS messages) registered improvements in the knowledge of the vocabulary that were statistically significantly greater than the improvements in the control group. Moreover, the participants felt overall positively about the use of SMS messaging in the instructional intervention (even though some of them did have technical problems). One may argue that the result could be because of the novelty effect—learners paid more attention to words that arrived as SMS messages because such mode of learning was new to them (so that in the long term, the advantage associated with using SMS messages would disappear as the novelty wears off). Alternatively, one can interpret the outcome as evidence that even very limited use of technology—technology used to deliver content in small, manageable chunks and to simultaneously remind learners of the need to learn the content—can be beneficial.

Blake (2009) conducted a study to compare the effectiveness of text chat and face to face instruction in improving the fluency of oral discourse (see Table 1 for the study details). The study used a true experimental design, with the participants assigned at random to groups. The participants were foreign university students and their spouses who volunteered to take part, native speakers of a range of languages; the instructional intervention used in the experiment was designed and executed for the purpose of the research. Pre and post tests were used to assess five dimensions discourse fluency. The instructional intervention involved discussions on unspecified topics under the guidance of an instructor; the pedagogical aim was to build the vocabulary; the discussions were conducted face to face (in the treatment group) and via text chat (in the control group). (A group that did not take part in the discussions was also involved. However, because this aspect is not directly relevant to the purpose of the review conducted in this section, this aspect is not discussed further.) A strong feature of the research was that a controlled experiment was conducted with a realistic instructional intervention. The instructional intervention was conducted over a relatively long period of time, six weeks, with two discussion sessions per week; thus, the

instructional intervention represented an experience similar to a typical course at a school or at a university.

It was found that the participants using text chat improved their fluency statistically significantly better than the participants in the control group according to two of the five dimensions considered (phonation time ratio and mean length of run). The design involved testing five hypotheses at the same time, and thus there was a risk of capitalization of chance (of hypotheses found statistically significant because of random fluctuations, see Chin & Todd, 1995). However, because the effect on the two dimensions was in the same and in the expected direction, the likelihood of the result being because of the capitalisation of chance was not very high. The difference was not statistically significant for the remaining three dimensions. For the participants in the treatment group (using text chat) the effect size for the improvement in fluency was strong (as one would expect, because only in the presence of a strong effect there would be statistically significant effects in an experiment with rather small number of participants). The authors suggest that greater effect of text chat discussions was because all of the participants in the treatment group could be involved in the discussions all the time, while in face to face learning the participants had to follow turn taking protocol. This explanation should be contrasted with the explanation by Sykes (2005), who also found text chat to be more effective than face to face teaching, and suggested that this was because learners communicating over text chat have to rely solely on language and cannot rely on communication cues such as intonation. The explanation by Blake (2009) is clearly more consistent with the communicative language teaching paradigm.

Suh et al. (2010) compared the effectiveness of video game based learning with face-to-face learning (see Table 2 for the study details). The study used a quasi experimental design—existing classes were assigned to different treatments. The participants were elementary school students. In the video game environment the learners (the treatment group) followed a scenario that required them to read and to write in English, as well as to discuss in English to reach collective decisions. In face to face learning, the learners (the control group) were involved in reading, writing, and discussions under the guidance of a tutor. In both groups, the instructional intervention was conducted over two months (involving two 40 min lessons per week). Pre and post tests were conducted, however, Suh et al. (2010) did not report the content of the

tests in detail, apart from stating that they were broad English language achievement tests.

There was no statistically significant difference between language proficiency gains in the treatment group and in the control group. Thus, the more engaging environment of the game did not result in better achievement, even in an instructional intervention of realistic duration (so that there was enough time for any novelty effects to wear out). The tasks in the game based environment, were not authentic, but rather designed to merely control the progress of the learners and to engage the learners. The no statistical difference outcome suggests that the increased engagement (if any—the engagement was not measured in the experiment) did not result in the game based environment being superior to the face to face environment. (An alternative explanation of this result—a possible maturation effect—was discussed earlier in this section.)

Alshumaimeri (2011) compared the effectiveness of teaching English writing skills in Wiki environment with the effectiveness of teaching English writing skills in conventional face to face classroom (see Table 2 for the study details). The study used a quasi experimental design—existing classes were assigned to different treatments. The participants were first year university students. In Wiki environment the participants (the treatment group) engaged in writing as a group (the topic of writing is not stated in the article). In face to face learning, the participants (the control group) engaged in similar writing activities using pen and paper. Unlike in Wiki environment, where the content is shared via the Wiki system, so that learners work on a shared artefact, it appears that in the face-to-face environment the learners worked mostly separately, but provided feedback on each other's writing (however, the exact nature of the instructional intervention in the face to face environment is not stated very clearly in the article). For both groups, the instructional intervention continued over 20 weeks, one hour per week. Pre and post tests involved individual learners writing on a given topic, with the resulting output rated on lexico-grammatical accuracy (e.g. spelling, word choice, word order, and punctuation) and on writing quality (organization, elaboration, coverage, clarity, links, and intent).

The treatment group performed better than the control group in terms of the improvements in both accuracy and in quality of writing. Alshumaimeri (2011) attributed the better performance of the treatment group (the group using the Wiki

environment) to closer collaboration between the learners, resulting in more discussions between them.

Fageeh (2011) compared the effectiveness of teaching English writing via blog writing with the effectiveness of teaching English writing skills in conventional face to face classroom (see Table 2 for the study details). It appears that the study used a true experimental design, with the participants assigned to groups by the researcher (it not clear from the article if the assignment was at random, but in the absence of any other principle for group assignment stated explicitly, it is likely that the assignment was at random). The participants in the treatment group posted blog entries (authored in a collaborative effort, as a group) and then discussed them with a tutor in a face to face classroom; explicit roles were assigned to the participants, such as secretary, coordinator, facilitator, language expert, and web expert, which served to facilitate and to structure interactions. The participants in the control group learned by using traditional face to face instruction (the article does not state explicitly the details of the instructional intervention for the control group). For both of the groups, the instructional intervention continued for six weeks (it is not stated explicitly in the article, how many hours per week). Pre and post tests involved the participants writing an essay on a set topic, which was rated on fluency in conveying ideas, grammar, and vocabulary, using a single score covering all of the three dimensions.

The treatment group participants demonstrated an improvement that was statistically significantly better the improvement in the control group. Fageeh (2011) suggested that the treatment group participants performed better that the control group participants because they paid closer attention to formal aspects of writing; however, because the details of the instructional intervention for the control group are not sufficiently clear in the article, it is difficult to interpret the results.

Overall, the results of the studies where statistically significant differences were found (which were always in favour of technology based learning) suggest that technologies enabling interactivity promote second language learning. Therefore, because MUVE enables particularly high levels of interactivity (as discussed in sections 1.1.1 and 2.7), MUVE is likely to promote second language learning.

2.7 Multi-user Virtual Environment as an E-learning Platform

With e-Learning, students can reach their learning goals at their own convenience, as e-Learning offers rich learning experiences at a distance, and the e-Learning tools and environments are usually available for use at all times (Dewar & Whittington, 2000; David & Govindasamy, 2005).

While a range of technologies, such as on-line forums and blogs, on-line video, or video conferencing, can be used to design e-Learning experiences, the focus of this thesis is on MUVE (first introduced in section 1.1.1). In MUVE, the ability of the users to interact not only with the highly realistic representation of the physical environment but also with each other results in high degrees of immersion of the users in the environment (Dickey, 2005; Monahan, McArdle, & Bertolotto, 2008). Users, represented by their avatars, are able to act on the objects made available and, at the same time, communicate with each other and observe the outcomes of their actions (Chee, 2007; Ciekanski & Chanier, 2008).

A high degree of immersion of users in MUVE is caused by the high degree of presence. Presence is defined as the feeling or sensation of actually being inside a simulated environment, which is, in fact, not real but just computer generated (Winn, Hoffman, & Osberg, 1995). Interactions within MUVE offer users a first person type experience, which Winn (1993) related to a direct and, possibly, even unconscious feeling of presence. A first person experience should be distinguished from a third-person one, which results from interactions through intermediate interfaces that require deliberate reflection, in the form of someone else's description of the world or a computer interface that stands between the environment and the user (the interface of which the user has to be aware). Because interactions in a virtual environment offer a first person experience, one can expect that such interaction could substitute for real experience; because the learners do not need to engage in deliberate reflection, virtual environments may enable spontaneous knowledge acquisition that requires less cognitive effort than the traditional practices (Bronack et al., 2008).

A high degree of presence distinguishes MUVE from other technologies used for e-Learning. In certain circumstances, a high degree of presence hinders learning by

causing learners to interact with each other and with the environment, rather than focusing on learning tasks (as, for example, highlighted by teachers in the study by Gamage, Tretiakov, & Crump, 2011). On the other hand, the high degree of presence in MUVE makes it an environment that can offer a close approximation to an on-campus experience for distance students (Lucia, Francese, Passero, & Tortora, 2009). Researchers have been interested in the pedagogical potential and the usefulness of MUVE in enabling distance learning (Dickey, 2005). MUVE can offer opportunities for teachers and students to engage in learning related social activities, even when teachers and students are not collocated (Bronack, Riedl, & Tashner, 2006). MUVE affords a context for distance students to be present and to socialise with their peers and with teachers. The feeling of presence and being involved in social activities afforded by MUVE leads to a positive learning experience for distance students (Edirisingha, Nie, Pluciennik, & Young, 2009).

MUVE offers a powerful platform for experiential learning because collaborative interactions and role-play can be combined with high quality simulations of physical contexts and physical experiences (Boulos et al., 2007). At the same time, the concept of situated learning (discussed in detail in section 2.9.2) offers a powerful lens for appreciating the affordances of MUVE in promoting learning. MUVE offers an environment where learners are situated in a realistic social and physical context. Because learning occurs in a realistic context, the skills and knowledge acquired should transfer easily into real world settings (Jones, Morales, & Knezek, 2005).

2.7.1 MUVE as an Environment Suitable for Digital Natives—the Effect of Video Games Experience

One of the themes that emerged in a recent qualitative study of educator perceptions around the use of Second Life MUVE for teaching (Gamage et al., 2011) was an educator expectation that students who enjoy playing video games are likely to find MUVE an attractive environment for learning. Educators believed that Second Life would appeal to this student cohort and lead to greater participation of some of the students who are not active in the traditional classroom setting. The view that "digital natives"—students who have used social media and video games as they were growing up—are likely to find MUVE an attractive

environment for learning has been voiced in a number of publications (Cabanero-Johnson & Berge, 2009; Spiegelman & Glass, 2009). Therefore, the physical and the social environments provided by MUVE are likely to be particularly suitable for the coming generation of "digital natives", who grew up exposed to the technology-mediated social environments of Facebook and texting, as well as to the physical (and social) environments of video games.

2.7.2 Specific Examples of MUVE Environments

This section introduces two technology platforms that are at present most commonly used to create MUVE based educational experiences: Active Worlds and Second Life. According to a recent survey by Dalgarno et al. (2011), in 2010 the overwhelming majority of MUVE based learning projects in New Zealand and Australia used Second Life (78%), followed by Active Worlds (5%).

2.7.2.1 Active Worlds

Active Worlds is a type of MUVE technology (Activeworlds, 2006), and Viras and SciCentre are two educational virtual world environments that utilised this platform. Viras (Prasolova-Forland & Divitini, 2003) is a collaborative virtual environment developed to support social awareness by creating environment bridging groups of learners formed in the real world. SciCentre, on the other hand, relied on Active Worlds to create a virtual museum with interactive artefacts representing science concepts (Corbit, 2002). Learners, represented by their avatars, interact with the exhibits and with each other, thus realising the social constructivist learning paradigm.

In a descriptive exploratory study by (Dickey, 2005), an Active Worlds MUVE was used as a medium for synchronous and asynchronous distance learning. Dickey focused on two aspects. First, how to use Active Worlds as a distance learning tool and second, what types of learning experiences are afforded by this 3D virtual environment. His findings were that Active Worlds provided an atmosphere of experiential and situated learning within a collaborative learning environment. Dickey (2005) concluded that this atmosphere enabled distance learners to converse and to construct meaning through interactions with each other

and with the environment, and thus, the environment supported constructivist learning. Dickey (2005) suggested that the immersive capability of Active Worlds cannot be replicated in a traditional classroom environment.

2.7.2.2 Second Life

While Second Life is not the only MUVE available for implementing e-Learning, as exemplified by the experience reports cited in section 2.7.2.1, currently most of the MUVE e-Learning research and practice relies on the Second Life MUVE. This can be seen, for example, from citation counts reported by Gamage et al. (2011).

Second Life applications for e-Learning have covered topics such environmental issues (Hackathorn, 2006), science (Cochrane, 2006; Doherty & Rothfarb, 2006), health and medicine (Boulos et al., 2007), multimedia design (Cargill-Kipar, 2009), combating sexual abuse of children (Feldman, 2006), instructional design issues and learning management (Doherty & Rothfarb, 2006), and urban planning (Mallan, Foth, Greenway, & Greg, 2010). These examples illustrate the diversity of Second Life e-Learning applications. For an analysis of the number of articles devoted to the use of Second Life for e-Learning, refer to Gamage et al. (2011). Further examples of MUVE applications for e-Learning (specifically, for language learning) are given in section 2.10, and involve the uses of Second Life, as well as of other MUVEs.

2.8 MUVE versus Other Technologies: Technology Trends in e-Learning

This section presents an overview of trends in technology in education and e-Learning, conducted to better understand the place of MUVE based e-Learning in the learning technology landscape.

The overview relies on an approach similar to the approach taken by Martin, Diaz, Sancristobal, Gil, Castro, and Peire (2011) in their recent article in Computers and Education titled “New technology trends in education: Seven years of forecasts and convergence”. Bibliometric analysis was conducted by using the Google Scholar search engine (scholar.google.com), with the number of hits returned for a

particular keyword interpreted as the influence of the technology associated with the keyword.

More specifically, the search was conducted by setting the "with all the words" field in advanced Google Scholar search interface to a combination of the keyword "education" with the keyword describing the technology of interest. For example, for wikis, the field was set to "education wikis" (see Figure 2). The "with at least one of the words" field was set to "student learner" (hence, only the articles specifically mentioning a "student" or a "learner" were included). The "return articles between" field was set to a year in the range between 2008 and 2012. For technologies described by more than one word, such as "social networking", these words were entered into the "with the exact phrase" field, rather than in the "with all of the words" field (but leaving the keyword "education" in the "with all of the words" field). Some of the articles returned by each search were examined to confirm that the content is, indeed, relevant to using the target technology for e-Learning; however, because most of the searches returned very large numbers of hits, it was not feasible to inspect all of the articles, even briefly. The approximate number of hits returned by Google Scholar was interpreted as the influence of the particular technology in e-learning research in a particular year. The searches were conducted on April 4 2013.

The absolute number of hits returned by Google Scholar for a particular keyword is difficult to interpret, because the keyword may be used in a variety of contexts (for example, in an article fully devoted to the topic suggested by the keyword or in an article just briefly mentioning the keyword). Therefore, similar to Martin et al. (2011), in the analysis presented in this section the absolute values of the numbers of hits are not interpreted, but rather the interpretation is based on the ratio of the number of hits for a particular keyword for a given year to the number of hits returned for the same keyword for the year 2008. Therefore, only the rates of growth of interest in the technologies are compared.

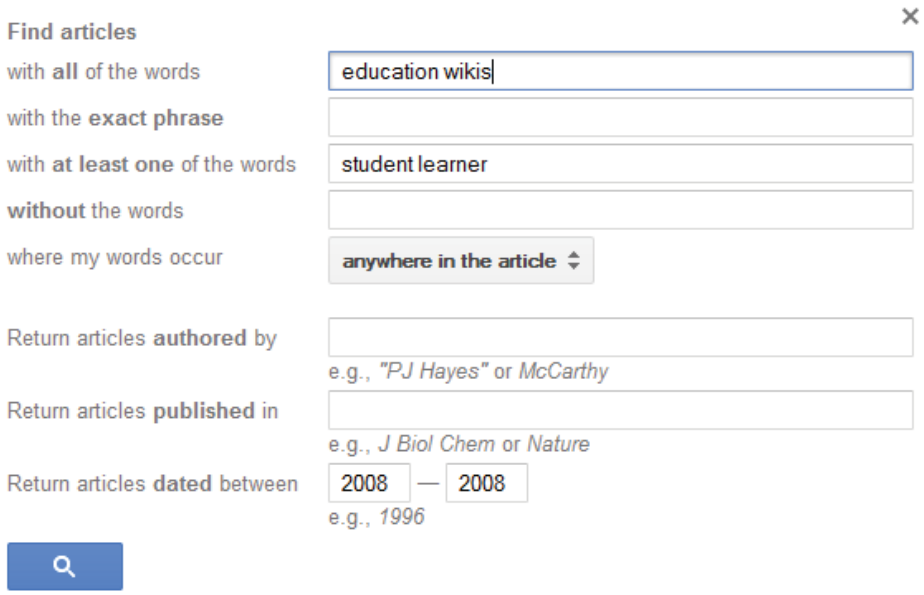

Figure 2. Advanced Google Scholar query used to assess the influence of topics devoted to wikis in 2008.

The technologies considered were wikis, blogs, and social networking (grouped together as social web technologies); learning objects and open content; augmented reality; mobile devices and ubiquitous learning; and virtual worlds and computer games. Overall, the technologies covered were similar to the ones covered by Martin et al. (2011), with one exception—semantic web was not covered because it was judged to be less related to virtual worlds and MUVE.

The rest of this section is organized as follows. Each subsection is devoted to a group of related technologies. The technologies are briefly introduced, which is followed by examples of the research on their use in e-Learning drawn from recent literature. Then, the results of the bibliographic analysis are presented.

2.8.1 Social Web

In the context of e-Learning, Social Web refers to technologies intended to enhance collaboration and communication (Martin et al., 2011). Blogs and wikis are specific Social Web technologies.

A blog is a sequence of articles (incorporating text and images), with the readers of the articles posting comments (for example, www.blogger.com is a system supporting blog creation made available by Google). Usually, an easy to use web

based interface is provided for posting articles and for commenting (although, it may also be possible to post blog entries from inside desktop applications, such as Microsoft Word).

Wikis allow multiple users to edit the same web based document, with version management functionality available to undo unintended or inappropriate updates (for an example of a system supporting wikis, refer to www.wikispaces.com, made available by Tangient; the most widely known wiki is Wikipedia, www.wikipedia.org—a free, collaboratively maintained online encyclopedia from Wikipedia Foundation). The shared document is presented as a web page, and the wiki functionality is normally provided via a web based interface.

Social networking (Ahn, Han, Kwak, Moon, & Jeong, 2007; Glaser, 2007) is a Social Web technology that allows a user to record information about the user and to define the members of the user's social network. Then, information about the user can be made available to the members of the social network. Some major examples of social networking systems are Facebook (www.facebook.com) and Google Plus (plus.google.com). The exact functionality made available around the representation of the users' social network in a social networking system varies, and is constantly evolving. Usually, similar to blogs, users can post articles (or messages), which are made available to the member of the users' social network, or, depending on the options set when the article is posted, to all Internet users, and, similar to blogs, the users viewing the posted articles can post comments.

Blogs, wikis, social networking, and other technologies enabling Internet users with no technical expertise to create and to share digital content and to collaborate with other users online are known as Web 2.0 technologies, a term that was coined by O'Reilly (O'Reilly, 2007). O'Reilly listed seven characteristics of Web 2.0 services, the characteristics that are particularly important from an e-Learning perspective are the first two in the O'Reilly's article: using the Web as a platform and harnessing collective intelligence. In terms of e-Learning one can interpret these characteristics as offering access to e-Learning materials and functionality anytime, anywhere and enabling the collaboration and interactions between the learners and between learners and teachers resulting in knowledge construction (in the sense of social constructivism).

Chu, Chan, and Tiwari (2012) reported the use of blogs to support learning during internship in information management and in nursing; students used blogs to share on their experiences and to comment on the experiences of others. Based on structured interviews with the participants and on the analysis of blog entries, it was found that students viewed the experience positively and engaged in a social-collaborative learning process by interacting via posting blog posts and comments. At the same time, it was found that student behaviour in blogging was influenced by the grading system and by instructor expectations.

Biasutti and El-Deghaidy (2012) reported the use of wiki to facilitate the design of interdisciplinary didactic projects by teacher education students. Based on a survey of the participants and on analysing the wiki content, it was found that the students were satisfied with the use of wiki and that the wiki helped to develop knowledge management processes.

Lin, Hou, Wang, & Chang (2013) described the use of Facebook as a platform to support online interactions by students of Appreciation of Arts and Design Aesthetics. Quantitative content analysis was used to analyse student posts. It was found that the most prominent knowledge dimension in the discussion (according to the Revised Bloom's Taxonomy, Anderson & Krathwohl, 2001) was the meta-cognitive dimension (knowledge about cognitive tasks, including the relevant contextual knowledge). Lin et al. (2013) noted that female students were more likely to engage in off-topic discussions that male students.

All of the three studies agreed in suggesting that Social Web technologies promote learning by enabling communication and collaboration between learners. The uses of blogs, wiki, and social networking applications to facilitate collaborative learning are highly consistent with the social constructivism learning paradigm (Cole, 2009). From the perspective of e-Learning researchers, the captured content of interactions between learners offers opportunities for detailed analysis of knowledge construction activities, including analysis by using quantitative methods, as in the study by Lin et al. (2013).

As seen in Figure 3, of Social Web technologies, the popularity of social networking in e-Learning research has been growing the fastest, followed by blogs, which enjoyed a steady growth. The popularity of wikis has been growing up to 2010, but then plateaued.

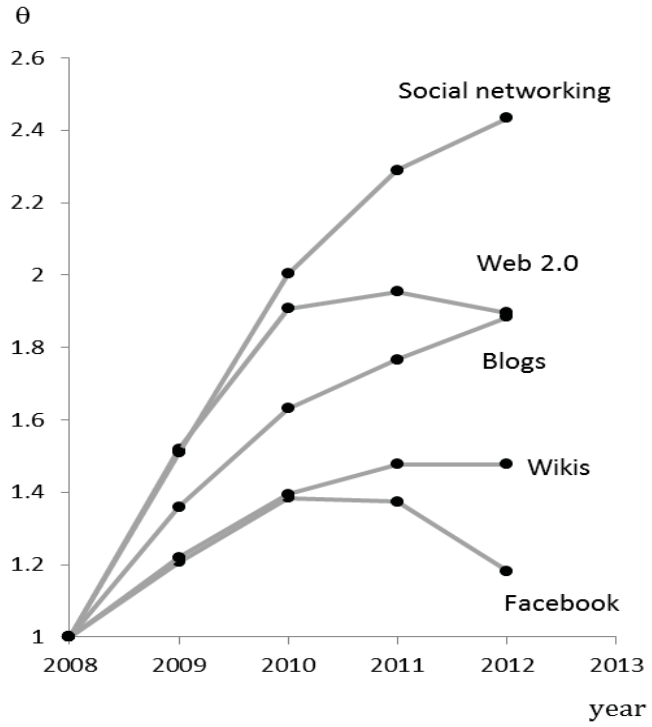

Figure 3. Popularity (θ) of Social Web learning technologies measured by the number of hits in Google Scholar relative to the number of hits for 2008.

2.8.2 Learning Objects

Learning objects (Mohan & Brooks, 2003) are reusable digital artefacts that can be used to enhance learning. Learning objects are not a technology per se, but rather a way to package digital content emphasizing its reusability. The IEEE 1484.12.1 – 2002 Standard for Learning Object Metadata (LOM) is an open standard for packaging digital content into leaning objects (http://ltsc.ieee.org/wg12).

Schibeci, Lake, Phillips, Lowe, Cummings, & Miller (2008) conducted a series of interviews of teachers and students to study the use of learning objects in Australian and New Zealand schools (learning objects, formatted in compliance with the LOM standard, were made available by the researchers, with the research conducted as a part of an initiative by Le@rning Federation, http://elearningasinquiry.tki.org.nz). The interview transcripts were analysed as qualitative data. Teachers found learning objects to be useful and intellectually stimulating, and wished to have access to more learning objects. At the same time, it was found that because of the ease with which learning objects could be

deployed, teachers sometimes used learning objects that were ill suited to the cognitive abilities of their students. Moreover, some of the learning objects inspected contained subject matter inaccuracies.

Highlighting the issue of intellectual property associated with the transfer of learning objects, Santos and Ramos (2004) proposed a framework for digital rights management of learning objects. Boyle (2010) proposed a multi-layered model for combining learning activities with learning objects intended to facilitate combining learning designs and to allow different designers to focus on different levels of abstraction in the process of learning design.

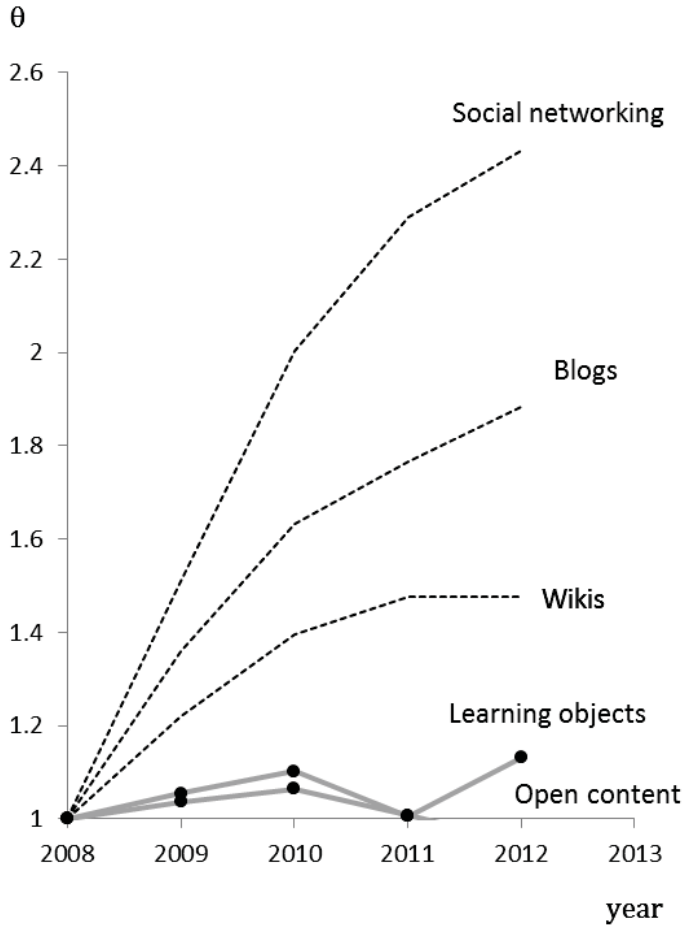

Figure 4. Popularity (θ) of learning objects measured by the number of hits in Google Scholar relative to the number of hits for 2008. Social networking, blogs, and wikis are included to enable comparison.

The study by Schibeci et al. (2008) demonstrated that the ease of deployment of learning objects might make them attractive to teachers; however, research providing a direct comparison of the use of LOM based learning objects (or any

digital content packaging technology) with e-Learning not relying on LOM is not available. Moreover, there is no evidence of the framework by Santos and Ramos (2004) and of the model by Boyle (2010) being used in practice.

As seen in Figure 4, the numbers of hits on Google Scholar suggest that the popularity of learning objects in e-Learning research did not grow over the last five years.

2.8.3 Augmented Reality

Augmented Reality (AR) refers to a technology involving overlaying in real time computer generated video and sound over video and sound representing the real world (Azuma, 1997). When experiencing augmented reality, one perceives computer content as part of the real world. An example of an augmented reality device that was highly prominent in the media at the time of writing this thesis was Google glasses (www.google.com/glass/start/how-it-feels).

Di Serio, Ibáñez, & Kloos (2012) reported the use of augmented reality to enhance paintings used in a visual arts class with annotations relevant to the course content. Student motivation was assessed by using a survey instrument; by comparing to students who did not use augmented reality, it was found that the use of augmented reality increased student motivation. Kamarainen, Metcalf, Grotzer, Browne, Mazzuca, Tutwiler, & Dede (2013) reported the use of augmented reality to provide directions and scaffolding to students on a field trip to study a pond ecosystem. Students' attitudes and student comprehension of subject matter were measured; students felt positively about the experience, and their knowledge of the subject matter has improved. Thus, both of the studies agreed that augmented reality promotes learning by enhancing learner experiences.

Wu, Lee, Chang, & Liang (2013) formulated a framework conceptualizing the educational benefits of augmented reality e-Learning by dividing augmented reality e-Learning applications in three categories: emphasizing roles (with computer generated content overlaid over the real world informing the learner about their role in the educational experience), emphasizing locations (with computer generated content providing information about the locations in the real world), and emphasizing tasks (with computer generated content providing

instructions). In terms of this classification, one can argue that the augmented reality e-Learning system in the study by Di Serio et al. (2012) can be classified as emphasizing locations, and the system in the study by Kamarainen et al. (2013) as emphasizing tasks. (Augmenting reality systems emphasizing roles appear to be less common—Wu et al. gave an example of an educational game simulating infection, with the infection status of an individual conveyed via computer generated content.) Wu et al. also highlighted the drawbacks of using augmented reality as students can be cognitively overloaded by the large quantities of information (experiencing the real world and the computer generated content at the same time) and the complex devices they are required to use, which may interfere with functioning in the real world (such as special glasses).

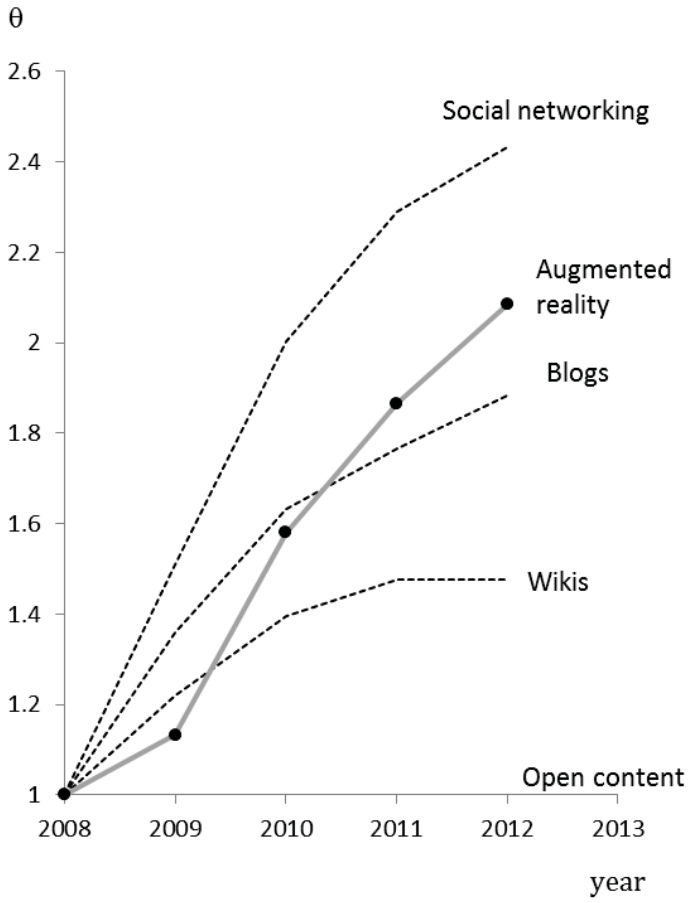

Figure 5. Popularity (θ) of augmented reality measured by the number of hits in Google Scholar relative to the number of hits for 2008. Social networking, blogs, and wikis are included to enable comparison.

As seen in Figure 5, the numbers of hits on Google Scholar suggest that the popularity of augmented reality in e-Learning research enjoyed very fast growth in the last four years, comparable to the grows in the popularity of social networks.

2.8.4 Mobile Devices

Mobile devices are portable devices with computing and digital communication capabilities, such as smartphones and tablets. Mobile devices offer ubiquitous learning capabilities (Caballé, Xhafa, & Barolli, 2010; Wong & Looi, 2011)—learners have access to the functionality they provide no matter where they are (e.g., at home or while commuting). Moreover, modern mobile devices not only offer functionality similar to conventional personal computers (such as email, Internet browsing, and document storage and viewing), but also functionality specific to mobile devices, such GPS (global positioning system) and accelerometers. Augmented reality can be implemented inexpensively by using mobile devices with digital cameras and screens—device generated content is mixed into the video signal and the resulting video is shown on the device's screen (Wagner, Reitmayr, Mulloni, A., Drummond, & Schmalstieg, 2010).

Wang, Wiesemes, & Gibbons (2012) conducted a study of using mobile devices to enable ubiquitous learning. Part-time mature doctoral students' used mobile devices in support of their doctoral studies. The students used the devices mainly to store and to view documents and to record ideas as audio; the students had all sort of technical problems using the devices, such as problems with WiFi connectivity. Based on qualitative analysis of the transcripts of semi structured interviews of the participants, it was concluded that mobile devices offered learning flexibility to the students, who had continued access to content relevant to their studies and were overall positive about the experience; at the same time, learning to use the mobile devices productively was an issue, particularly for the older students.

Furió, GonzáLez-Gancedo, Juan, Seguí, & Costa (2013) compared the use of mobile devices of different size and weight to conduct an educational game devoted to water cycle, water composition, and water pollution. The game made use of the device's touch screen and accelerometer capabilities and included

augmented reality features (but no collaborative learning features). Based on a survey of the participants, it was found that the device's size and weight did not affect the acquired knowledge, engagement, satisfaction, ease of use, or AR experience.

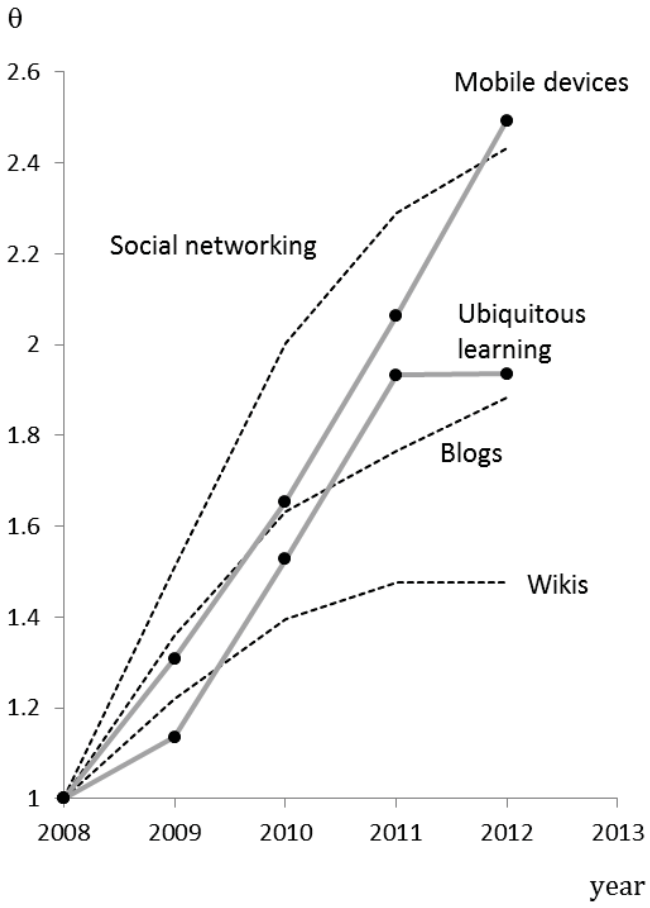

Figure 6. Popularity (θ) of mobile devices measured by the number of hits in Google Scholar relative to the number of hits for 2008. Social networking, blogs, and wikis are included to enable comparison.

By synthesising the results by Wang et al. (2012) and by Furió et al. (2013), one can conclude that the ease of use of a mobile device is more important than the device's physical characteristics for the device to be suitable for e-Learning.

Wu, Wu, Chen, Kao, Lin, & Huang (2012) conducted a systematic literature review of mobile learning studies and concluded that the most highly cited articles devoted to mobile e-Learning focus on mobile system design rather than on mobile e-Learning use in practice. This may reflect the emerging nature of mobile learning—mobile learning going beyond just using generic mobile device

functionalities (mobile learning as in the study by Furio, rather than mobile learning as in the study by Wang) is yet to become a mainstream approach to e-Learning.

As seen in Figure 6, the numbers of hits in Google Scholar suggest that the interest in mobile e-Learning has been growing very fast, even faster than the interest in social networking, and, unlike the interest in social networking, did not show signs of slowing down over the last two years.

2.8.5 Educational Computer Games and Virtual Worlds

As introduced in section 1.1.1, virtual worlds (or MUVE—multiuser virtual environments) present to their users a representation of a simulated physical environment and of the users themselves (as avatars). In virtual worlds, users are free to interact with each other and with the physical environment—there are no goals or scenarios built into the environment. Some of the computer games are similar to MUVE in offering the representations of a physical environment and of the users, but, unlike in MUVE, in computer games the users are expected to pursue specific goals and scenarios programmed into the environment (in educational computer games, such goals may be associated with achieving learning outcomes) (Charles, Charles, McNeill, Bustard, & Black, 2011). More broadly, computer games are interactive environments in which users pursue goals, in interaction with other users or not. Both MUVE and computer games can be used to implement e-learning. While in educational computer games the educational experience may be shaped by the environment, in MUVE the interactions have to be shaped by instructors and by the learners themselves.

Examples of educational uses of MUVE are presented in sections 2.10 and 2.11 of this thesis. Therefore, this section focuses on the uses of educational computer games.

Ke (2013) examined the potential of using computer games as an anchor for tutoring mathematics, creating a learning experience that at the same time is perceived by children as a game and results in achieving educational outcomes. The approach was trialled with middle schoolers at a low performing urban school and at a rural school. Based on observations, interviews of the participants, and on

the results of pre and post tests of subject matter knowledge, it was found that there was a high level of student involvement; however, after 10 gaming sessions, only the students at the rural achieved statistically significant learning gains.

Meluso, Zheng, Spires, & Lester (2012) compared collaborative game learning and single player game learning in terms of knowledge gains and in terms of subject matter self-efficacy. The game was based on narrative-centred learning, with the learners having to interact with in-game characters to gain science related knowledge needed to complete the game scenario. The participants were middle school pupils. Based on pre and post tests, no statistically significant difference was found between collaborative learning and single player learning gain, even though there were statistically significant increases in both subject matter knowledge and in self-efficacy when aggregated pre and post scores are compared.

Overall, the results by Ke (2013) and Meluso et al. (2012) suggest that computer games may be effective in promoting learning, but not necessarily for all types of learners.

As seen in Figure 7, the numbers of Google Scholar hits suggest that the popularity of computer games has been growing, but very slowly (slower than the popularity of wikis). Similar to wikis, the interest in computer games in e-learning research plateaued over the last two years.

As to the popularity of virtual worlds, even though it has been growing fast around 2009 (when the data for the present study was collected), is plateaued, and even somewhat declined, over the last three years. The number of hits for the keyword MUVE has been falling very fast over the last three years, possibly reflecting a shift from using the rather technical term MUVE to the more colloquial “virtual worlds”.

The rest of this section interprets the implications of the technology trends suggested by Figure 7 for the significance of the present study.

Social networking, blogs, and wikis (as seen from the studies introduced in this section) support social constructivist approach to teaching and learning (Cole, 2009), and in this respect are similar to virtual worlds. The rising interest in the use of these Social Web technologies for learning suggest that the pedagogical principles underlying MUVE based learning are gaining in influence.

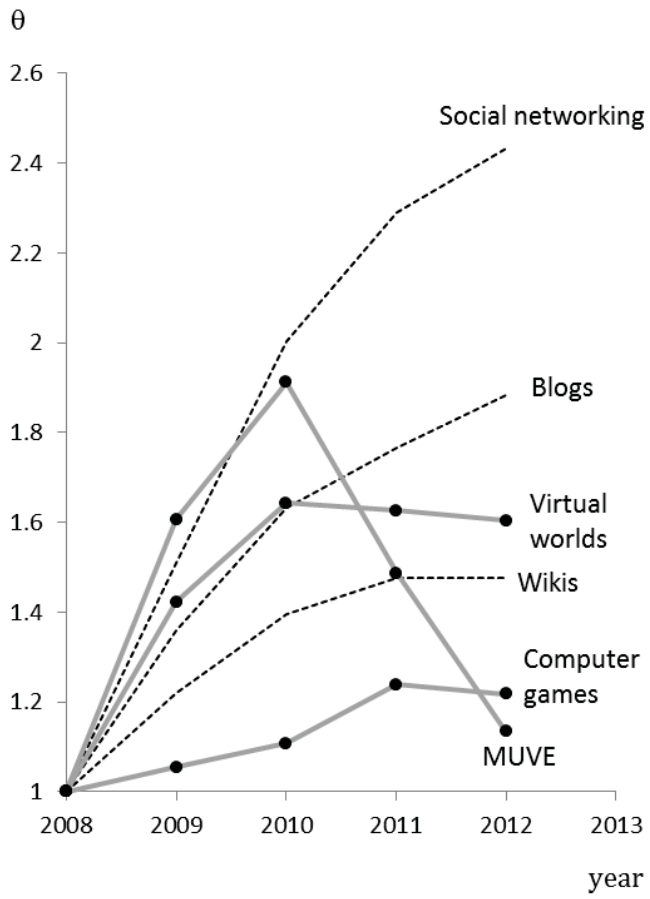

Figure 7. Popularity (θ) of virtual worlds and computer games measured by the number of hits in Google Scholar relative to the number of hits for 2008. Social networking, blogs, and wikis are included to enable comparison.

The current learning objects metadata (LOM) standard for reusable content (briefly described in section 2.7.2) does not cover MUVE content. Yet, reusability of artefacts created in MUVE (such as avatar appearance) and their transferability is an important topic for MUVE in general and for MUVE based e-learning in particular (see http://www.metaversestandards.org). Creating high quality e-Learning content for use in MUVE is a considerable investment, and having to depend on the continued availability of a particular MUVE environment (such as Second Life) may prevent organization from investing in MUVE based e-Learning. Thus, in the medium term, the apparent lack of energy in the research on learning objects is somewhat of concern for the future of MUVE as an e-Learning platform.

Augmented reality is similar to MUVE in terms of exposing the learners to artificial three dimensional artefacts. One can expect that in the future augmented

reality technology and virtual worlds technology are going to merge, enabling different extents of mixing between the real world and the virtual world (a vision behind the Mixed Reality Project, http://www.architecturemixedreality.com). Thus, the growing interest in augmented reality suggests a growing interest in virtual worlds in the future.

It is likely that the current interest in mobile e-Learning is driven by the increased availability of mobile devices, rather than by pedagogical principles suggesting their use. In principle, mobile devices can be used to enable a broad range of educational experiences, from rote memorisation via virtual flash cards (as offered by the Flachcards Deluxe product, see http://orangeorapple.com/Flashcards) to complex group learning based on social constructivism principles, with the interaction between group members, as described in the study by Cochrane and Bateman (2010). Currently, the readily available virtual worlds technology (reviewed in section 2.7.2) relies on desktop computing. With the growth in the power of mobile technology, virtual worlds are likely to become accessible via mobile devices (and indeed, availability of virtual worlds via mobile devices may facilitate the convergence between virtual worlds and augmented reality). Moreover, the distinction between virtual worlds and computer games is already somewhat uncertain, as games can be implemented inside a virtual world (see http://secondlife.com/destinations/games), and therefore MUVE can be used as platforms for implementing educational computer games.

As to the apparent lack of growth in the interest in virtual worlds over the last three years, it is likely to be associated with the particular technology used to implement virtual worlds, namely the desktop virtual reality used in virtual world products such as Second Life and Active Worlds. It is likely that the advances in mobile technology and in augmented reality will lead to new generations of virtual and mixed worlds, accessible via products such as Google glasses. Therefore, the study of the effectiveness of teaching in virtual worlds (which is not strongly associated with a particular way of implementing virtual worlds) remains important.

2.9 Learning Paradigms Relevant to MUVE

Learning theorists suggest that developing courses with new and emerging technologies promotes a shift from rote learning to constructivist learning. Technology can influence how learners approach learning. Previous research claims that the traditional (instructional) methods can be supplemented or even replaced by constructivist learning methods when using new technology, such as MUVE (Dede, 1995; Dickey, 2005; Antonacci & Modaress, 2008).

The constructivist paradigm emphasises the active role of learners in acquiring knowledge, which is viewed as knowledge construction (Piaget, 1954; Mayer, 1987; Kalina & Powell, 2009). In this research, constructivism is seen as a learning paradigm that is most appropriate for the MUVE environment. Situated learning is a form of constructivism that is particularly relevant to teaching ESP, including oral business English communication skills. The situated learning approach emphasises the immersion of learners in realistic social and physical environments, which in the case of ESP could match the specific application areas targeted by the instruction. One way to implement situated learning is by designing broad scenarios involving real-world-like activities in a real-world-like environment (and MUVE can provide such an environment including any artefacts and social encounters needed for the activities). The rest of this section details the constructivist paradigm and the situated and scenario-based approaches to learning, in view of their implementation in MUVE.

2.9.1 Constructivism

Constructivism is a term found in the philosophy of science and in education (Philips, 1995). In the philosophy of science, constructivism refers to the process of construction of scientific knowledge by the society. This aspect is not relevant to this thesis, which, henceforth, uses the term constructivism as it is understood in the field of education. In education, constructivism refers to a theory positing that learners construct their knowledge. When the focus is on an individual, constructivism suggests that individual learners build their own mental models, rather than receive knowledge from instructors. The mental models built by individual learners are anchored in their prior knowledge and experiences, and,

thus, are not the same. The role of an instructor, then, is to assist learners in their knowledge construction, rather than to ensure that "correct" knowledge is transferred. An alternative interpretation of the term constructivism in education puts focus on learning in a group, so that knowledge is constructed in a collaborative manner, in negotiation between learners with different levels of proficiency in the subject of learning. This view is known as social constructivism and is often attributed to Vygotsky (1978), while the constructivist view of learning focusing on an individual is often attributed to Piaget (1980). In education research and practice, it is common not to distinguish between Vygotsky's social constructivism and Piaget's constructivism but to use the term 'constructivism' either to refer to both views, or to refer to social constructivism which is particularly influential in education as done in the following works (Moreno, Gonzalez, Castilla, Gonzalez, & Sigut, 2007; Huang, Rauch, & Liaw, 2010).

Constructivism is defined as an approach where learners construct new knowledge based on their prior knowledge through active learning when immersing in and embracing a new learning experience (Vygotsky, 1978; Kalina & Powell, 2009). Learners reflect on the new knowledge and articulate their thoughts about it through sharing and negotiating with other learners. Sharing and negotiation foster collaboration and support social negotiation of meanings among learners (Lave & Wenger, 1991; Jonassen, 1999). This, in turn, allows learners to test understandings and to change their constructions based on feedback (Duffy & Cunningham, 1996; Jonassen, 1999). This learning paradigm can be contrasted with objectivism, according to which knowledge is transferred to learners via drills and conditioning (Mayer, 1998).

In constructivism, the role of an instructor changes from defining and conveying knowledge to acting as a facilitator of learning. The instructor guides and supports learners in the process of constructing new knowledge. The guidance and support from the instructor to a learner fades away as the learner is increasingly able to learn, understand, reflect, and apply the knowledge independently (Huang, 2002; Agostinha, Meek, & Herrington, 2005; Cheung, 2006).

A learning environment based on the constructivist paradigm must be authentic, personally relevant to learners, challenging and interesting to learners, and must provide a physical representation of the real-world task environment. In such a

simulated environment, learners engage in collaborative activities with other learners as well as with the instructor, who acts as a facilitator and guide. Through working in groups, learners would be able to tap into other learners' skills and use a variety of tools and information resources while pursuing learning goals and participating in problem solving activities (Jonassen, 1999).

In his highly influential work (with the citation count on Google Scholar close to 15000), Vygotsky (1978) asserted that social interaction guides learner thinking and enhances the cognitive development process by encouraging learners to ask relevant questions to check their understanding on particular subject matter. He further added that interactions promote learning. According to him, when learners take part in the activities that are slightly more challenging than ones already mastered, their thinking skills will be further developed. Moreover, their thinking skills would also develop when they are interacting with others who are more knowledgeable than themselves. In terms of instructional scaffolding, Vygotsky's theory of zone of proximal development stressed the requirement to provide learners with the appropriate learning environment, which would offer challenges, interaction with peers and instructors, and the appropriate level of support that is gradually removed.

Supyan (2008) relied on Vygotsky's theory of zone of proximal development when conducting research on the use of online forums for e-Learning. He reflected on the nature of the relationship between technology and constructivism, and he agreed with the view of Duffy and Cunningham (1996) that technology serves as a tool to expand cognition by providing learners with appropriate learning support.

2.9.2 Situated Learning

Situated learning is an approach to implementing constructivist learning relying on using authentic social and physical contexts relevant to learning goals, which assist learners in constructing knowledge (Huang et al., 2010). Brown, Collins, and Duguid (1989) and Lave and Wenger (1991) defined situated learning as learning occurring when learners are embedded within a relevant activity. In situated learning, learners are not separated from participating in the practice they aim to learn. The more learners are separated from the social and physical context of the

targeted practice, the less valuable the knowledge they gain is (Duffy & Cunningham, 1996; Alessi & Trollip, 2001).

Unlike the linear approach of instruction in a classroom, in the situated learning approach, learning occurs naturally and is practise-oriented. Namely, learners simultaneously gain knowledge and acquire the ability to apply it (Hung & Chen, 2002).

The situated learning approach also encourages collaboration among learners—learners interact as members of a community sharing common interests, with instructors also forming part of the community (Duffy & Cunningham, 1996; Wilson & Cole, 1996; Henning, 1998). Thus, to an extent, learners themselves provide some of the social context in which the learning is situated. The relationships that exists among the members of the community tend to be peer-based, rather than formal teacher-student relationships as in the traditional classroom. As members of the community, learners participate in learning tasks throughout the period of instruction, and as their knowledge and skills increase, the role and status of a learner as member of the community gradually evolves from that of a novice or an apprentice to that of an expert (Lave & Wenger, 1991; Hislop, 2003). Situated learning relies on modelling (target performance demonstrated by the instructor or by learners who already achieved expert status), collaboration (negotiation of meaning and joint practice), exploration (perception of the context and experimentation with the context), and scaffolding (help and direction from the instructor that is gradually removed).

When experts or practitioners participate in the community of practice in which the situated learning is embedded, they can transfer their tacit knowledge to the learners. Tacit knowledge is knowledge that cannot be fully articulated and explained, and can only be transferred when learners engage in real-world activities along with the experts (Nonaka, 2007). Thus, situated learning can achieve learning goals that cannot be achieved in traditional (instructional method) classrooms because situated learning facilitates the acquisition of tacit knowledge (Brown et al., 1989).

2.9.3 Scenario-based Learning

Scenario-based learning is a kind of situated learning (Brown et al., 1989). In scenario-based learning, learner interaction with the social and physical context in which the learning is situated is structured via loosely defined scenarios in which learners are expected to engage.

Scenario-based learning can be seen as role-play—learners assume roles suggested by the scenarios. In this style of learning, learners can be assessed based on their performance in their assigned scenario roles (McLellan, 1986). Thus, the assessment reflects more accurately the students' likely performance in the job roles corresponding to the scenario roles than traditional classroom based assessment.

In terms of language learning, the use of role-play or simulation within a situated learning framework enables learners to rehearse the required range of language functions necessary for a particular role in a safe environment. Moreover, the use of role-play or simulation is claimed to be an effective technique to develop communicative competence, as it encourages learners to actively explain, elaborate, or defend ideas on a particular matter. Therefore, the approach is believed to enhance learner understanding of appropriate expressions and prepare them to use those expressions in real world environments (Krashen & Terrell, 1983; Shi, Corcos, & Storey, 2001).

Yan's study (2006) described an example of using the scenario-based learning approach in language learning. She adopted scenario-based learning to help improve poor communicative skills in oral and written English among some Chinese learners. Learners were involved in authentic activities that stimulated their creativity while speaking and writing. Using scenario-based learning, she was able to engage and manage a large class, with all students engaged in language production. Group work enabled all learners to be involved in activities and reduced shyness among the learners who were reluctant to speak in front of the whole class in a traditional setting**.**

2.10 Applications of Multi-user Virtual Environments to Promote Language Learning

There are a number of studies describing the use of MUVE for language learning. Peterson (2005) conducted an observational study that demonstrated the opportunities virtual world learning environments created to engage learners in the kind of interaction that facilitates the development of second language (L2) competencies. Using Active Worlds as a platform, he discovered that through using the communication features of avatars, learners employed both interactional strategies (such as small talk and apologies) and transactional strategies (such as abbreviations, acronyms, and clarification requests) when interacting in the virtual learning environment, resulting in rich, unimpeded communication.

In another study using the same type of MUVE (Active Worlds), Peterson (2006) provided further evidence that the use of avatars in MUVE facilitates non-native speaker interaction in the target language; once again, both interactional and transactional strategies were observed. The study also reported the use of avatars for negotiation of meaning focusing on new vocabulary.

Wang, Song, Xia and Yan (2009) reported Chinese students' perspectives of an English as Foreign Language programme in Second Life. Participants from both a state university in the United States of America and a major provincial university in Eastern China were involved in the study, with the US students offering the students in China an opportunity to interact with native speakers. (For the US students, the study was part of their coursework requirements in a technology course.) Both qualitative and quantitative data (pre-program and post-program surveys, interviews, and documentation) were collected from 61 participants from China. It was found that the Chinese students perceived Second Life as a useful and interesting language learning platform, and that they were prepared to use Second Life for English language learning.

In MUVE, students can assume the roles of virtual characters by changing the appearance of their avatars. This masks their true identities and, thus, reduces anxieties associated with face-to-face interactions—anxieties that are common in traditional classroom settings. In this way, MUVE helps to create an environment

that is more forgiving when foreign language students make mistakes during attempts to communicate with others (Rankin, Gold, & Gooch, 2006).

Similar findings on foreign language students who are struggling to develop their target language proficiency due to inhibitions towards using the new language were shared by other researchers. According to these researchers, MUVE can provide a kind of support to anxious students learning the language. MUVE acts as a scaffolding element to these anxious students by giving them greater confidence to experiment with language in the virtual environment compared to the classroom environment (Bradley. & Lomicka., 2000; Roed, 2003).

Besides reducing inhibitions, MUVE is also believed to have the ability to provide ample opportunities for students to practise, improve, and test their emergent communicative abilities in an authentic environment. MUVE grants a social interactive environment that makes it possible for people of different ethnicities and cultures to learn a language together. MUVE can be a social space involving people from diverse groups, residing all around the world, meeting and communicating with one another; therefore, MUVE can offer L2 students access to communication opportunities in the target language (Jakobson & Taylor, 2003; Steinkuehler & Williams, 2006).

2.11 MUVE e-Learning Effectiveness Studies

e-Learning effectiveness studies are quite common. Some studies employ true experimental design, involving experimental and control groups, and random assignment of participants to groups. For example, a recent study by El-Deghaidy and Nouby (2008) found that the use of a blended e-learning cooperative approach resulted in learning gains in science that were statistically significantly better than the learning gains in a traditional classroom. Another study, (Kuo, Shadiev, Hwang, & Chen, 2012), found that learning gains by L2 students using speech to text recognition technology were statistically significantly better than learning gains by students in a traditional classroom. Both of these studies relied on relatively small groups (approximately 20 students). Other studies claiming to investigate e-Learning effectiveness rely on less robust quasi-experimental designs. For example, Abdous, Facer and Yen (2011) examined the relationship

between the self-reported use of podcasting by L2 students and the student grades. While a significant positive relationship was found, it is not clear if using podcasts caused students to get better grades. An alternative hypothesis that diligent students used podcasts more, and also obtained better grades (due to their diligence, rather than due to the use of podcasts), also explains the outcome and cannot be ruled out.

Providing a literature review for e-Learning effectiveness studies in general is outside the scope of this study. Rather, to address the first research question of this study (see section 1.4), the existing studies of MUVE effectiveness involving objective systematic comparisons with face to face teaching were reviewed. There were only two studies that examined the effectiveness of MUVE in promoting learning by using objective measurements, and these are discussed in the following subsections.

2.11.1 Learning Gains in a Virtual Lab

Cobb et al. (2009) examined the learning gains in, and the students' perceptions of, a learning module using a virtual biosciences laboratory developed in Second Life. Specifically, the purpose of the study was to determine whether conducting a molecular biology experiment called Polymerase Chain Reaction (PCR) in Second Life produced similar learning gains to those observed in a real-world world laboratory. Participants were allocated into two groups. The first 50 students to arrive made up the experimental group using a virtual lab in Second Life, while the rest of the students were the control group. The 35 students in the control group viewed a demonstration of the PCR experiment by a member of teaching staff, while students in the experimental group conducted a simulated experiment in Second Life (Cobb et al. did not report the exact duration of the instructional intervention; however, from the nature of the intervention, it appears that it was brief, close to the duration or a typical lecture). Henceforth, this stage of the experiment is referred to as "treatment". As students in the experimental group performed, in the virtual world, activities according to a scenario, the intervention can be described as scenario-based learning. Then, both groups conducted the experiment in a real lab. Learning gains were measured via pre-, mid-, and post-tests administered to the students as in-class quizzes before the treatment (the pre-

test), after the orientation (in Second Life for the experimental group and as a lecture-based orientation for the control group) (the mid-test), and after both groups of students conducted the experiment in a real lab (post-test). In addition, the number of questions asked by students while conducting the experiment in the real lab was counted, with the assumption that students who asked fewer questions had better learning gains during the treatment stage.

Both groups achieved statistically significant learning gains (at a strict level of significance, $p < 0.001$). However, the difference in learning gains between the two groups was not significant. Unfortunately, the article does not provide a p value for the comparison, but by considering other tests reported in the article and the common practice, it appears to be likely that they found $p > 0.05$.

As students were not allocated to the experimental and to the control group at random, the study should be described as a quasi-experiment, rather than a true experiment. The design favoured Second Life, because the students who arrived early are likely to be more motivated and, thus, likely to perform better irrespective of the environment. Indeed, the authors noted that there was a significant difference ($p < 0.001$) between the pre-test scores in the two groups, with the experimental group getting higher pre-test scores.

The use of quiz scores as measures of knowledge levels may have favoured the students in the control group. One would expect that active, situated learning in Second Life should lead to students being better prepared to actually engage in the target practice than students who were involved in passive learning. On the other hand, the argument that students involved in situated learning should memorize facts better than students who attended a lecture is considerably weaker. Sixteen of the 50 students in the experimental group used Second Life before the study and the treatment included an orientation session in which the students learned to use Second Life. Students in the experimental group with no Second Life experience may have been at a disadvantage, as they had to learn the subject matter and the Second Life environment at the same time, while students in the control group used a familiar learning mode (lecture) for a familiar purpose (to achieve a good score in a paper and pencil quiz).

Students in the experimental group asked significantly fewer questions while conducting the experiment in the real lab ($p < 0.001$). This can be viewed as

evidence that situated learning in Second Life resulted in the acquisition of knowledge that was more transferable to the relevant real-world situation. However, this difference may have been merely due to the students in the experimental group being better motivated and organized (because of the way participants were allocated to groups).

2.11.2 Learning Gains in a Virtual Ecosystem

A comparative study of learning gains in a MUVE and in the classroom was reported by Wrzesien and Raya (2010). A MUVE was used to represent an ecosystem of the Mediterranean Sea in the teaching of science and ecology. The environment was a custom developed system, and did not rely on Second Life, Active Worlds, or any other established platform. It was not a desktop environment, but rather, the learners interacted with it by entering a special room with projection facilities and tracking cameras. The educational experience involved children controlling avatars representing fish. The aim was to collect specific objects in the environment following instructions provided by a fish avatar controlled by the system.

The participants were 48 children between the age of 10 to 11 years old who were allocated to an experimental group (consists of 24 participants) and to a control group (consists of 24 participants) at random, thus, making it a true experiment. Participants in the control group listened to a lecture on Mediterranean sea ecology. For both of the groups, the instructional intervention involved a single session of 25 min duration. Knowledge levels were measured as scores in pen-and-paper tests of factual knowledge.

There was a statistically significant knowledge gain in both groups (at 0.05 significance level). However, the difference in knowledge gains between the experimental and the control groups was not significant ($\rho = 0.408$).

Thus, the overall result was the same as in the study by Cobb et al. (2009)—the difference between the learning gains in MUVE and in the classroom was not significant. Unlike in the Cobb et al. study, the learning design was not clearly scenario-based. The experience in MUVE did not match a real-life experience, and the activity (collection of artefacts) was, arguably, too narrow in scope to be seen

as a realization of constructivist learning. The measure of learning gain via pen-and-pencil tests had the same weaknesses as the similar measure in the study by Cobb et al. Cobb et al. also reported the count of requests for support when conducting a real-world task which appears to be an appropriate measure—the aim of the learning session was to prepare students for the real-world task and to enable them to conduct it as autonomously as possible, and the number of requests for support was directly related to the resource (support staff time) the learning session was intended to save. Wrzesien and Raya (2010) reported only the knowledge gain via a pen-and-pencil test.

In conclusion, both studies had strong and weak points. While Cobb et al. (2009) used an appropriate measure for detecting knowledge gains expected for scenario-based learning in MUVE, the validity of the findings was undermined by biased assignment of participants to groups, as noted by Cobb et al. (2009) in their article. On the other hand, Wrzesien and Raya (2010) used a random assignment, but used a measure of knowledge gain considered even by the authors themselves as not entirely appropriate. Wrzesien and Raya did not explain the reasons for choosing pen-and-pencil tests to measure knowledge gains. One can speculate, though, that an alternative was not feasible—authentic activities involving a real world under-water eco-system are difficult to implement.

2.12 Theories Relevant to Technology Acceptance

There has been a considerable amount of research examining the determinants of information technology acceptance and utilization among users. This is because technology acceptance is a major success factor—if technology is not used, any investment made into acquiring the technology is lost. Theories (or models) of technology acceptance are commonly presented in terms of depicting factors that lead to user acceptance or adoption of a particular information technology (Venkatesh & Davis, 2000; Chang & Tung, 2007). The terms "adoption" and "acceptance", when applied to individuals, are used in this thesis as synonyms, which is a common practice in the literature. Theories explaining adoption of information technology originated from a range of disciplines, such as psychology, sociology, and information systems. This section discusses the theories that are either particularly influential or particularly relevant to this study, especially as

they relate to the factors likely to influence the acceptance of MUVE for learning oral business English communication skills by Malay learners (research question two, see section 1.4). They include the innovations diffusion theory (IDT) (Moore & Benbasat, 1991; Rogers, 2003), the technology acceptance model (TAM) (Davis, Gagozzi, & Warshaw, 1989; Venkatesh & Davis, 2000), the social cognitive theory (SCT) (Compeau & Higgins, 1995a, 1995b), and the unified theory of acceptance and use of technology (UTAUT) (Venkatesh, Morris, Davis, & Davis, 2003).

2.12.1 Innovation Diffusion Theory

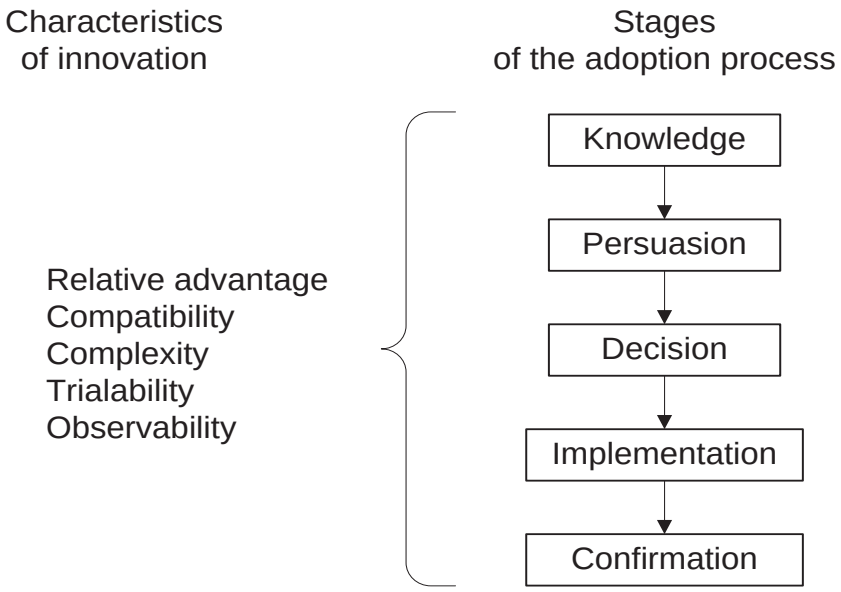

Figure 8. The innovation diffusion theory—stages and characteristics (based on Rogers, 2003).

In IDT, innovation refers to information technology (or, more generally, an idea, a practice, or any kind of object) that is perceived as new by an individual or by another unit of adoption, such as an organisation (Rogers, 2003). Diffusion is the process by which an innovation is communicated over time from one unit of adoption to another (Rogers, 2003).

IDT defines five stages of the innovation adoption process (see the illustration in Figure 8). The stages occur in chronological order: (1) knowledge of the innovation by the individual, (2) persuasion—the individual becomes interested

and actively seeks information, (3) decision—the individual weighs advantages versus disadvantages and makes a decision on whether to adopt the innovation, (4) implementation—the individual starts using the innovation, and (5) confirmation—the individual finalises the decision to use the innovation and starts using it to its full potential. The innovation adoption process is influenced by the five innovation characteristics, including relative advantage, compatibility with the current practice, complexity, trialability, and observability of the innovation by others (Rogers, 2003). For information technology innovations, (Rogers, 2003) modified and extended the IDT innovation characteristics by adding voluntariness, image, and result demonstrability, and by replacing complexity by ease of use.

Certain key constructs in IDT are similar to the constructs of the TAM (see section 2.12.2). In particular, relative advantage is similar to perceived usefulness, and complexity is similar to perceived ease of use. In technology acceptance research, TAM constructs are often mixed with constructs corresponding to perceived characteristics of innovation from IDT, and both of the theories are taken into account in UTAUT (see section 2.12.4).

2.12.2 Technology Acceptance Model (TAM)

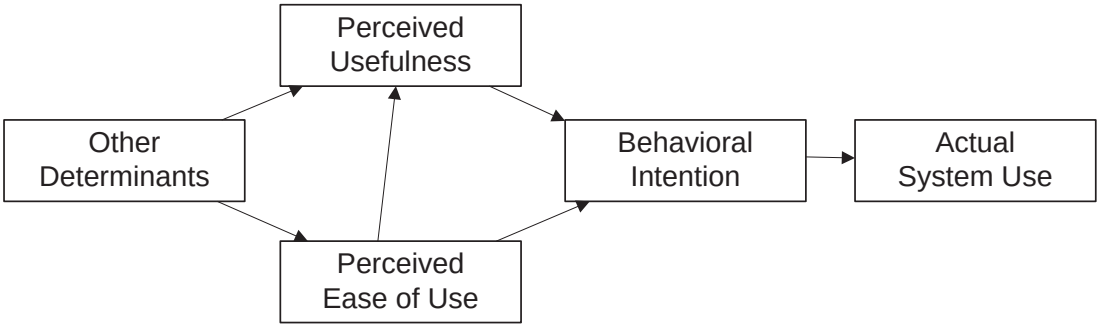

Figure 9. Technology acceptance model (taken from Davis,1989).

Another theory explaining user acceptance of information technology is TAM (Davis, 1989) (see the illustration in Figure 9). Originating from Ajzen and Fishbein's (1980) theory of reasoned action, TAM identifies two major factors determining technology acceptance as perceived usefulness and perceived ease of use. In TAM studies, behavioural intention to use is often used as a proxy for the actual behaviour (Ajzen, 1991; Agarwal & Karahanna, 2000). While early versions

of TAM included Attitude Towards Using a Technology as a mediating variable, most of the later studies, such as (Hong, Thong, Wong, & Tam, 2002) did not include it and followed the pattern of the model in Figure 9 by assuming that Perceived Usefulness and Perceived Ease of Use mediate the effect of any other variables.

Perceived usefulness is defined as the degree to which a person believes that using the system would enhance his or her performance. It relates to the users' beliefs that the system will be effective in helping the users to achieve their goals. On the other hand, Perceived Ease of Use refers to the degree to which the users believe that using a particular system would be free of effort in terms of physical and mental effort, as well as the ease of learning.

TAM has been validated in a wide range of contexts, including user acceptance of word processors (Davis et al., 1989), spreadsheet applications (Mathieson, 1991), e-mail (Szajna, 1996), telemedicine technology (Hu, Chau, Sheng, & Tam, 1999), commercial web sites (Koufaris, 2002), and web based learning systems (Ngai, Poon, & Chan, 2007).

Many studies found that Perceived Ease of Use acts as an antecedent of Perceived Usefulness and does not affect the Intention to Use directly (Mathieson & Eileen, 2001; Chau & Hu, 2002; Amoako-Gyampah & Salam, 2004; Hu, Lin, & Chen, 2005). However, there were also other studies that show just the opposite. The studies conducted by Moon and Kim (2001), Brown, Massey, Montoya-Weiss, and Burkman (2002), Gong and Xu (2004), and Shih (2004) found that Perceived Ease of Use directly affected the Intention to Use.

Studies that aim to explain and predict technology use, rather than just validate TAM, often consider other direct determinants of Intention to Use or of actual use, and do not make the assumption that the effect is mediated by Perceived Usefulness and Perceived Ease of Use. A particularly influential example of such approach is the TAM2 model (Venkatesh & Davis, 2000) that added the Subjective Norm (the perception that it is important if others think that the user should use the technology) to the basic TAM model to explain technology use in non-voluntary circumstances.

2.12.3 Social Cognitive Theory

SCT originated in psychology and asserts that human behaviour is determined by environmental factors and by the cognitive state of the individual (Bandura, 1991). With regards to cognitive factors, individuals are more likely to undertake behaviours they believe will result in desired outcomes and the behaviours they believe they are able to perform (the latter belief known as Self-efficacy). Based on the work of Bandura (1991), Compeau and Higgins (1995b) introduced the concept of Self-efficacy in technology acceptance research. The concepts of social cognitive theory included in the model by Compeau and Higgins (1995a) and Compeau et al. (1999) are Self-efficacy, Affect (a liking for particular behaviours), and Anxiety (feelings of apprehension with respect to performing the behaviours), along with the constructs representing performance outcome expectations. The illustration in Figure 10 shows SCT constructs from Compeau et al. (1999) and their relationships to System Use. As SCT strongly influenced the way the second research question was addressed in this study (see section 1.4), the rest of this section discusses it in greater detail.

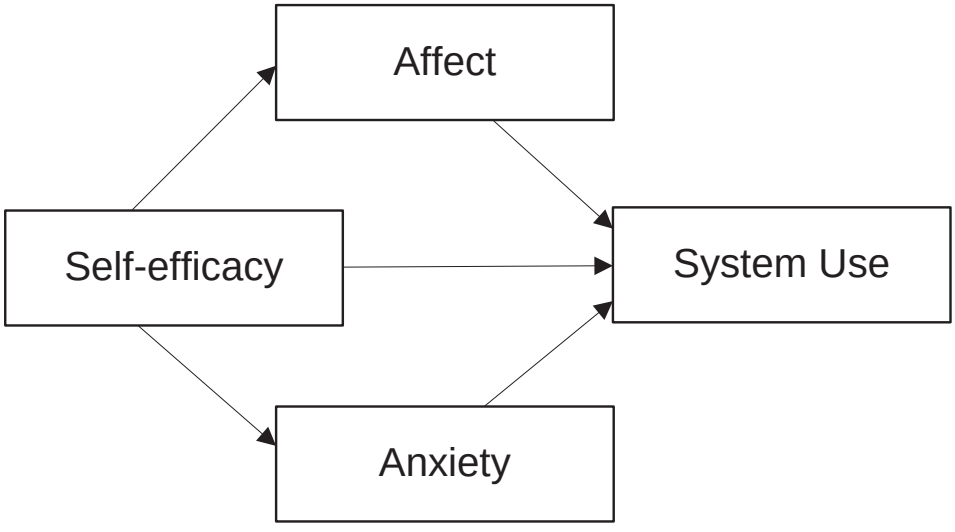

Figure 10. Social cognitive theory and technology usage (based on Compeau et al., 1999).

Self-efficacy is defined as one's capabilities to organise, perform, and execute the activities required to produce the desired results. It is concerned not with the actual

capabilities one possesses, but with the extent to which one believes in one's capabilities (Bandura, 1997). Self-efficacy influences activities; people undertake and perform confidently the activities that they judge themselves capable of managing, but avoid those they believe exceed their ability. These judgements also help people to decide the required effort and the time needed to overcome obstacles or difficulties at hand. Individuals with lower Self-efficacy tend to have a more negative view of their environment, focus on the negative aspects of their environment and on barriers they have to overcome, and thus often neglect the task at hand (Stumpf, Brief, & Hartman, 1987; Bandura, 1997). Self-efficacy impacts feelings of Affect and Anxiety with respect to performing the required behaviours (Stumpf et al., 1987; Bandura, 1997), which can in their own right impact behaviour.

Affect is the positive or negative feeling that is based on experience and that governs one's reaction to people, objects, and situations. In application to technology use, the affective component refers to the emotion or the feeling reflecting how much information technology is liked. Affect for computer use was found to influence the computer use (Compeau et al., 1999).

Anxiety refers to a state of unpleasant emotion resulting from the feelings of tension, worry, and apprehension. Individuals experience these feelings when they perceive conditions in their environment to be threatening (Bandura, 1986; Stumpf et al., 1987). Generally, there are three types of anxiety: state anxiety, trait anxiety, and situational anxiety. State anxiety is the feeling of apprehension experienced within a specific period of time. Trait anxiety is part of one's personality, and one who is suffering from this type of anxiety views the world as an unsafe and threatening place no matter what the actual level of risk is. Situational anxiety refers to individuals who are anxious only in specific environments; feeling anxious only when certain factors are present (Bandura, 1986; Stumpf et al., 1987).

In information system research, Self-efficacy is believed to be a determinant of technology use or acceptance (Bandura, 1986; Compeau & Higgins, 1995a; Compeau et al., 1999). Self-efficacy was found to have a direct effect on system usage (Compeau & Higgins, 1995b; Compeau et al., 1999) and on usage intention (Luarn & Lin, 2005; Wang, Lin, & Luarn, 2006). As to the effects of Anxiety on

system usage, Compeau and Higgins (1995b) reported that anxiety had a statistically significant, but small size effect on system usage. Compeau et al. (1999) found that Affect had a moderate yet statistically significant effect, while the effect of Anxiety was not statistically significant.

There are studies (Hill, Smith, & Mann, 1986, 1987; Gist, Schwoerer, & Rosen, 1989; Burkhart & Brass, 1990; Webster & Martocchio, 1993) that examined the impact of Self-efficacy on a variety of computer related behaviours. These studies found evidence of the effect of Self-efficacy on registration in computer courses at universities (Hill et al., 1987), on adoption of high technology products (Hill et al.,1986) and innovations (Burkhart & Brass, 1990), as well as on performance in software training (Gist et al. 1989; Webster & Martocchio, 1993). Moreover, a number of studies reported that repeated computer use has a positive effect on Self-efficacy (Clark, 2003; Emurian, 2004; Cheong, Pajares, & Oberman, 2004).

2.12.4 Unified Theory of Acceptance and Use of Technology

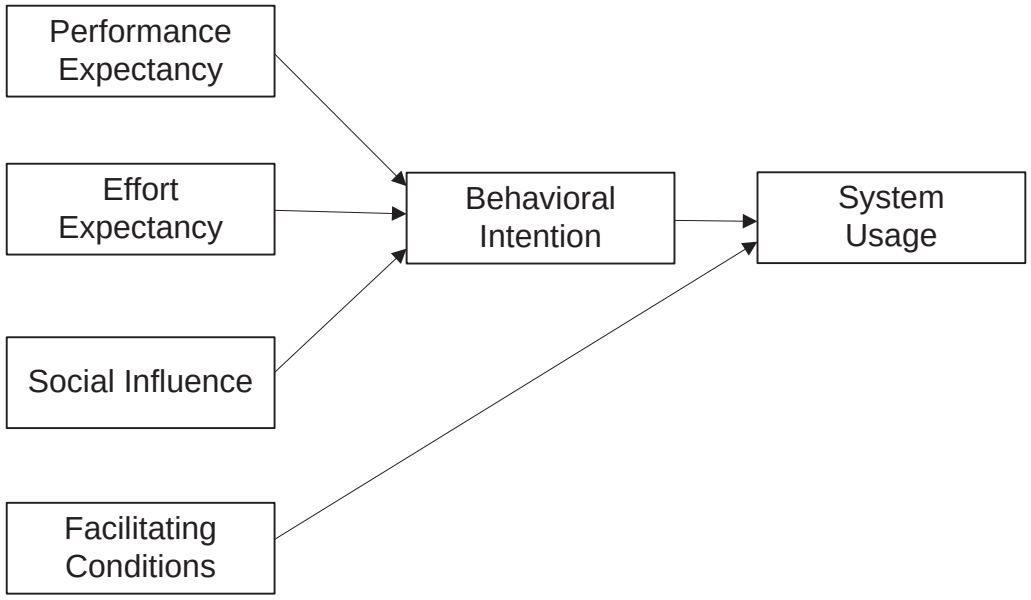

Figure 11. Unified theory of acceptance and use of technology (Venkatesh et al., 2003).

Venkatesh et al. (2003) made an attempt to integrate the existing technology acceptance models, including IDT, TAM, SCT, and a number of others into a single model known as the UTAUT. The main constructs and relationships of UTAUT are illustrated in Figure 11.

UTAUT includes four core constructs determining the technology acceptance by users. The four constructs are listed below along with comments on the constructs from other theories that the UTAUT constructs incorporate.

1. Performance expectancy is the degree to which individuals believe that using the system will help them to improve their job performance. Performance acceptance corresponds to Relative Advantage in IDT, and to Perceived Usefulness in TAM.
2. Effort expectancy is the degree to which individuals believe that using the system is free from effort. Effort expectancy corresponds to Complexity in IDT and to Perceived Ease of Use in TAM.
3. Social Influence is the degree to which individuals believe that important others (such as superiors or co-workers) expect them to use the system. Social influence corresponds to Subjective Norm in TAM2.
4. Facilitating Conditions is the degree to which individuals believe that an organization and technical infrastructure exists to support their use of the system. Facilitating Conditions correspond to Compatibility in innovation diffusion theory.

Besides the core constructs shown in the Figure 11, the full UTAUT includes moderators: Gender, Age, Experience, and Voluntariness of Use (Venkatesh et al., 2003). These variables moderate the impact of the four main determinants on System Usage.

The influence of the UTAUT model, however, was somewhat limited, because it appears that even after its publication the studies on technology acceptance often used TAM, rather than UTAUT, as the basis for extended models incorporating new variables. For example, a search of the content of the Computers and Education (a major e-Learning journal) using scholar.google.com revealed that (up to April 2013) UTAUT has been mentioned in only 16 articles; in contrast, TAM has been mentioned in 128 articles.

2.13 MUVE Acceptance Studies

The existing MUVE acceptance studies are summarized in Table 3 by listing the relationships they hypothesised, and are discussed one by one and compared in the rest of this section. As seen in Table 3, the existing studies considered a wide range of determinants and were not consistent in their approaches, as determinants added to extend the basic TAM model were related to different constructs in different studies.

Table 3: Relationships Tested in Prior Studies of MUVE Acceptance

Independent variable	Dependent variables	**β**	***p***	Sources
Anxiety [a]	Perceived Usefulness	0.05	> 0.05	(Fetscherin & Lattemann, 2008)
Attitude Towards Technology	Perceived Usefulness	0.29	< 0.05	(Fetscherin & Lattemann, 2008)
Community	Perceived Usefulness	0.58	< 0.05	(Fetscherin & Lattemann, 2008)
Social Norms	Perceived Usefulness	0.06	< 0.05	(Fetscherin & Lattemann, 2008)
Perceived Media Richness	Perceived Usefulness	Not given	< 0.001	(Saeed et al., 2008)
Perceived Media Richness	Perceived Ease of Use	Not given	< 0.001	(Saeed et al., 2008)
Perceived Enjoyment	Intention to Use	Not given	<.0.001	(Saeed et al., 2008)
Perceived Role Projection	Intention to Use	Not given	< 0.5	(Saeed et al., 2009)
Personal Emotional Involvement	Intention to Use	Not given	< 0.01	(Saeed et al., 2009)
Subjective Norms	Intention to Use	Not given	< 0.5	(Saeed et al., 2009)
Subjective Norms	Perceived Usefulness	Not given	< 0.01	(Saeed et al., 2009)
Computer Anxiety [a]	Perceived Ease of Use	Not given	> 0.05	(Shen & Eder, 2009)
Computer Playfulness	Perceived Ease of Use	0.29	< 0.01	(Shen & Eder, 2009)
Computer Self-Efficacy	Perceived Ease of Use	0.35	< 0.01	(Shen & Eder, 2009)

Note. In all sources, TAM variables were included (Perceived Usefulness, Perceived Ease of Use, and Intention to Use). Relationships between TAM variables are not listed in this table.

[a] In Fetscherin and Lattermann (2008), Anxiety corresponded to anxiety with respect to using Second Life specifically, while in Shen and Eder (2009), Computer Anxiety referred to anxiety with respect to using computers in general.

The studies by Fetscherin and Lattermann (2008), Saeed et al. (2008), and Saeed, Yang and Sinnappan (2009) focused on the acceptance of MUVE by current Second Life users. Therefore, it is likely that the participants in these studies used Second life for purposes other than learning.

In contrast, the study by Shen and Eder (2009) is particularly relevant to the problem addressed in this thesis (in particular, research question two, see section 1.4), as the study was conducted with students and specifically asked about the acceptance of MUVE e-Learning. The participants in the study by Shen and Eder used Second Life for learning; moreover, the study explicitly focused on the acceptance of Second Life as an environment for e-Learning.

2.13.1 The Study by Fetscherin and Lattemann (2008)—the Effects of Community, Attitude towards Technology, Social Norms and Anxiety

Fetscherin and Lattemann (2008) conducted a study of factors affecting user acceptance of Second Life. They assessed the effects of Attitude towards Technology, Social Norms, Performance Expectancy, and Anxiety within the framework provided by the Technology Acceptance Model. Fetscherin and Lattemann (2008) took a view that the effect of these factors on Intention to Use is intermediated by Perceived Usefulness (the details of the model are given in

Figure 12). The items used to measure the constructs in the model were in part adopted from the literature, and in part formulated anew, to reflect the specifics of Second Life and MUVE. The study (the study by Fetscherin and Lattemann) used a large convenience sample; the Second Life residents were contacted via a variety of social media outlets, including an in-world announcement within the Second Life itself, and 249 Second Life users participated in the study. The data (in the study by Fetscherin and Lattemann) was collected in 2007. It is not known how

many potential participants saw the invitation to participate, so the response rate cannot be determined.

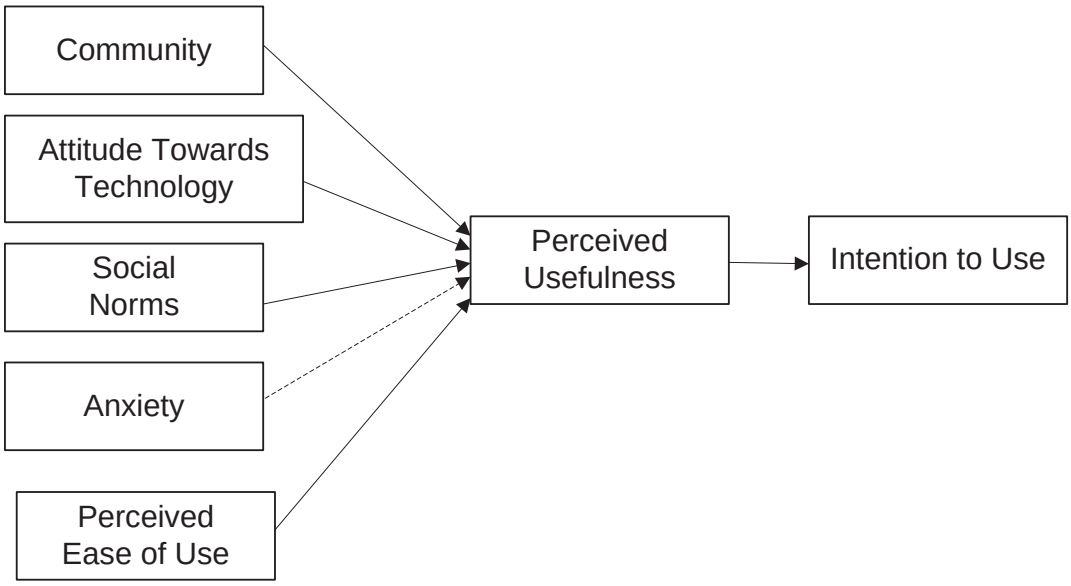

Figure 12. Fetscherin and Lattemann (2008, p. 239) research model. Solid arrows show paths found to be significant at $p < 0.05$.

As the study (the study by Fetscherin and Lattemann, 2008) used a convenience sample, the population for the study is somewhat vaguely defined—it is difficult to judge how representative it is of the Second Life residents, as the respondents self-selected to participate from a very large population. The study does not employ any analysis to demonstrate that the respondents represent the target population (such as comparing early responses with late responses, or comparing characteristics of the participants with the characteristics of the population). Such analysis was probably impossible, because it was probably difficult to distinguish early respondents from late respondents because of the way the data was collected, and because information on the characteristics of Second Life residents is difficult to obtain.

Based on the data collected, an exploratory factor analysis was conducted to assess the basic structure of the proposed research model. Based on this analysis, Performance Expectancy and Attitude Toward Technology were combined into one construct (named Attitude Toward Technology), and a new construct, Community, was added to the research model (using a subset of items originally

allocated to Perceived Usefulness). The research model was fitted using covariance based SEM. Community, Attitude Towards Technology, Social Norms, and Perceived Ease of Use affected Perceived Usefulness. Moreover, Perceived Usefulness and Perceived Ease of Use significantly affected Intention to Use. The effect of anxiety on Perceived Usefulness was not statistically significant.

There is a possibility that the results of the study were due to the capitalisation of chance, because the model in the study was re-formulated to fit the data, and was not fully based on a priori theoretical considerations.

2.13.2 The Study by Saeed et al. (2008)—the Effects of Perceived Media Richness on Perceived Usefulness and Perceived Ease of Use

Saeed et al. (2008) conducted a study of the effect of Perceived Media Richness on user acceptance of Second Life. They assessed the effect of Perceived Media Richness within the framework provided by the Technology Acceptance Model (and thus, the study is motivated by the Media Richness Theory, Daft & Lengel, 1986). (Saeed et al. 2008) took a view that the effect of Perceived Media Richness on Intention to Use is intermediated by Perceived Usefulness and Perceived Ease of Use (the details of the model are given in Figure 13). The items used to measure the constructs in the model were based on previously published scales. Saeed et al. (2008) invited survey participants through various educational forums and groups in Second Life. The data were collected in March 2008. The number of respondents was 112. The majority of respondents had at least six months experience in Second Life. It is not known how many potential participants saw the invitation to participate, so the response rate cannot be determined.

Based on an analysis using PLS approach, it was found that Perceived Media Richness affected Perceived Usefulness and Perceived Ease of Use. Moreover, Perceived Usefulness affected Intention to Use. But the effect of Perceived Ease of Use on Intention to Use was not statistically significant (Saeed et al., 2008).

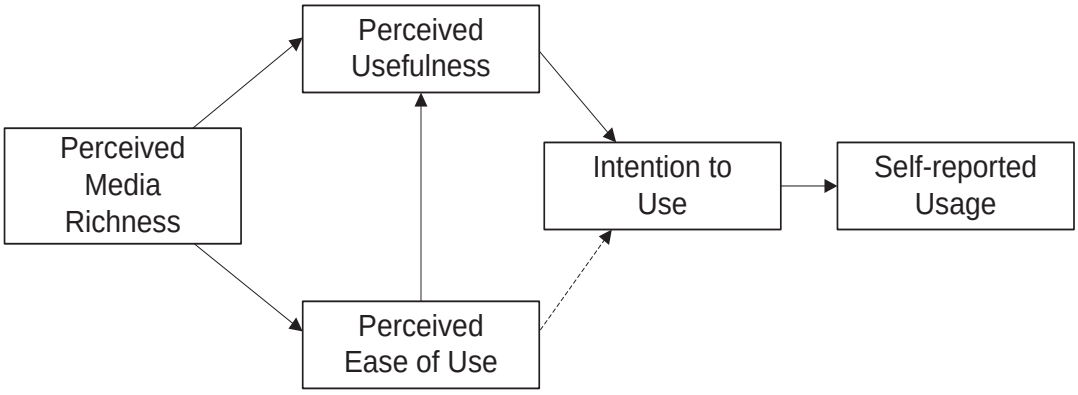

Figure 13. Saeed, Yang, and Sinnappan (2008, p. 853) research model. Solid arrows show paths found to be significant at $p < 0.01$.

The study predicted only 15% of variance in Intention to Use ($R^2 = 0.15$), which is relatively low for TAM studies. For example, a later study by the same authors, using a different set of independent variables, predicted 51% of variance in Intention to Use. The study differed from the rest of the TAM acceptance studies discussed in this section by also measuring the self-reported Actual Use of the Second Life. Although it is common in TAM studies to assume that Intention to Use is a good proxy of the actual use, in this particular case the Intention to Use predicted only 18.5% of variation of the Actual Use.

The study (Saeed et al., 2008) used a convenience sample, and the approach to recruiting participants was very similar to Fetscherin and Lattemann (2008), with the same criticism applying in this respect (see section 2.13.1). The study does not report standardised values of β path coefficients, which makes it somewhat difficult to compare the results to similar studies.

2.13.3 The Study by Saeed et al. (2009)—the Effects of Subjective Norms, Perceived Emotional Involvement, Perceived Enjoyment and Perceived Role Projection on Intention to Use

Saeed et al. (2009) conducted another study, following Saeed et al.(2008), of factors affecting the user acceptance of Second Life in 2009. They assessed the effect of Perceived Emotional Involvement, Perceived Enjoyment, Perceived Role Projection, and Subjective Norms within the framework provided by the Technology Acceptance Model. The choice of the first of these three variables was motivated by the hedonistic consumption theory (Hirschman & Holbrook, 1982), a theory from marketing suggesting that consumers make choices not solely based of the utility of a product, but also based on multisensory images, fantasies, and emotional arousal associated with using the product. The authors took a view that these variables directly affect Intention to Use (the details of the model are given in

Figure 14). The items used to measure the constructs of the model were based on previously published scales. The study (Saeed et al., 2009) invited survey participants through in-world notification (within Second Life), Second Life educators mailing list, and Second Life research listserv.

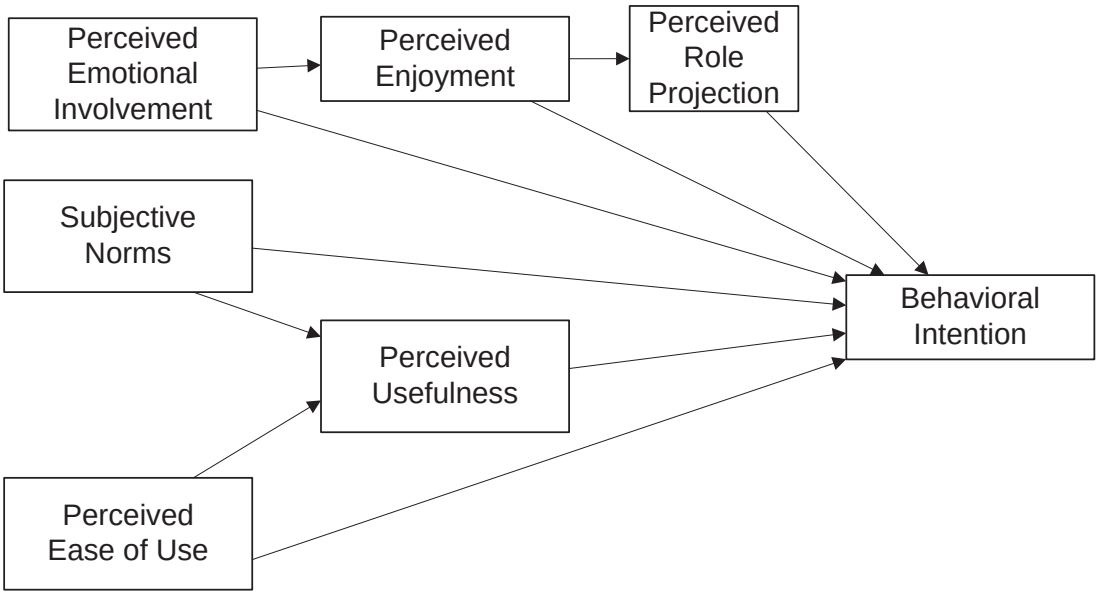

Figure 14. Saeed et al. (2009, p. 6) research model. All paths were significant at $p < 0.01$.

The data were collected in 2008. The number of respondents was 122, with the majority having at least six months experience in Second Life. However, it is not known how many potential participants saw the invitation to participate, so the response rate cannot be determined. Even though some of the characteristics of the data set are similar to (Saeed et al. 2008), the details such as the exact number of participants and the number of males and females in these studies differ, suggesting that the studies used different data sets.

The study (Saeed et al., 2009) used a convenience sample, and the approach to recruiting participants was very similar to (Fetscherin & Lattemann, 2008) and (Saeed et al., 2008), so that in this respect the same criticism applies (see section 2.13.1).

Based on an analysis using PLS approach, Perceived Role Projection, Perceived Enjoyment, Perceived Emotional Involvement, and Subjective Norms significantly affected Intention to Use. Moreover, Subjective Norms affected Perceived Usefulness.

Overall the model explained 51% of variance in Intention to use. Comparing with the results of Saeed et al. (2008), there was a considerable difference in variance explained in Intention to Use. It appears that Perceived Enjoyment was a particularly strong (and direct) determinant of Intention to Use, stronger than Perceived Usefulness. This is not surprising, as it is quite likely that the participants primarily used Second Life for entertainment. (Entertainment is the main aim of Second Life. The front page at http://secondlife.com states: "Experience endless surprises and unexpected delights in a world imagined and created by people like you.") However, as the study does not report standardized values of β path coefficients, it makes it difficult to interpret the results or to compare the results to similar studies.

2.13.4 The Study by Shen and Eder (2009) —the Effects of Computer Playfulness, Computer Self-Efficacy and Computer Anxiety on Perceived Ease of Use

Shen and Eder (2009) conducted a study of factors affecting business students' acceptance of Second Life as an e-Learning environment. They assessed the effect of Computer Playfulness, Computer Self-efficacy, and Computer Anxiety within the framework provided by TAM. Shen and Eder (2009) took a view that the effect of these factors on Intention to Use is intermediated by Perceived Ease of Use (the details of the model are given in Figure 15). It should be noted that Computer Anxiety and Computer Self-efficacy are the dimensions of the framework originally used by Compeau et al. (1999) to represent the social cognitive theory concepts in the context of technology acceptance research (Compeau et al., 1999). The items used to measure the constructs of the model were adopted from the literature. The study (Shen & Eder, 2009) was conducted in the autumn of 2007, and the data were collected from the students enrolled in e-Commerce and Management Information Systems classes of a business school at a university. A total number of 77 responded to the survey, with a response rate of 85.6%.

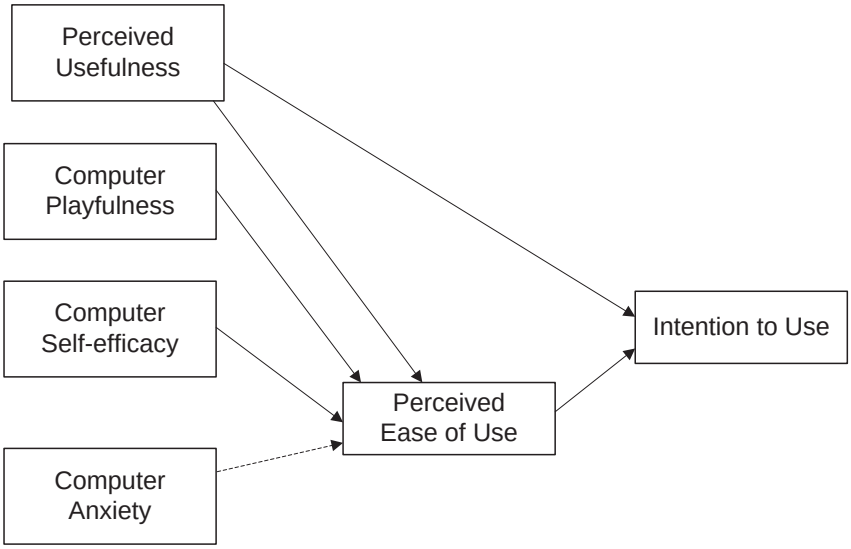

Figure 15. Shen and Eder (2009, p. 227) research model. Solid arrows show paths found to be significant at $p < 0.01$.

Based on an analysis using the PLS approach, Computer Playfulness and Computer Self-efficacy were found to affect Perceived Ease of Use. However, the effect of Computer Anxiety was not found to be significant.

Of the Second Life acceptance studies, the study by Shen and Eder (2009) is the closest one to the issues raised by the second research question of this thesis (see section 1.4). It does not answer the research question, though, because findings for business students in the US cannot be easily generalized to L2 Malay learners. In addition, Shen and Eder's study does not address the issues identified in the introduction and the literature review as particularly relevant to Malay learners and to the users of Second Life in general, namely, the learners' English Class Anxiety (see section 1.1.3) and the effects of video game experience (see section 2.7.1).

2.14 Knowledge Gaps to be Addressed in This Study

Even though there are numerous claims in the literature that MUVE is an effective e-Learning environment, there are very few empirical studies of MUVE effectiveness involving controlled experiments or quasi experiments (and thus, unlike experience reports, providing objective results). Only two studies of this kind were identified in the literature review of the present study. The first of these studies, by Cobb et al. (2009), was a quasi-experiment with a design problem acknowledged by the authors. Students who arrived earlier were assigned to the MUVE environment, while student who arrived later were assigned to a control group. Thus, it is likely that more motivated students were assigned to the MUVE environment. The second study, by Wrzesien and Raya (2010), used a controlled experiment with random assignment but the nature of the MUVE environment involved was quite different from the desktop-based MUVE that is potentially of benefit to Malay learners and, thus, is the focus of the present thesis. Neither of these two studies was in the context of learning business communication by using a scenario-based approach. This knowledge gap (lack of MUVE effectiveness studies in the relevant context) has to be addressed to answer the first research question of this study.

Even though there are a number of studies of factors affecting MUVE acceptance (reviewed in section 2.13), only one of these studies (Shen & Eder, 2009) addressed MUVE acceptance as an e-Learning environment, in context of learning Management Information Systems, rather than in context of second language learning. While all of these studies tested models based on TAM, only two of them involved variables relating to SCT. The first of these studies, by Fetscherin and Lattemann (2008), included anxiety with respect to using MUVE, while the study by Shen and Eder (2009) included Computer Anxiety and Computer Self-efficacy. Self-efficacy and anxiety with respect to learning the subject matter and with respect to video games are not covered in the existing studies. In view of the features of Malay learners, it is believed that these variables are particularly relevant for this study. In addition, none of the existing studies of MUVE acceptance was in context of teaching business English communication skills using a scenario-based approach, and none of the existing studies used Malay learners as participants. This knowledge gap (lack of MUVE acceptance studies involving subject matter and video games Self-efficacy and Anxiety and lack of MUVE acceptance studies in the relevant context) has to be addressed to answer the second research question of this study.

2.15 Summary

This chapter started by discussing business English and ESP, followed by a discussion of different methods used to teach English and the technology used to support them. Then, it discussed the features of MUVE, and the use of MUVE as an e-Learning platform, with a focus on using MUVE for teaching language skills and on teaching strategies particularly appropriate for MUVE, and existing quantitative studies of MUVE effectiveness. The chapter proceeded to a discussion of information systems acceptance and the related theories, including IDT, SCT and UTAUT. This was followed by a detailed discussion of the existing studies of MUVE acceptance. The chapter concluded by analysing gaps in the current literature that need to be addressed to answer the research questions of this study.

Chapter 3: Research Hypotheses

3.1 Introduction

This chapter introduces the research hypotheses employed in this study.

As stated in the introduction (section 1.4), the study addresses the following research questions, covering the two complementary facets of the overall research problem addressed by the thesis: (1) effectiveness and (2) acceptance of MUVE as an e-Learning environment in the Malaysian context:

1) Is MUVE based learning effective in facilitating situated scenario-based learning of oral business English communication skills by Malay learners?
2) Which factors influence the acceptance of MUVE based learning of oral business English communication skills by Malay learners?

Correspondingly, two sets of research hypotheses are posed. Research question one is addressed by the hypotheses E1 and E2, labelled with the letter E for "effectiveness". These hypotheses are introduced in section 3.2.

Research question two is addressed by a SEM research model introduced in section 3.3. The hypothesis in the research model (A1 to A14, labelled with the letter A for "acceptance") are introduced in the sub-sections of the section 3.3.

3.2 Hypotheses Relating to MUVE Effectiveness: E1 and E2

As discussed in the introduction and in the literature review (sections 1.1.1 and 2.7), MUVE offers a combination of physical and social environment that is highly realistic—MUVE users experience a high degree of presence as they have the perception of actually being in the environment. Situated learning involves learning in a realistic environment representing the target on-the-job situation (see section 2.9.2). Scenario based learning involves implementing job-relevant learning scenarios in a realistic environment (see section 2.9.3). In terms of the language teaching paradigms and instructional designs introduced in section 2.3,

scenario based language learning, because of the emphasis on executing realistic scenario tasks and because of the emphasis on realistic environment, can be seen as a realization of task based and communicative language leaning. MUVE can provide an environment for situated scenario-based learning, offering features that are difficult or impossible to offer in a traditional classroom (see sections 2.7 and 2.3).

Therefore, for MUVE based situated scenario based learning to make most effective use of the physical and social environment provided by MUVE, the following hypotheses are formulated:

E1: Students' communication proficiency will improve after a session of situated scenario-based learning in MUVE.

E2: Students' communication proficiency after a session of situated scenario-based learning in MUVE will improve to a greater extent than after a session of classroom learning.

3.3 Hypotheses Relating to MUVE Acceptance: A1 to A14

The research model for the purpose of the study to measure user acceptance MUVE was developed based on combining technology acceptance model and social cognitive theory. In the model (Figure 16), the Intention to use MUVE e-Learning is the main dependent variable in this study. It stands as a proxy variable for Actual Use of e-Learning, which could not be measured. Perceived Ease of Use and Perceived Usefulness, along with the Intention to Use are Technology Acceptance Model (TAM) constructs. While validating TAM in context of MUVE was not a goal of this study, TAM constructs were added to ensure that important known predictors of Intention to Use are included. Video Games Self-efficacy, Video Games Affect, and Video Games Anxiety correspond to computing-related Self-efficacy, Affect, and Anxiety in Compeau and Higgins (1995b), and represent the feeling of learners with respect to virtual visual environments. Similarly, Subject Matter Self-efficacy, Attitude Towards Using English, and English Class Anxiety represent the feelings towards learning English. Finally, Desire to Learn English accounts for learners' motivation. The constructs are presented in Figure 16, along with the research hypotheses. The model comprises 10 constructs and 14

hypotheses (labelled from A1 to A14). Detailed introductions of the constructs and the hypotheses included in the model are given in the rest of this section.

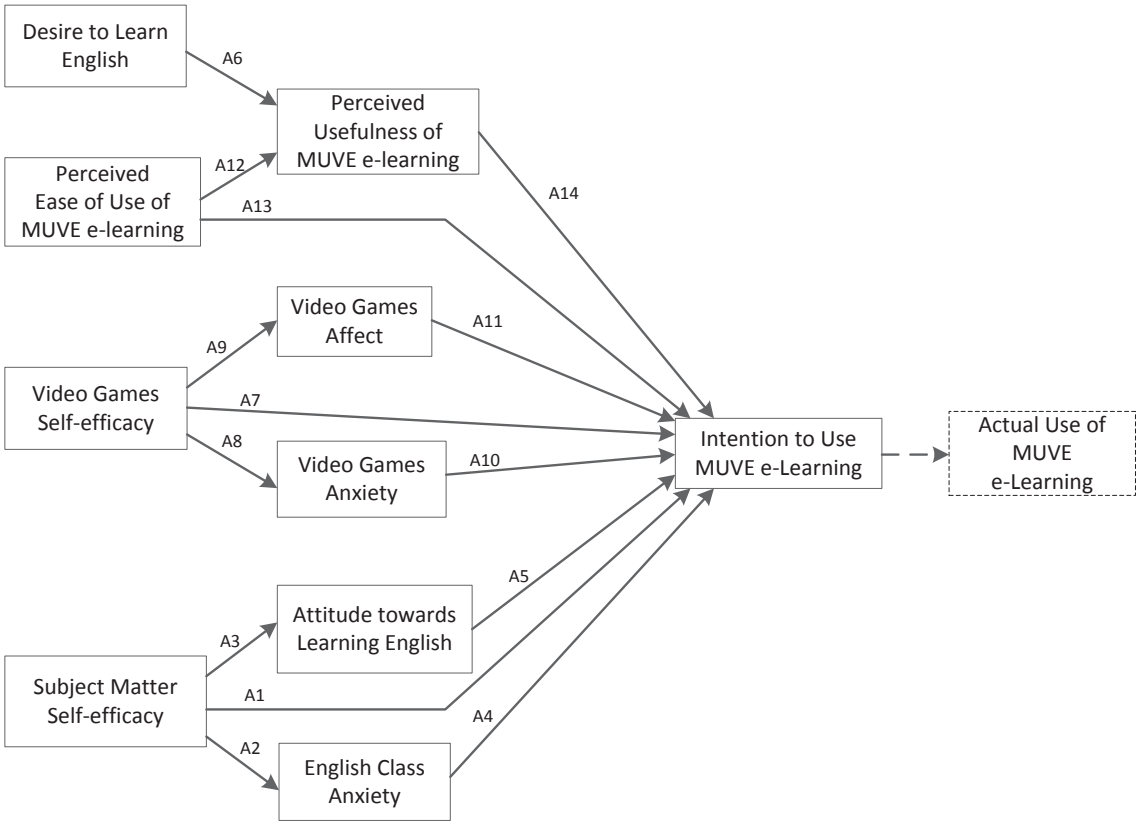

Figure 16. Research model presenting in a single diagram the hypotheses relating to MUVE acceptance introduced in section 3.3 of the present study.

3.3.1 A1, A2, and A3: The Impact of Subject Matter Self-Efficacy

Self-efficacy is a term introduced by (Bandura, 1997), the founder of social cognitive theory, to refer to one's belief in one's ability to perform a particular behaviour. In this section, self-efficacy is seen in the context of usage of the English language and defined as the belief in one's ability to use the English language. (Bandura, 1997) suggested that individuals with low self-efficacy beliefs will be less likely to perform related behaviours in the future. Low self-efficacy also hinders individual participation in learning activities (Bandura, 1986; Schunk, 1991). Consequently, individuals' perceptions of their own competence in the second language (L2) will either hinder or facilitate users' intentions to accept or refuse a specific activity.

Therefore, lack of confidence in learning the English language (conceptualized as Subject Matter Self-efficacy) was considered to have a potential effect on the Intention to Use MUVE e-Learning. The present study also follows Compeau and Higgins (1995b) in assuming that Subject Matter Self-efficacy will affect Attitude Towards Learning English and English Class Anxiety. Attitude Towards Learning English corresponds to Affect in the work by Compeau and Higgins (1995b), while English Class Anxiety corresponds to Anxiety in the same work. (See sections 3.3.3 and 3.3.2 for detailed introductions of these concepts).

A1: Subject Matter Self-efficacy affects Intention to Use MUVE based e-Learning.

A2: Subject Matter Self-efficacy affects English Class Anxiety.

A3: Subject Matter Self-efficacy affects Attitude Towards Learning English.

3.3.2 A4: The Impact of English Class Anxiety

MacIntyre and Gardner (1994a) described L2 anxiety experienced by second language speakers—the feeling of tension and apprehension when they were compelled to communicate orally in a L2. Many students feel more tense and nervous in Foreign Language (FL) class than in any other classes (Horwitz, Horwitz, & Cope, 1986; MacIntyre & Gardner, 1989; Campbell & Ortiz, 1991), and their anxiety seems to come predominantly from speaking situations (Horwitz et al., 1986; Young, 1990; Koch & Terrell, 1991; Price, 1991).

Anxiety is common among FL students (Aida, 1994), and is associated negatively with language performance (Gardner & MacIntyre, 1993; Brown, 2001) and the ability to self-manage in an online learning environment (Mirjam & Stella, 2005). As such, anxiety towards language learning may hinder student acceptance of a learning activity.

One of the themes that emerged in a recent qualitative study of educator perceptions around the use of the Second Life MUVE for teaching (Gamage et al., 2011) was an educator expectation that students who are shy in a traditional classroom, and thus reluctant to participate, are going to feel more comfortable in a MUVE, where intermediation by an avatar leads to a degree of anonymity. This view was also voiced in the literature (Baker, Wentz, & Woods, 2009; Cheong,

2010). Therefore, in relation to this study, individuals' anxiety to use English language in class was thought to positively affect the Intention to Use MUVE e-Learning.

A4: English Class Anxiety affects Intention to Use MUVE e-Learning.

3.3.3 A5: The Impact of Attitude Towards Learning English

Attitude Towards Learning English refers to the feelings experienced by the learner while learning the language. It refers to the feeling of like or dislike towards the English language (Gardner, 1985). It is hypothesised that as a positive Attitude Towards Learning English translates into the Intention to Use a new environment intended to facilitate such learning.

A5: Attitude Towards Learning English affects Intention to Use MUVE e-Learning.

3.3.4 A6: The Impact of Desire to Learn

Desire to Learn the target language (motivation to learn the target language) is concerned with whether an individual strives to work hard in learning the language because of his or her interest to do so (Gardner, 1985). This study hypothesises that individuals who Desire to Learn English will view MUVE useful, as MUVE is a tool allowing them to achieve their goal. This is based on a view that users who are motivated to pursue the goals facilitated by an information system are more likely to view the information system as useful; the close connection between perceived usefulness of an information system and the motivation to achieve goals facilitated by the information system was highlighted by Venkatesh et al. (2003).

A6: Desire to Learn English affects Perceived Usefulness.

3.3.5 A7, A8, and A9: The Impact of Video Games Self- efficacy

Self-efficacy is expected to have a direct influence on task choice and the persistence in achieving the task (Bandura, 1997). Based on Bandura's definition of self-efficacy, Video Games Self-efficacy is defined in this study as a belief of one's capability to play video games. Following Bandura's argument, individuals

have low self-efficacy beliefs if they have little confidence in their skills or experience in video games and are dissatisfied or uncomfortable using video game skills. In this sense, those with low self-efficacy beliefs will be less likely to perform related behaviours such as playing computer games in the future. In contrast, individuals with high self-efficacy will tend to experience less video games related anxiety, use video games more often, and enjoy using them more (Compeau & Higgins, 1995b). It is assumed that student confidence or feelings in respect to playing video games will apply to MUVE, which offers a similar environment. This study also follows Compeau and Higgins (1995b) in assuming that Video Games Self-efficacy will affect Video Games Affect and Video Games Anxiety (see sections 3.3.7 and 3.3.6 for detailed introductions of these concepts)—learners confident about playing video games will have more positive feelings with respect to playing video games and will be less anxious about playing video games.

A7: Video Games Self-Efficacy affects Intention to Use MUVE e-Learning.

A8: Video Games Self-Efficacy affects Video Games Anxiety.

A9: Video Games Self-Efficacy affects Video Games Affect.

3.3.6 A10: The Impact of Video Games Anxiety

In this study, Video Games Anxiety refers to feelings of unease and worry with respect to playing video games. Computer Anxiety was shown to have a negative effect on information system usage (Compeau & Higgins, 1995b; Compeau et al., 1999). Moreover, Computer Anxiety was also found to result in avoidance of computer usage among users (Maurer, 1994). As MUVE is similar to computer games, individuals with a high degree of Video Game Anxiety are expected to have a lower degree of Intention to Use MUVE e-Learning.

A10: Video Games Anxiety affects Intention to Use MUVE e-Learning.

3.3.7 A11: The Impact of Video Games Affect

Affect is the feeling of like or dislike towards video games. Torkzadeh, Pflughoeft and Hall (1999) commented that individuals generally proceed with activities that

they like and avoid activities that they do not like. It should, however, be noted that affect (like or dislike) is a different construct than anxiety (Kernan & Howard, 1990). An individual could simultaneously like video games and have lots of anxiety towards them. One study, for example, found no significant relationship between negative affect and anxiety (Thatcher & Perrewe, 2002). In this study, in view of similarity between video games and MUVE, a person's affect towards video games was hypothesized as affecting the acceptance of 3D MUVE e-Learning.

A11: Video Games Affect impacts Intention to Use MUVE e-Learning.

3.3.8 A12 and A13: The Impact of Perceived Ease of Use

Perceived Ease of Use is the degree to which a person believes that using a particular system will be free of effort. In other words, it refers to the individual's evaluation of the effort involved in the process of using the system (Davis, 1989). Perceived Ease of Use is expected to influence Perceived Usefulness and the Intention to Use MUVE. A lot of research in the past provided evidence that Perceived Ease of Use significantly affects the behavioural Intention to Use an information system, either directly or indirectly, through its effect on Perceived Usefulness (Agarwal & Prasad, 1999; Venkatesh, 1999; Venkatesh et al., 2003). There are also multiple studies exploring the impact of Perceived Ease of Use on the use or the behavioural Intention to Use e-Learning (Ong, Lai, & Wang, 2004; Chang & Tung, 2007; Ngai et al. 2007; Lee, Yoon, & Lee, 2009). Therefore, in this study Perceived Ease of Use is hypothesised to affect Intention to Use MUVE e-Learning and Perceived Usefulness.

A12: Perceived Ease of Use affects Perceived Usefulness.

A13: Perceived Ease of Use affects Intention to Use MUVE e-Learning.

3.3.9 A14: The Impact of Perceived Usefulness

Perceived Usefulness is defined as the degree to which a person believes that using a particular system will increase his or her job performance. Davis et al. (1989) stated that people tend to either use or not use an application until they believe it will enhance their job performance by decreasing the time taken, and achieving

more efficiency as well as accuracy in doing the job (see also Venkatesh et al. (2003) for an extensive review of TAM research). There are also multiple studies exploring the impact of Perceived Usefulness on the use or the behavioural Intention to Use e-Learning (Yi & Hwang, 2003; Gong & Xu, 2004; Lee et al., 2009; Friedrich & Hron, 2010).

A14: Perceived Usefulness affects Intention to Use MUVE e-Learning.

3.3.10 The Target Dependent Construct - Intention to Use

It has been shown across a variety of studies that Intention to Use is a strong predictor of actual use (Chau, 2001; Fusilier & Durlabhji, 2005). In this study, it was not possible to measure actual use. The participants were not users of MUVE, but experienced MUVE e-Learning for the first time as the study participants. They were not involved in any course involving MUVE e-Learning either before or after the study. Therefore, Intention to Use was adopted as a proxy variable.

3.4 Summary

This chapter presented the research hypotheses addressing the research questions of this study (as formulated in section 1.4). Hypotheses E1 and E2 addressed research question one, related to MUVE effectiveness. Hypotheses A1 to A14 addressed research question two, related to MUVE acceptance. The following chapter presents the research design and methodology.

Chapter 4: Research Design and Methodology

4.1 Introduction

This chapter presents a detailed description of the research design of this study. The research setting and the research participants (Malay learners taking a course in Business Communication) are discussed first, followed by a discussion of data collection. The data in this study was obtained by using an experiment with random assignment for MUVE effectiveness and a cross-sectional survey for MUVE acceptance. Both the experimental data and the survey data were collected from the same participants, in the same session. The research instruments for measuring communication proficiency and user acceptance are then identified, and approaches taken to assess their validity are discussed. For the MUVE acceptance survey, existing instruments were adapted for the purposes of the study, while for measuring communication proficiency, a new rating scale, based on an existing scale, was created. Then, the methods of data analysis are identified (standard *t*-test for MUVE effectiveness and structural equation modelling for factors affecting MUVE acceptance, with exploratory factor analysis and co-variance based SEM employed to confirm the PLS results). Potential threats to the validity of the findings are also discussed. Then, the human ethics issues are considered. The chapter concludes by introducing a pilot study used to confirm that the research design and the procedures are adequate and workable. The steps followed in the present research are outlined in Figure 17.

The present thesis relies on using research methods well established in MIS, behavioural, and e-Learning research and does not intend to contribute to the field of research methodology. Rather, established research methods are employed as tools. Therefore, the discussion in this chapter is focused and relies on established methodological sources. The full richness of the methodological literature is not presented because it would distract from addressing the purpose of the research. Moreover, following the recommendations of the Publication Manual of the American Psychological Association (American Psychological Association, 2010, p. 116), no literature references are given for statistics in common use.

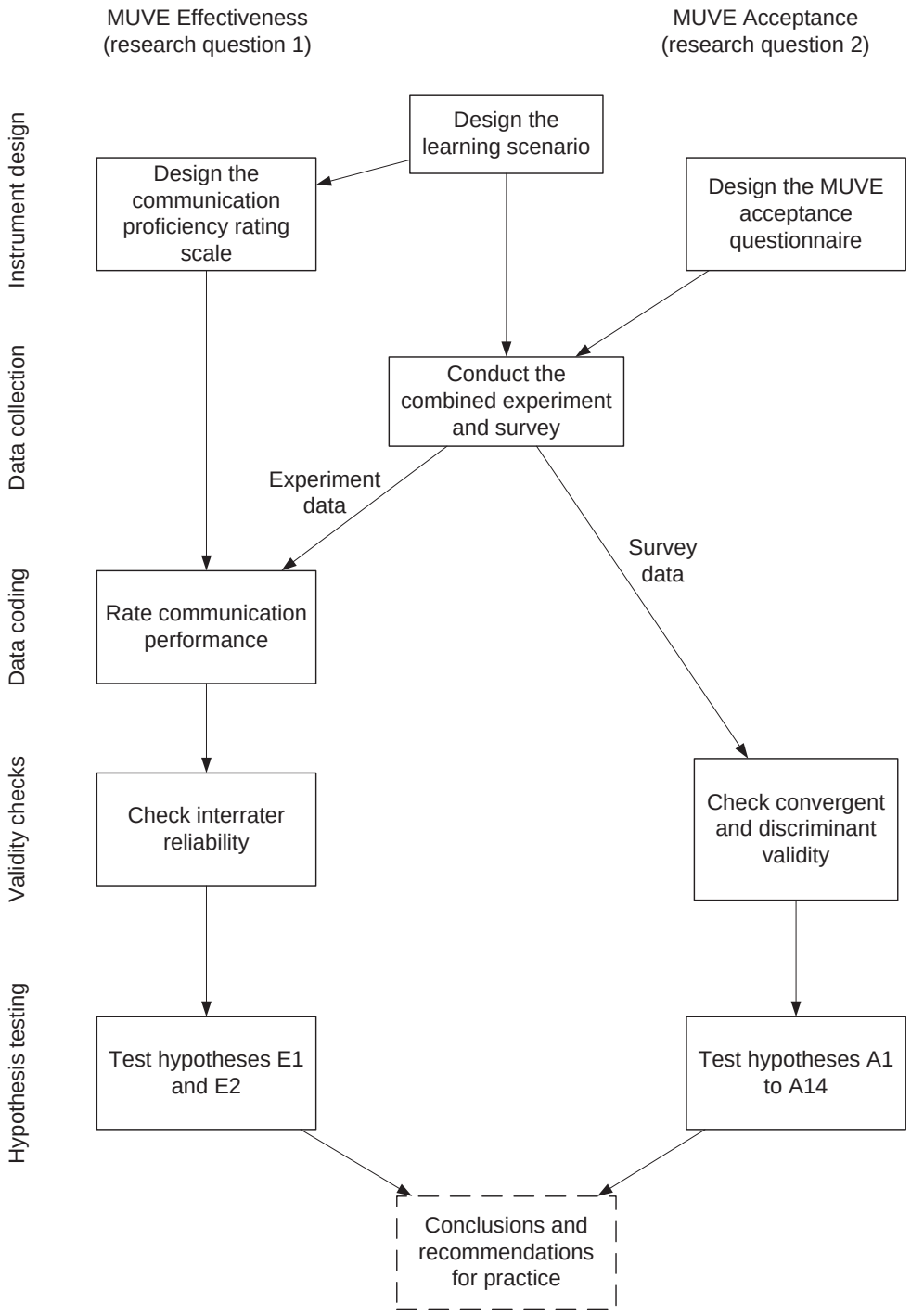

Figure 17: The steps followed in the present study.

4.2 Choosing the Research Paradigm

This section justifies the choice of the positivist research paradigm as the underlying paradigm for the present study. First, it introduces the two dominant

research paradigms: positivist and intepretivist before making an argument for the choice of research paradigm in view of both the problem addressed in this research and the research questions. The discussion in the following two paragraphs is based on Bryman and Bell (2003).

The positivist research paradigm suggests that knowledge is independent from the people who possess it. Positivist researchers formulate theories and test them by fitting them to data. If a theory is not rejected in a test, this is considered to be evidence in favour of the theory. It is never possible to prove that a theory is correct. It is only possible to present further evidence that it is not wrong. Theories are viewed as objective and generalizable—valid for a particular range of circumstances independently from the subjective opinions of the researcher or research participants. Most commonly, the positivist paradigm is implemented by using a quantitative research strategy. Quantitative research relies on measuring quantities. In the context of positivist research, theories are formulated as relationships between quantities and are tested by using statistical analysis. The specific research methods most relevant to positivist research are experiments and quasi-experiments, cross-sectional surveys, longitudinal surveys, and quantitative content analysis involving instruments such as rating scales and purporting to generate objective ratings.

Even though this study did not involve conducting interpretivist research, it is essential to introduce this type of research and the qualitative research strategy to justify the choice of the positivist research paradigm as well as to provide background information for the discussion of further research in Chapter 6. The interpretivist research paradigm suggests that knowledge is constructed by people and cannot be separated from the people who constructed it. Interpretivist researchers capture subjective meanings of research participants. Theories are viewed as subjective constructions created by researchers and research participants. Most commonly, the interpretivist paradigm is implemented by using a qualitative research strategy. Qualitative research involves collecting and analysing rich data from multiple sources, prolonged involvement with the research participants, and in-depth understanding of a particular situation and its context. The specific research methods most relevant to interpretivist research are

participant observation, semi structured and unstructured interviews, and analysis of content relying on a researcher's subjective understanding.

The present study aimed to assess the effectiveness or MUVE as a learning environment and to assess the effects of factors potentially affecting the acceptance of MUVE as a learning environment. The aim was to conduct an objective assessment, rather than to explore subjective meanings. Therefore, the study relied on the positivist research paradigm and on quantitative research.

4.3 Choosing the Overall Research Design

4.3.1 Experiments

Experimental research seeks to determine if a specific treatment influences an outcome. In this type of research, at least one of the variables is deliberately manipulated or varied by the researcher to determine the effects of that variation. In addition, in order to have what has become to be accepted as a "true experiment", the participants in the experiment must be assigned randomly to the experimental treatments (Wiersma & Jurs, 2009; Weathington, Cunningham, & Pittenger, 2010). The advantage of true experimental research is that it allows the distinguishing of cause from effect.

To address the first research question (see section 1.4), a controlled experiment was designed, with random assignment of participants. This allowed to establish a clear cause and effect connection between the use of MUVE for e-Learning and the measured learning gains.

4.3.2 Quasi-experiments

This research does not involve conducting quasi-experiments. This section discusses quasi-experiments as background information necessary to justify the choice of research design, so as to understand the limitations of this research, and to make suggestions for further research. The discussion is based on Weathington, Cunningham, & Pittenger (2010).

Quasi-experimental research is similar to experimental research in that one or more experimental variable is involved. However, instead of having participants

randomly assigned to experimental treatments, "naturally" assembled groups, such as classes, are used in the research. Members can also self-select themselves into the groups. Because of the difficulty often encountered when attempting to form groups by random assignment, quasi-experimental research is quite common (Wiersma & Jurs, 2009; Weathington et al., 2010)

Quasi-experimental research is similar to experimental research in allowing one to make inferences regarding cause and effect. However, compared with true experimental research, the conclusions made from quasi-experimental research are less certain. In quasi-experimental research it is uncertain if the effect is due to manipulating the experimental variables, or due to some other circumstances. For example, learners of class A may perform better than learners of class B because they, unlike learners in class B, used e-Learning, or, as an alternative explanation, they might just be more interested in the subject matter than the learners of class B.

A quasi-experiment may be easier to implement in a real organisation comparing to a real experiments, because a quasi-experiment requires less control by the researcher. It may be possible, for example, to make an arrangement to employ somewhat different approaches to learning in two different classes. On the other hand, it is unlikely for a researcher to have control of how learners are assigned to classes in a real organisation. The outcome of a quasi-experiment conducted in a real organisation can be viewed as more relevant (better able to be generalised) to real practice, because the treatments were applied in more realistic circumstances. Thus, both real experiments and quasi-experiments have advantages and disadvantages. This study did not use a quasi-experiment because the researcher did not have access to a group of learners engaged in MUVE based e-Learning in the context of a real organisation.

4.3.3 Surveys

This section discusses surveys as a research methodology. Similar to section 4.3.2, the discussion is based on Weathington et al. (2010).

This research included a cross-sectional survey. This section, along with cross-sectional surveys, discusses longitudinal surveys as background information

necessary to justify the choice of research design, to understand the limitations of this research, and to make suggestions for further research.

Surveys are studies involving the collection of data via questionnaire or interview, without subjecting the participants to any treatments arranged by the researcher. Survey data analysis deals with the distribution, incidence, and relationships of the relevant variables.

Survey designs can be categorised into two types: longitudinal and cross-sectional. Longitudinal designs involve the collection of data over time and at specified points in time. In contrast to longitudinal designs, cross-sectional designs involve the collection of data at one point in time from a random sample representing a given population at that time.

In cross-sectional survey research, it is often difficult to distinguish between A causing B and B causing A. For example, if one measures Perceived Ease of Use and Perceived Usefulness at the same time, one cannot tell if the participants decided not to use the system because they think it is not useful because an alternative explanation (users first decided not to use the system for some other reason, and then formed the belief that the system is not useful to justify the decision). In longitudinal surveys, if A was measured before B, it constitutes evidence in favour of A causing B.

To address the second research question (see section 1.4), a cross-sectional survey was designed and executed, because experimental or longitudinal research of MUVE acceptance (for example, measuring participant attitudes and later observing if they are use MUVE) was not feasible. In particular, the researcher did not have access to a group of learners engaged in MUVE based e-Learning over a long period of time.

4.3.4 The Overall Research Design of This Study

As indicated earlier in this section, the combination of a true experiment and a survey was used to test the hypotheses introduced in Chapter 3. The experiment was used to test the hypotheses relating to MUVE effectiveness (E1 and E2), and the survey was used to test the hypotheses relating to the determinants of MUVE acceptance (A1 to A14). The experiment and the survey were conducted

simultaneously, with the same set of participants (the participants are introduced in section 4.4). The detailed experimental and survey procedures are described in section 4.5.

4.4 Research Participants

The research was carried out at a public institution of higher learning in Malaysia in September and October, 2008. For the purpose of maintaining confidentiality to fulfil human ethics requirements, the setting and the participants of the research are kept anonymous and only the relevant details are provided. The selected institution is a university offering diploma and degree courses, enrolling mostly local students at the diploma level. However, foreign students are also eligible to enrol at diploma, degree, master, and doctorate levels.

The participants were involved in the English for Professional Communication course, which was the final English course and was compulsory for all diploma and bachelor degree students at the institution. Prior to taking the course, participants passed the Malaysian University English Test (MUET) and completed two English courses, Basic English Proficiency and English for Academic Communication at the same university.

Basic English Proficiency emphasises the basic skills of listening, speaking, reading, and writing, whereas English for Academic Communication focuses on more specific areas of grammar, reading texts, and academic writing skills. The English for Professional Communication course focuses on oral business English communication skills and on report writing.

The course in which the participants were enrolled, was a typical English language course—the course did not have any special characteristics that would distinguish it from other, similar courses. (This statement relies on the judgment of the researcher who has been involved in teaching English at a university in Malaysia for many years. Population data that would allow statistical comparisons to objectively verify this claim was not available.) Thus, the results of the study are likely to be of relevance to all Malay university students in Malaysia. It is, however, not possible to conduct statistical tests to demonstrate that the participants did not differ from the full population of Malay university students in

Malaysia. Sampling at random from the full population of Malay university students in Malaysia would fulfil the statistical requirement for generalizability to the target population. This, however, was not considered, because this was clearly not feasible in the present study.

From the 180 students taking the English for Professional Communication course, 152 participated in the study. All of the students enrolled in the course were Malay; therefore, all of the participants were Malay. Thus, the sample was self-selected (as is the case in almost all experimental and survey research). On the other hand, because the overwhelming majority of students in the course participated in the research, the participants well represented the Malay students involved. Thus, the response rate in the study was 84%. Among the respondents, 55% were male (n=84) and 45% were female (n=68). The majority of the respondents were between 18 and 21 years old, with the average age being 20 years old. Respondents were also asked about their use of Second Life before the study. The great majority had never heard about Second Life before, and none of the participants had experience in using Second Life or any other MUVE.

4.5 Experimental and Survey Procedures

Students were assigned at random in pairs to Group A (control group) and Group B (experimental group) as they arrived. The random assignment procedure involved assigning odd-numbered pairs, such as the first or the third pair, to Group A or Group B based on an outcome of a random numbers generator. Then, the following even-numbered pair was assigned to the alternative group.

In both groups, students experienced both MUVE e-Learning and classroom learning. Students in Group A had a session in the classroom first, and then a session in Second Life. Group B students had a session in Second Life first, and then a session in the classroom. Teaching followed the scenario-based teaching paradigm - students executed the target scenario (Appendix C) in different roles (switching roles from time to time), with the teacher providing comments and corrections and modelling the behaviour as appropriate.

For all participants, communication performance was assessed before and after the first session, thus implementing the classic experimental design with a control group and with random assignment. The MUVE e-Learning survey acceptance

questionnaire was administered after the second session. Thus, by the point when the survey questionnaire was distributed, all participants had experience of MUVE e-Learning.

Table 4: Experimental and Survey Procedures

Step	Procedure	Duration	Groups A and B	
1.	Introduction	5 minutes	The researcher explained the nature of the study and the procedures involved, while the participants read information sheets and signed release forms (see Appendix A and Appendix B).	
2.	Technology acceptance survey, part 1	3 minutes	The participants filled in the technology acceptance questionnaire, Part 1 (see Appendix E for the questionnaire and see Figure 16 for the research model the questionnaire was based on). This part included items about video games (this part of the questionnaire does not rely on the participants having an experience of MUVE e-Learning). The participants were asked to complete this part of the survey at this stage to avoid confusion between Second Life and video games.	
3.	Technology demonstration	10 minutes	The researcher demonstrated the technology to be used (Second Life) and let the participants try it out.	
4.	Modelling	10 minutes	The participants watched a demonstration of the target scenario, which involved a seller showing a house to a potential buyer (see Appendix C for a description of the scenario). A doll house was used in this scenario.	
5.	Pre-test (audio recorded in digital format)	10 minutes	Participants were asked to execute the scenario twice in the classroom, switching roles - each participant acted once as a seller and once as a buyer (in random order).	
6.	Teaching session 1	40 minutes	Teaching in the classroom for Group A	Teaching in the virtual environment for Group B
7.	Post-test (audio recorded in digital format)	10 minutes	The same procedure was applied as in the pre-test, in the same order.	
8.	Teaching session 2	40 minutes	Teaching in the virtual environment for Group A	Teaching in the classroom for Group B
9.	Technology acceptance survey, part 2	7 minutes	Participants filled in the technology acceptance questionnaire (see Appendix F for the questionnaire).	
10.	Wrap-up	10 minutes	The researcher (who was also an experienced English language teacher) gave encouraging comments / advice on how to further improve English language proficiency of the participants.	

The details of the data collection procedures are given inTable 4. Useable digital recordings of pre and post test performances were obtained for 68 pairs of students—34 pairs in each group. Thus, the pre-and post-test scores were available only for 136 of the 152 students who volunteered to participate.

As pairs of students were allocated to groups at random, the male / female ratio in the two groups was not the same (see Table 5).

Table 5: Gender Composition of the Two Groups

Gender	Group A	Group B
F	22	32
M	46	36

Even though females were better represented in Group B, the result of Pearson's Chi-squared test of homogeneity suggested that the difference in male / female ratio between the two groups was not statistically significant ($p = 0.11$).

A longer instructional intervention in the experiment, covering a broad range of scenarios, would result in stronger external validity because it would closer resemble a course of study at a real educational institution and would assess the effectiveness of MUVE for a broader area of content. In the present study, however, a longer instructional intervention was not possible because it was not possible to ensure the continued involvement by the participants.

A short-term intervention may have biased the results in favour of MUVE (with participants more engaged in learning due to novelty effect, something that would have weared down in a longer study) or against MUVE (as adjusting to the new environment takes cognitive effort, the participants might have performed better in a longer study because they'd have time to get accustomed to the new environment).

4.6 Intervention–the Learning Scenario

The scenario chosen for the study was "House Selling", involving the learners taking turns at assuming the role of house seller or house buyer. Learners used the English language in a simulated real life situation. The scenario-based approach is especially beneficial for the evaluation of oral performance related to specific work

requirements (Weir, 1988). Assessment relying on learners' performance in such scenarios is a better indicator of real life performance than interview-based assessment because it allows a variety of language activities similar to real life situations.

In this study, an existing house in Second Life was used for the "House Selling" activity. The house in Second Life was highly realistic, with minute details carefully represented. The participants, represented by their avatars, could walk inside and outside the house, viewing and presenting it from different vantage points. In the classroom, a low fidelity doll house was used for the same purpose.

For a learner participating in the scenario, the house, its interiors, and the area outside constituted a physical environment closely representing the real world environment of the task represented in the scenario. At the same time, the other learner and the instructor represented the social environment, in which the task was tried out, and the approaches to better completing the task were negotiated. Thus, the intervention had the characteristics of constructivist, situated, scenario-based learning (see section 2.3).

4.7 Measurements Instruments

4.7.1 Oral Business English Communication Proficiency Assessment Scale

An oral assessment scale was used to measure the learning gains by students in the experiment—to determine if the exposure to a particular learning method (learning in the classroom or learning in the MUVE environment) generated a learning gain and to compare the learning gains between the two environments (hypotheses E1 and E2).

An oral assessment scale was designed to suit the purpose of the study based on an existing scale by Oliveira (2004). The following components were used to measure gain scores: organising ideas, vocabulary, accuracy, and fluency. Each component had a number of items, with each item rated from 1 to 4. The details of the scale are given in Appendix D.

Oliveira's rating scale was developed based on Bachman and Palmer's Language Testing Model, which is an established model supported both theoretically and empirically (Bachman, 1990; Bachman and Palmer, 1996; Chalhoub-Deville, 1997). The rating scale used in this study is an analytic one—it is focussed on the specific language features encountered in the test. Analytic scales tend to be more reliable than holistic scales that are based on statements relating to overall language proficiency, because analytic scales allow raters to focus on narrowly defined, specific criteria in assigning the score (Brown & Bailey, 1984; Hamp-Lyons, 1991).

The rating scale follows Oliveira's scale in using 4 levels of descriptors for each item, which is fewer than in some other tests, including the well-known International English Language Testing System (IELTS) (Oliveira, 2004). As noted by Underhill (1987), fewer levels make assessments easier to do while also increasing the reliability of the rating scale. Underhill further explained that the rating scale works well when the assessor can hold the scale levels in his or her mind while listening to the learners. The rating scale developed in Oliveira's work specifically caters for a task-based oral assessment in English for Specific Purposes (ESP), which further justifies using it as a basis for the scale employed in this study.

4.7.2 MUVE e-Learning Acceptance Questionnaire

A questionnaire was developed to gather data to test the relationships of the MUVE e-Learning user acceptance research model (see section 3.3).

In developing the questionnaire for the survey, a combination of existing validated scales was used to develop the survey questionnaire for this study. Reusing existing instruments is common in survey research, because existing instruments have already been assessed for validity and reliability and because it facilitates comparisons between studies. In information system research, the use of existing instruments is extensive (Leidner & Jarvenpaa, 1995; Kitchenham & Pfleeger, 2002).

TAM constructs covered by the questionnaire were Perceived Usefulness, Perceived Ease of Use, and Behavioural Intention to Use. Items used to measure these constructs were adapted from Davis (1989) and Hong et al. (2002).

Perceived Ease of Use was measured by six items, with items reworded to mention MUVE (see Table 6).

Table 6: Perceived Ease of Use Scale (PEOU)

Item	Original Wording	Modified Wording
PEOU1	My interaction with the system would be clear and understandable.	My interaction with the interactive visual computer environment is clear and understandable.
PEOU2	I would find the system to be flexible to interact with	Interacting using the interactive visual computer environment does not require a lot of mental effort.
PEOU3	I would find the system easy to use.	I find the visual computer environment easy to use.
PEOU4	I would find it easy to get the system to do what I want it to do.	I found it easy to get the interactive visual computer environment to do what I want it to do.
PEOU5	Learning to operate the system would be easy for me.	Learning to operate the visual computer environment is easy for me.
PEOU6	It would be easy for me to become skillful at using the system.	It is easy for me to become skilful at using the interactive visual computer environment.

Perceived Usefulness was measured by six items, with items reworded to mention MUVE (see Table 7).

Table 7: Perceived Usefulness Scale (PU)

Item	Original Wording	Modified Wording
PU1	Using the system would enhance my effectiveness on the job.	Using the interactive visual computer environment improves my learning efficiency.
PU2	Using the system would improve my job performance.	Using the interactive visual computer environment will increase my learning performance.

Item	Original Wording	Modified Wording
PU3	Using the system in my job would increase my productivity.	Using the interactive visual computer environment increases my learning output.
PU4	I would find the system useful in my job.	Using the interactive visual computer environment is useful for my learning.
PU5	Using the system in my job would enable me to accomplish tasks more quickly.	Using the interactive visual computer environment will allow me to accomplish learning tasks more quickly.
PU6	Using the system would make it easier to do my job.	Using the interactive visual computer will make it easier to learn course content.

Intention to Use was measured by five items, with items reworded to mention MUVE (see Table 8).

Table 8: Intention to Use Scale (INT)

Item	Original Wording	Modified Wording
INT1	I intend to use the system as often as possible.	I intend to use the visual computer environment as often as needed.
INT2	Whenever possible, I intend to use the system in my study.	Whenever possible, I intend to use the visual computer environment for learning.
INT3	I plan to use the system in the future.	I plan to use the visual computer environment in the future.
INT4	I intend to continue using the system in the future	I intend to continue using the visual computer environment in the future.
INT5	I expect my use of the system to continue in the future.	I expect my use of the visual computer environment to continue in the future.

The items measuring social cognitive theory constructs in application to video games were adapted from Compeau and Higgins (1995b).

Video Games Self-efficacy was measured by 10 items, adapted to mention video games (see Table 9).

Table 9: Video Games Self-efficacy Scale (VGSE)

Item	Original Wording	Modified Wording
VGSE1	I could complete the job using the software if there was no one around to tell me what to do as I go.	I could play a video game if there was no one around to tell me what to do as I play.
VGSE2	I could complete the job using the software if I had never used a package like it before.	I could play a video game if I had never played this type of game before.
VGSE3	I could complete the job using the software if I had only the software manuals for reference.	I could play a video game if I had only the manuals for reference.
VGSE4	I could complete the job using the software if I had seen someone else using it before trying it myself.	I could play a video game if I had someone else playing it before trying it myself.
VGSE5	I could complete the job using the software if I could call someone for help if I got stuck.	I could play a video game if I could call someone for help if I got stuck.
VGSE6	I could complete the job using the software if someone else had helped me get started.	I could play a video game if someone else had helped me get started.
VGSE7	I could complete the job using the software if I had a lot of time to complete the job or which the software was provided.	I could play a video game if I had a lot of time to complete it.
VGSE8	I could complete the job using the software if I had the built-in help facility for assistance.	I could play a video game if I had the built-in help facility for assistance.
VGSE9	I could complete the job using the software if someone showed me how to do it first.	I could play a video game if someone showed me how to do it first.
VGSE10	I could complete the job using the software if I had used similar packages before this one to do the same job.	I could play a video game if I had played a similar video game.

Video Games Affect was measured by five items, adapted to mention video games (see Table 10).

Table 10: Video Games Affect Scale (VGAFF)

Item	Original Wording	Modified Wording
VGAff1	I like working with computers.	I like playing video games.
VGAff2	I look forward to those aspects of my job that require me to use a computer.	I look forward to playing video games.
VGAff3	Using a computer is frustrating for me.	Playing video games is fun for me.
VGAff4	Once I start working on the computer, I find it hard to stop.	Once I start playing video games, I find it hard to stop.
VGAff5	I get bored quickly when working on a computer.	I get excited quickly when playing video games.

Video Games Anxiety was measured by four items, adapted to mention video games (see Table 11).

Table 11: Video Games Anxiety Scale (VGAN)

Item	Original Wording	Modified Wording
VGAN1	I feel apprehensive about using computers.	I feel apprehensive about playing video games.
VGAN2	It scares me to think that I could cause the computer to destroy a large amount of information by hitting the wrong key.	Playing video games scares me.
VGAN3	I hesitate to use a computer for fear of making mistakes I cannot correct.	I hesitate to play video games in case I look stupid
VGAN4	Computers are somewhat intimidating to me.	Playing video games makes me feel uncomfortable.

As to the language aspects of the model, the socio-educational model of L2 acquisition (Gardner, 1985) was used to measure individual differences in the learning of the L2. The socio-educational model is an established one and focuses on the role of various individual characteristics of the student in the learning of a L2 in the classroom. In the socio-educational model of L2 acquisition, the Attitude/Motivation Test Battery (AMTB) is commonly used to measure individual differences in the learning of the L2 (Gardner, 1985).

The scales for Attitude Towards Learning English, Subject Matter Self-efficacy, English Class Anxiety, and Desire to Learn English were adapted from Attitude/Motivation Test Battery (AMTB) (Gardner, Tremblay, & Masgoret, 1997).

Attitude Towards Learning English was measured by nine items, adapted to mention English (see Table 12). It should be noted that the first four items were positively keyed, and the remaining five, negatively keyed.

Table 12: Attitude Towards Learning English Scale (ATLE)

Item	Original Wording	Modified Wording
ATLE1 Positively keyed	French is really great.	*English* is really great.
ATLE2 Positively keyed	I really enjoy learning *French.*	I really enjoy learning *English.*
ATLE3 Positively keyed	I love learning *French*.	I love learning *English*.
ATLE4 Positively keyed	I plan to learn as much *French* as possible.	I plan to learn as much *English* as possible.
ATLE5 Negatively keyed	I hate *French*.	I hate *English*.
ATLE 6 Negatively keyed	I would rather spend my time on courses other than *French.*	I would rather spend my time on courses other than *English*.
ATLE7 Negatively keyed	I find the study of *French* very boring.	I find the study of *English* very boring.
ATLE 8 Negatively keyed	Learning *French* is a waste of time.	Learning *English* is a waste of time.

Item	Original Wording	Modified Wording
ATLE 9 Negatively keyed	When I finish this course, I shall give up the study of *French* entirely because I am not interested in it.	When I finish this course, I shall give up the study of *English* entirely because I am not interested in it.

Subject Matter Self-efficacy was measured by 10 items, adapted to mention English (see Table 13).

Table 13: Subject Matter Self-efficacy Scale (SMSE)

Item	Original Wording	Modified Wording
SMSE1	I'm sure I could speak French well in almost any circumstances.	I'm sure I could speak English well in almost any circumstances.
SMSE2	When the French language is spoken to me, I feel I can understand practically everything.	When the English language is spoken to me, I feel I can understand practically everything.
SMSE3	I feel comfortable conducting myself in French almost any time and any place.	I feel comfortable conducting myself in English almost any time and any place.
SMSE4	I believe that I can competently read and understand most books and articles written in French.	I believe that I can competently read and understand most books and articles written in English.
SMSE5	I may not be completely fluent in French, but I feel confident speaking it.	I may not be completely fluent in English, but I feel confident speaking it.
SMSE6	Despite the fact that I may not be completely proficient in French, I am self-assured conducting myself in French.	Despite the fact that I may not be completely proficient in English, I am self-assured conducting myself in English.
SMSE7	Even when I make mistakes speaking French, I still feel sure of myself while trying to communicate.	Even when I make mistakes speaking English, I still feel sure of myself while trying to communicate.
SMSE8	I am confident when having conversation with French-speaking people despite any errors I may make.	I am confident when having conversation with English-speaking people despite any errors I may make.

Item	Original Wording	Modified Wording
SMSE9	Regardless of how much French I know, I feel confident about using it	Regardless of how much English I know, I feel confident about using it
SMSE10	I feel confident using French regardless of my ability.	I feel confident using English regardless of my ability.

English Class Anxiety was measured by 10 items, adapted to mention English (see Table 14). The first five items were positively keyed, and the rest, negatively keyed.

Table 14: English Class Anxiety Scale (ECAN)

Item	Original Wording	Modified Wording
ECAN1 Positively keyed	I never feel quite sure of myself when I am speaking in our French class.	I never feel quite sure of myself when I am speaking in our English class.
ECAN2 Positively keyed	It embarrasses me to volunteer answers in our French class.	It embarrasses me to volunteer answers in our English class.
ECAN3 Positively keyed	It worries me that other students in my class seem to speak French better than I do.	It worries me that other students in my class seem to speak English better than I do.
ECAN4 Positively keyed	I get nervous and confused when I am speaking in my French class.	I get nervous and confused when I am speaking in my English class.
ECAN5 Positively keyed	I am sometimes afraid the other students will laugh at me when I speak French.	I am sometimes afraid the other students will laugh at me when I speak English.
ECAN 6 Negatively keyed	I don't usually get anxious when I have to respond to a question in my French class.	I don't usually get anxious when I have to respond to a question in my English class.
ECAN 7 Negatively keyed	I feel confident when asked to participate in my French class.	I feel confident when asked to participate in my English class.
ECAN 8 Negatively keyed	I do not get anxious when I am asked for information in my French class.	I do not get anxious when I am asked for information in my English class.

Item	Original Wording	Modified Wording
ECAN 9 Negatively keyed	I don't understand why other students feel nervous about using French in class.	I don't understand why other students feel nervous about using English in class.
ECAN10 Negatively keyed	Students who claim they get nervous in French class are just making excuses.	Students who claim they get nervous in English class are just making excuses.

Desire to Learn English was measured by nine items, adapted to mention English (see Table 15). The first five items were positively keyed, and the rest—negatively keyed.

Table 15: Desire to Learn Scale (DLE)

Item	Original Wording	Modified Wording
DLE1 Positively keyed	I wish I had begun studying French at an early age	I wish I had begun studying English at an early age.
DLE2 Positively keyed	If it were up to me, I would spend all of my time learning French.	If it were up to me, I would spend all my time learning English.
DLE3 Positively keyed	I want to learn French so well that it will become second nature to me.	I want to learn English so well that it will become second nature to me.
DLE4 Positively keyed	I would like to learn as much French as possible.	I would like to learn as much English as possible.
DLE5 Positively keyed	I wish I were fluent in French.	I wish I were fluent in English.
DLE6 Negatively keyed	Knowing French isn't really an important goal in my life.	Knowing English isn't really an important goal in my life.
DLE7 Negatively keyed	I find I'm losing any desire I ever had to know French.	I find I'm losing any desire I ever had to know English.
DLE8 Negatively keyed	To be honest, I really have little desire to learn French.	To be honest, I really have little desire to learn English.
DLE9 Negatively keyed	I haven't any great wish to learn more than the basics of French.	I haven't any great wish to learn more than the basics of English.

Prior to the selection of the AMTB, two other instruments that measure constructs related to L2 learning were examined. The instruments were Foreign Language Classroom Anxiety (FLCAS), which specifically measures anxiety (Horwitz et al. 1986), and the Beliefs about Language Learning Inventory (BALLI), which specifically measures beliefs about language learning. However, AMTB (Gardner, 1985, 2001) was chosen for this study because it covers all of the language-related constructs needed for this study, and because these constructs were validated together in prior research.

4.8 Methods of Data Analysis

This section discusses how the experiment (relating to MUVE effectiveness) and the survey (relating to MUVE acceptance) data were analysed. The discussion covers both the approaches to hypotheses testing and the associated validity checks.

4.8.1 MUVE Effectiveness

This section discusses the approach taken to measure oral business English communication proficiency in the context of the adopted learning scenario, as well as the statistical procedures employed to test the hypotheses E1 and E2 (see section 3.2). Then, the section highlights the approach taken to measure inter-rater reliability, which is necessary to demonstrate the validity of the measurements of oral business English communication proficiency.

4.8.1.1 Assessing Learning Gains

In order to test the effectiveness of MUVE as a learning environment, oral outputs before and after teaching in the classroom environment and MUVE were recorded and analysed. The details of the procedure are described in section 4.5.

The oral output before and after exposure to both classroom teaching and MUVE was recorded separately, and later coded blindly by two raters using the rating scale designed for the study (see section 4.7.1). The raters did not know if

the recordings they were rating were from post or from pre tests, or from students who learned in MUVE or in the classroom. Further, they were not provided with any information that would allow them to determine if two recordings were for the same learner. Inter-rater reliability is discussed in section 4.8.1.2.

A paired samples *t*-test was used to compare pre-test and post-test scores separately for learners who learned in MUVE and in the classroom. Moreover, a paired samples *t*-test was also conducted to compare learning gains between the two environments.

4.8.1.2 Ensuring Inter-rater Reliability

Inter-rater reliability is achieved when different coders arrive at the same or similar coding decisions when evaluating content (Rourke, Anderson, Garrision, & Archer, 2001).

To establish inter-rater reliability for this study, the recorded conversations were evaluated by two raters using the oral interaction assessment scale: the main rater, who rated all recordings, and the reliability rater, who rated only a subset of the recordings to provide data for inter-rater reliability calculations. The raters were selected based on their academic qualifications in teaching English as a second language. The raters never met and worked on the basis of the rating scheme (see Appendix D for details of the scheme).

There are a number of indexes used to report inter-rater reliability; in this study, a measurement of correlation coefficient — the Pearson R was used, along with Cronbach Alpha. Reliability was sought for the overall score calculated by adding up the scores for the individual items of the scale. Inter-rater reliability indexes that emphasise correlation between numerical values, rather than matching categorical variables, were chosen for this study because the measure of communication performance was numeric, rather than categorical.

4.8.2 MUVE Acceptance

This section discusses the approach taken to test the research model for learner acceptance of MUVE based e-Learning formulated in section 3.3, thus testing the

hypotheses A1 to A14. Approaches taken to assess construct validity and to minimize the possibility of common method bias are also discussed.

4.8.2.1 Structural Equation Modeling

To test the model of MUVE e-Learning acceptance (see section 3.3), structural equation modelling (SEM) was used. SEM involves the use of latent and manifest variables. Latent variables (constructs) are the variables which cannot be measured (not observable) directly, so latent variables are inferred from a set of variables called manifest variables (observed variables) that researchers obtain and measure through using tests, surveys, and other methods (Schumaker & Lomax, 2004). A SEM describes how latent variables are measured, as well as hypothesised cause-effect relationships between the latent variables. Thus, by fitting data to a SEM, one can test several hypotheses at the same time.

A SEM can be viewed as comprising two separate models: a measurement model and a structural model (Kline, 2005). The measurement model defines the relationships between the latent variables and the manifest variables (also known as indicators) used to measure them. In case of data obtained via Likert-style survey questions (as in the survey used in this study), the values of indicators correspond to the participants' responses to survey items.

The structural model defines the relationships between different latent variables. Structural model evaluation is the assessment of causal relationships between constructs in the model (in an assumption the statistical patterns in the data are due to cause-effect relationship). Fitting the structural model results in estimates of the path coefficients (ß) for each hypothesised cause-effect relationship (a path in the model) and of explained variance (R^2) for each dependent variable (variable hypothesised to depend on at least one other variable).

There are two alternative approaches to testing SEM: covariance-based analysis (implemented in software such as AMOS and Mplus) and component-based analysis implemented in the software such as Smart PLS (Falk & Miller, 1992; Barclay, Higgins, & Thompson, 1995). Component based analysis is more suitable for smaller samples, for large SEM models, and for non-normal data. Component based analysis is recommended for use as an exploratory technique in cases where

the theory is not well established (Henseler, Ringle, & Sinkovics, 2009). Covariance based analysis is more suitable for larger samples and requires multivariate normal data (although data deviating from normality is commonly used in practice). Covariance based analysis is recommended for use as a confirmatory technique—in cases where the theory is well established (Henseler et al. 2009). In accordance with these recommendations, this study uses component based analysis to explore a larger initial research model, and then confirms the results by applying a covariance based technique to fit a smaller model retaining only the relationships found to be significant in the initial PLS analysis.

4.8.2.2 Partial Least Squares Limitations

MUVE e-Learning acceptance survey data were analysed with PLS (Chin & Newsted, 1996; Chin, 2003). PLS has been used extensively in MIS research, especially in testing technology acceptance models. PLS has an established record as a reliable technique (Cool, Dierickx, & Jemison, 1989; Venkatesh & Davis, 2000; Yoo & Alavi, 2001; Kahai & Cooper, 2003; So & Bolloju, 2005).

In view of the recent criticism of PLS claiming that its estimates may be biased (Marcoulides & Saunders, 2006), the present study used covariance based approaches to confirm the results of PLS analysis (see sections 5.3.2 and 5.3.3).

4.8.2.3 Ensuring Construct Validity

Construct validity is the extent to which latent variables are correctly represented by their manifest variables. It is not possible to assess the construct validity directly. To ensure construct validity, this study reused measurement items from prior studies. In addition, convergent validity (and related reliability measures) and discriminant validity were assessed by using PLS (see section 5.3.1.1) and EFA (section 5.3.2) analyses, see the respective sections for details. Convergent and discriminant validity serve as evidence of construct validity (Bagozzi, Yi, & Phillips, 1991).

4.8.2.4 Guarding for Common Method Bias

In this study, data were obtained from a single questionnaire, which may have inflated some of the relationships found due to common method bias (Whitman & Woszczynski, 2004; Burton-Jones & Hubona, 2005;). However, since all measures were derived from established instruments with good psychometric properties, it may reduce the risk of common method bias undermining the study results (Podsakoff et al., 2003).

Podsakoff et al. (2003) provided guidelines to reduce common method bias. One guideline was to ensure anonymity in the survey administration; and another was to improve the items used to measure constructs. In this study, both of these suggestions were followed. Participants were assured of anonymity; anonymity means greater levels of calmness and honesty. Because of these measures, they are less likely to change their responses to make them more socially acceptable (Podsakoff et al., 2003). Also, extensively tested previous test scales were used in the study. The use of validated scales reduces item ambiguity and thus reduces the possibility of response sets (behaviour when participants rate many items at the same level to deal with cognitive load imposed by the survey), meeting a key recommendation by Podsakoff et al. (2003).

4.9 Software Used for Statistical Analysis

This section introduces the software used for statistical analysis. A range of software tools was used, including R, SmartPLS, AMOS, and Mplus.

4.9.1 Software Used to Compare Learning Gains

Comparing learning gains in the experiment used to address research question one involved conducting paired and unpaired *t*-tests (see section 5.2.2). In addition, a number of tests were conducted to verify that the data meets the conditions for using *t*-tests (see section 5.2.2.2). All of these statistical tests were conducted using R version 2.13.

4.9.2 Software Used for PLS and Covariance-based SEM Analysis

Testing the MUVE e-learning acceptance model (Figure 16) involved SEM modelling using both components based and covariance based techniques. Component based partial least squares analysis and covariance based analysis approaches to structural equation modelling are introduced in section 4.8.2.1. This section discusses the specific software used for the analysis.

The analysis was conducted using SmartPLS (version V2M3) software (Ringle, Wende, & Will, 2005) for PLS analysis and AMOS 7.0 for covariance-based SEM analysis. AMOS is a covariance-based tool that uses a maximum likelihood function to obtain parameter estimates, whereas SmartPLS is a component-based tool that uses a partial least squares estimation procedure in its analysis. SmartPLS reports internal composite reliability and average variance extracted (AVE) needed for convergent validity and discriminant validity. As the underlying algorithms in PLS and in covariance based analysis are very different, when results obtained by these two approaches coincide, it constitutes a stronger evidence of their validity compared with when only one of the approaches is used.

While PLS analysis is best suited for exploratory analysis (tentative testing of models when the theory is weak), covariance based analysis is suitable for confirmatory analysis (testing established theories). Unlike PLS, covariance-based analysis provides various goodness-of-fit indices to assess overall model fit and allows formulation and testing of the measurement model separately from the structural model (Byrne, 2001). Correspondingly, PLS was used for initial testing of the model, which in its initial form was too big for covariance based analysis. Then, to confirm PLS results, covariance based analysis was applied using AMOS. To test data for multivariate normality (a prerequisite for using covariance based SEM with a maximum likelihood estimator), R version 2.13 was used, packages fUtilities version 2110.78 and QRMlib version 1.4.5.1.

As discussed in section 4.8.2.2, the partial least squares approach was recently a subject of criticism. To counter such criticism, it was particularly important to verify PLS results by using alternative approaches. Arguably, the consistency of the results that were obtained by using different approaches may be seen as a contribution to the debate about the usefulness of PLS in management information

systems research. In particular, consistent results can be seen as a contribution in favour of continuing the use of PLS.

4.9.3 Software Used for EFA Analysis

As a separate step to confirm PLS measurement model results, exploratory factors analysis was employed using Mplus 6.0 with a maximum likelihood estimator to independently verify the convergent and the discriminant validity of the measurement instruments (see section 5.3.2). To verify that the data meets distributional assumptions for the use of the maximum likelihood estimator, R version 2.13 was used, package fUtilities version 2110.78.

4.10 Human Ethics

Before carrying out the research, human ethics approval is a part of the procedure to ensure that no harm is done to the participants or the setting of the research. In this study, the application for permission to conduct research went through three phases. The first phase was filling in the screening questionnaire to determine the approval procedure. Since the questionnaire result indicated that this study fell into the category of a low risk research, it did not require approval from Massey University Human Ethics Committee. The second phase was filling in the low risk notification document, which involved the research being peer reviewed by the researcher's supervisors. The peer review process involved reviewing the intention of the research, the research methodology, and other relevant documents, such as research instruments, students' biodata forms and questionnaires, the information sheet, and the consent form. In the third phase, once the peer review process was completed, the application document was submitted to the Massey University Ethics committee, an acknowledgement was received indicating that the low risk notification had been considered and accepted. In addition to the acknowledgement from the Massey University Ethics committee, permissions to conduct this research were obtained from the Economic Planning Unit of the Prime Minister's Department in Malaysia and from the Deputy Vice-Chancellor as well as the Dean of the faculty of the institution in Malaysia where the study was conducted.

All the students who participated in this study were volunteers. Before the initiation of this research, written consent of the participants was obtained for the collection and analysis of the data. During the meeting with each pair of students, an information sheet containing details about the purpose, participant involvement, and participant's rights was distributed to the participants (Appendix A). The consent form to obtain the participant's permission to audio record scenario-based learning activities was also distributed (Appendix B). Each participant read the information sheet and the consent form and signed it to indicate agreement. The consent forms were collected after being signed.

The participants were assured of strict confidentiality of their identities. Any information about them was used only for the present research and for publications arising from the project, in an aggregated form. Participants' names were changed in the audio files and procedures were implemented to ensure safe storage of data. All audio recordings and answered questionnaires were carefully labelled and stored in a secured, locked cabinet. The consent forms were sealed and stored separately in another cabinet. In compliance with Massey University policy on research practice, the data is to be retained for five years and at the end of the retention period will be disposed of.

Care was taken to ensure that participants understood the information about the study. Information sheets and consent forms were written as clearly as possible in English. It was not necessary to translate the information sheet and consent form into the participants' first language because all of them had a fairly good command of English, having studied the language for at least eleven years in school.

When the volunteers were identified, full information was provided to them about the purpose, the extent of their involvement in the study, and what was expected of them (in the form prescribed by Massey University human ethics regulations). The participants were told that they would not be exposed to any risks other than what normally happened in a usual classroom situation. Whatever social or physical discomfort they might experience as a result of the participation would be disagreements of ideas with other group members. In the form prescribed by Massey University human ethics regulations, the participants were informed that they can decline participation, refuse to answer any particular questions, request

their dialogue not be recorded, or withdraw from the study at any time. None of the participants in the study, however, requested any of these options.

4.11 Pilot Study

Before actual research is conducted, it is generally advisable to try out the proposed research instruments and procedures on a few respondents, who are representative of the target population as a whole (Bryman & Bell, 2003). Therefore, prior to carrying out the research, a pilot study was conducted to find out whether the research procedure was workable and achievable within the time frame suggested. Ten Malay learners from a university in New Zealand (all speakers of English as a second language) were invited to take part in the pilot survey. They were not involved in the actual study. Learners commented on the scenario-based learning activity and on items in the MUVE e-Learning acceptance questionnaire. Comments covered the understandability of the scenario-based learning activity instructions, wording of scales, length of the instrument, and format of the questionnaire. Feedback received from the participants was positive; they commented that they felt comfortable with the time given to complete the survey and could understand the content of the questionnaire, and there was no need to modify the research instruments and the procedures. The pilot study demonstrated that the research was feasible.

4.12 Summary

This chapter described the research methodology used in this study. The research procedures and data collection methods were discussed in detail. In addition, the chapter outlined the human ethics approval procedure and described a pilot study conducted prior to undertaking the actual research.

Chapter 5: Results

5.1 Introduction

This chapter begins by presenting the results of the experiment on MUVE effectiveness. Learner conversations on the effectiveness of MUVE recorded in the experiment were rated by two raters using the rating scale introduced in Chapter 4. The consistency between the two raters was confirmed by using inter-rater reliability analysis and was found to be satisfactory. Learning gains were measured based on the classroom and MUVE group performance, respectively. According to a t-test, there were statistically significant learning gains both in MUVE and in the classroom. However, the difference between the learning gains in the classroom and in MUVE was not statistically significant.

The chapter then moves to a discussion of the results of the cross-sectional survey used to measure factors that are likely to influence Malay learner acceptance of MUVE. Before assessing the fit of the structural model, the measurement model was assessed by estimating the reliability and the convergent and discriminant validity of the model's constructs. The assessment of the measurement model resulted in a number of items that caused reliability and validity problems being eliminated from the model, to meet the commonly accepted criteria for reliability and validity (the criteria, as well as the possible reasons why the items have led to reliability and validity problems, are discussed in section 5.3.1.1). Then, the structural model was assessed, providing results for hypothesis testing. The TAM hypotheses were confirmed (with the exception of the relationship between the Perceived Ease of Use and Intention to Use). For constructs representing social cognitive theory, the Video Games Affect and the English Class Anxiety affected the Intention to Use (although the effect size was rather small). However, the direct effects of other social cognitive theory constructs on Intention to Use were not found to be statistically significant. The initial assessments of both the measurement model and the structural model were conducted by using the PLS approach, using SmartPLS software.

The chapter concludes by presenting additional analyses of MUVE e-Learning acceptance data conducted to confirm the result of the analysis using PLS. To confirm the results obtained by PLS, covariance-based techniques were applied. To confirm the discriminant and convergent validity, exploratory factor with a maximum likelihood (ML) estimator was applied, using Mplus software. The results were consistent with the PLS results for the measurement model. A reduced version of the research model (with only the relationships found to be significant in PLS analysis retained) was tested by using a covariance-based technique (using Amos software). Both the measurement and the structural model were assessed. The results were consistent with the result obtained by PLS.

5.2 Results for MUVE Effectiveness

This section discusses the results related to research question 1 (see section 1.4): Is MUVE based learning effective in facilitating situated scenario-based learning of oral business English communication skills by Malay learners? This research question was addressed by hypotheses E1 and E2 (see section 3.2).

These hypotheses were tested by using an experiment. The corresponding data collection procedures are described in sections 4.5 and 4.8.1.1. The rating scale used to measure communication performance is described in section 4.7.1, and the procedures for checking inter-rater reliability (needed to support the validity of the measurement) are described in section 4.8.1.2. The remainder of this section discusses the outcomes of verifying inter-rater reliability, followed by a presentation of the results of the experiment, with descriptive results followed by the outcomes for hypotheses E1 and E2.

5.2.1 Inter-rater Reliability

Inter-rater reliability was verified to confirm the validity of the rating scale. Post-test recordings for the first 24 pairs of participants were rated by two raters, as discussed in section 4.8.1.2. (rating for a pair was obtained by adding up the ratings for individual participants in the pair.) The scores are given in Table 16.

Table 16: Scores Between the Two Raters

	Reliability rater	Main rater
Pair 1	45	42
Pair 2	45	48
Pair 3	46	45
Pair 4	46	49
Pair 5	48	52
Pair 6	50	48
Pair 7	50	44
Pair 8	52	45
Pair 9	52	52
Pair 10	52	45
Pair 11	53	54
Pair 12	53	49
Pair 13	54	49
Pair 14	55	54
Pair 15	55	47
Pair 16	55	56
Pair 17	56	51
Pair 18	57	56
Pair 19	57	54
Pair 20	58	57
Pair 21	59	55
Pair 22	60	52
Pair 23	60	53
Pair 24	60	58

The inter-rater reliability values were 0.70 (using the Pearson correlation coefficient) and 0.82 (using Cronbach's alpha). According to Straub, Boudreau, and Gefen (2004), who suggested a cut-off value of 0.70 for inter-rater reliability in MIS research, these values correspond to acceptable levels of inter-rater reliability.

5.2.2 Learning Gains

This section discusses the descriptive statistics for gains in oral business English communication skills in both MUVE and the traditional classroom. This is followed by the results of hypothesis testing.

5.2.2.1 Descriptive Statistics for Learning Gains—Outcome for Hypothesis E2

Pre-test and post-test scores for both groups (A and B, corresponding classroom and Second Life teaching, respectively) are presented in the Table 17.

Table 17 shows that both groups (classroom and MUVE) achieved higher average scores (arithmetic means) in the post-test than in the pre-test. The average scores of the learners in the classroom teaching environment increased from 48.32 to 52.97. On the other hand, the average scores of the learners in the Second Life teaching environment increased from 49.15 to 53.41. Both in the classroom and in MUVE, the Cohen's effect size (obtained by dividing the mean learning gain by the standard deviation) was close to 0.8. Thus, according to Cohen (1988), the effect size for the learning gains both in the classroom and in MUVE were large. The mean learning gain in the classroom was slightly higher than in MUVE. The Cohen's effect size for this difference was close to 0.2, and thus, according to Cohen's classification, was small.

The slightly higher mean learning gain in the classroom meant that the hypothesis E2 was not confirmed. The experiment provided no evidence in favour of MUVE being more effective than the classroom environment.

Table 17: Pre-test and Post-test Scores of Classroom and Second Life

Group A: Classroom Teaching				Group B: Second Life Teaching			
Pair	Pre-test score	Post-test score	Difference	Pair	Pre-test score	Post-test score	Difference
1	36	49	13	2	39	45	6
3	52	54	2	4	50	55	5
5	45	58	13	6	48	48	0
7	45	44	-1	8	49	49	0
10	44	53	9	9	50	59	9
11	42	54	12	12	51	45	-6
14	55	57	2	13	49	55	6
16	54	58	4	15	54	53	-1
17	47	56	9	18	49	56	7
19	44	44	0	20	53	49	-4
21	51	47	-4	22	43	48	5
24	48	54	6	23	49	38	-11
25	59	52	-7	26	33	55	22
28	54	49	-5	27	52	58	6
30	40	49	9	29	51	55	4
31	52	59	7	32	54	54	0
34	47	55	8	33	53	55	2
35	57	49	-8	36	51	55	4
37	49	51	2	38	47	55	8
39	54	60	6	40	54	57	3
42	55	55	0	41	44	48	4
44	50	59	9	43	53	62	9
45	50	59	9	46	54	58	4
48	50	49	-1	47	58	57	-1
50	51	56	5	49	42	49	7
51	52	51	-1	52	55	62	7
53	39	47	8	54	61	61	0
56	46	55	9	55	33	48	15
57	46	60	14	58	50	58	8
59	51	54	3	60	51	57	6
61	42	57	15	62	57	59	2
63	37	40	3	64	43	50	7
65	46	52	6	66	48	56	8
67	53	55	2	68	43	47	4
Mean	48.32	52.97	4.65	Mean	49.15	53.41	4.27
Standard deviation	5.68	4.98	5.96	Standard deviation	6.28	5.53	5.79

5.2.2.2 Checking Prerequisites for Using *t*-tests

Quantile-quantile plots and the Shapiro-Wilk test were used to check for normality of pre- and post-test score distributions. Only the distribution of pre-test scores for Group B (Second Life) did not pass the Shapiro-Wilk test (with $p = 0.03$), but was judged to be close enough to normal based on its quantile-quantile plot. This demonstrated that it was justifiable to use *t*-tests to test for statistical significance of learning gains in each group, and to test the difference in learning gains between the two groups for statistical significance.

5.2.2.3 Testing if There Were Learning Gains in the Classroom

A paired *t*-test was employed to test the significance of the learning gains in the classroom. The results are shown in Table 18.

Table 18: The Results of The Paired t-test for Classroom Learning

Mean difference	*t*	df	*p* value
4.65	-4.55	33	<0.001

As seen from the table, the learning gains in the classroom were statistically significant. This confirms the validity of the experiment design: The length of the intervention was sufficient to result in an improvement of the communication skills of the participants in the targeted application area (selling a house).

5.2.2.4 Testing if There Were Learning Gains in MUVE—Testing the Hypothesis E1

A paired *t*-test was employed to test the significance of the learning gains in MUVE. The results are shown in Table 19.

Table 19: The Results of The Paired *t*-test for Second Life

Mean difference	*t*	df	*p* value
4.27	-4.29	33	<0.001

As seen from the table, the learning gains in the Second Life were statistically significant. This confirms hypothesis E1: learning in Second Life resulted in a statistically significant improvement of the communication skills of the participants.

5.2.2.5 Exploring the Difference Between the Learning Gains in MUVE and in the Classroom—Testing the Inverse of the Hypothesis E2

Higher learning gains in the classroom meant that hypothesis E2 was not supported, and suggested an alternative hypothesis—chat classroom was a more effective environment than Second Life. An unpaired *t*-test was employed to test this hypothesis. The results are shown in Table 20.

Table 20: The Results of The Unpaired t-test Between Classroom Learning and Second Life

	Mean	*t*	df	*p* value
Classroom teaching	4.65			
		0.26	66	0.6
Second Life teaching	4.27			

As seen from the table, the difference was not statistically significant (p=0.60). Thus, the statistical power of the experiment was not sufficient to conclude which environment (Second Life or classroom) is more effective.

5.3 Results for MUVE Acceptance

This section discusses the results related to research question 2 (see section 1.4): Which factors influence the acceptance of MUVE based learning of oral business English communication skills by Malay learners? This research question was addressed by the research model introduced in section 3.3, which comprises hypotheses A1 to A14.

These hypotheses were tested by using a cross-sectional survey. The corresponding data collection procedures are described in section 4.5. The questionnaire used to collect data relating to MUVE acceptance is described in section 4.7.2, and the procedures for checking the instrument validity are introduced in section 4.8.2.3. The rest of this section discusses the outcomes of testing the measurement model (thus verifying the reliability and the validity of the measurement instruments) and presents the results of testing the structural model (and thus testing the hypotheses A1 to A14).

Testing the measurement model involved assessing convergent and discriminant validity. Convergent and discriminant validity were verified by using partial least squares (section 5.3.1.1) and exploratory factor analysis (section 5.3.2). In addition, for a subset of the research model including only the constructs found to be affecting MUVE acceptance in partial least squares analysis, covariance based confirmatory factor analysis was applied (section 5.3.3), resulting in acceptable values of global indices of fit.

The structural model (comprising the hypotheses A1 to A14) was first tested by using partial least squares (section 5.3.1.2). Then, for a subset of the research model including only the constructs found to be affecting MUVE acceptance in partial least squares analysis, covariance based structural equation modelling was applied (section 5.3.3).

5.3.1 Testing the Research Model Using PLS

The following two sections are devoted to testing the measurement model and the structural model for the model of MUVE e-Learning acceptance (see Figure 16)

using PLS. PLS fits the measurement model simultaneously with the structural model.

First, the PLS algorithm was applied to the data collected in the MUVE e-Learning acceptance survey with the construct measure defined as described in section 3.3. The measurement model was analysed and it was found that some of the indicators were unreliable and had to be removed. Subsequent testing of the structural model was conducted by using the updated construct measures. An approach in which the measurement model and the structural model are fitted to different data sets would be preferable, as it would reduce the possibility of the capitalisation of chance. (When the measurement model is adjusted, by removing indicators, to fit a particular data set and then the structural model relying on the adjusted measurement model is fitted to the same data set, the fit of the structural model may be artificially high.) However, it was not possible to obtain enough data to fit the two models separately; therefore, similar to the studies by Bourgonjon et al. (2013) and Wojciechowski & Cellary (2013), the same data set was used.

The validity of PLS analysis does not depend on distributional assumptions (Tenenhaus, Vinzi, Chatelin & Lauro, 2005). Tenenhaus et al. (2005) suggest unidimensionality requirements for all constructs should be met for PLS analysis results to be meaningful. The definition of unidimensionality according to Tenenhaus et al. (2005) coincides with the definition of internal consistency reliability which is demonstrated for the MUVE acceptance data in section 5.3.1.1.

5.3.1.1 Test of the Measurement Model—Convergent and Discriminant Validity

Convergent validity reflects the extent to which the indicators of a construct are similar to the other indicators of the same construct. Using the SmartPLS software, the convergent validity of each construct was assessed by (a) computing the reliability of indicators (item reliability of each measure), (b) composite reliability of each construct, (c) Cronbach's alpha, and (d) average variance extracted by the construct (Fornell & Larcker, 1981; Hair, Anderson, Tatham, & Black, 1998).

The reliability of indicators (item reliability of an item) was assessed by their loadings onto their underlying constructs. On the basis of the recommendations by

Gefen & Straub (2005), indicators with loadings of 0.7 or greater were considered as reliable.

In the following tables, item reliability in constructs is discussed starting from constructs related to Video Games Self-efficacy, followed by constructs relating to Subject Matter Self-efficacy, and by TAM constructs. The main dependent construct—Intention to Use MUVE for e-Learning—is discussed last.

Table 21 shows that three of the ten items of the construct Video Games Self-efficacy had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study (Gefen & Straub, 2005). Items VGSE1, VGSE2, VGSE3, VGSE4, VGSE7, VGSE8 and VGSE10 were removed from the construct due to poor loading. It appears that the approach used to formulate items for this measure (heavily relying on conditional statements) made them difficult for the participants to understand. The three items that were retained were the ones with the clearest meaning and all referred to getting help from someone with playing the game. Therefore, even with just these three items remaining, the measure still had face validity. There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 21: Factor Structure Matrix of Loadings and Cross-loadings for the Video Games Self-efficacy Construct

	VGSE	VGAN	VGAFF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
VGSE1	**-0.03**	-0.11	0.24	0.02	0.06	0.03	0.18	0.15	0.16	0.11
VGSE2	**-0.23**	-0.20	0.36	-0.01	0.08	0.09	0.24	0.17	0.13	0.13
VGSE3	**0.40**	0.10	0.15	-0.08	-0.00	0.06	-0.02	-0.04	0.07	0.01
VGSE4	**0.62**	0.30	0.02	-0.00	0.02	-0.05	0.10	0.00	0.02	0.01
VGSE5	**0.70**	0.28	0.04	-0.03	0.01	-0.05	0.04	0.05	0.06	0.05
VGSE6	**0.77**	0.23	-0.01	-0.13	0.16	0.05	0.12	0.11	0.06	0.05
VGSE7	**0.44**	0.14	0.11	0.06	-0.07	0.20	0.22	0.20	0.16	0.13
VGSE8	**0.64**	0.21	0.01	-0.03	0.02	0.18	0.23	0.20	0.26	0.22
VGSE9	**0.79**	0.27	-0.04	-0.10	0.25	0.08	0.18	0.08	0.09	0.07
VGSE10	**0.52**	0.12	0.07	0.01	0.13	0.20	0.24	0.20	0.24	0.19

Table 22 shows that three of the four items of the construct Video Games Anxiety had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study. Item VGAN4 was removed from the construct due to poor loading. To fit

the video games playing context, this item was reworded considerably compared to its wording in the instrument on which this instrument was based and the new wording did not work. There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 22: Factor Structure Matrix of Loadings and Cross-loadings for Video Games Anxiety Construct

	VGSE	VGAN	VGAFF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
VGAN1	0.34	**0.77**	-0.10	0.02	-0.16	0.00	0.03	-0.07	-0.05	-0.14
VGAN2	0.29	**0.76**	-0.21	0.01	-0.05	0.01	0.03	-0.11	-0.12	-0.09
VGAN3	0.28	**0.76**	-0.15	0.03	-0.05	-0.04	0.03	-0.07	-0.06	-0.02
VGAN4	0.17	**0.64**	-0.39	-0.06	0.04	-0.13	-0.12	-0.19	-0.18	-0.14

Table 23 shows that all of the five items of the construct Video Games Affect had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study. Thus, none of the items were removed from the construct. There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 23 : Factor Structure Matrix of Loadings and Cross-loadings for the Video Games Affect Construct

	VGSE	VGAN	VGAFF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
VGAFF1	-0.03	-0.26	**0.87**	-0.07	0.21	0.03	0.07	0.21	0.17	0.29
VGAFF2	-0.09	-0.17	**0.84**	0.00	0.06	0.03	-0.01	0.14	0.11	0.22
VGAFF3	-0.04	-0.24	**0.89**	0.04	0.09	0.07	0.07	0.28	0.22	0.33
VGAFF4	0.05	-0.12	**0.70**	0.02	0.15	0.00	-0.02	0.07	0.05	0.14
VGAFF5	0.01	-0.23	**0.78**	0.13	0.01	0.06	0.04	0.24	0.21	0.29

Table 24 shows that six of the ten items of the construct Subject Matter Self-efficacy had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study. Items SMSE2, SMSE4, SMSE6, and SMSE7 were removed from the construct due to poor loading. (Again, the removed items appeared to be the ones using more complex wording, although the difference was not clear-cut).

There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 24 : Factor Structure Matrix of Loadings and Cross-loadings for the Subject Matter Self-efficacy Construct

	VGSE	VGAN	VGAFF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
SMSE1	-0.09	0.02	-0.02	**0.70**	-0.54	0.24	0.11	-0.00	-0.03	-0.08
SMSE2	-0.05	0.01	0.20	**0.47**	-0.22	0.15	0.18	0.21	0.12	0.17
SMSE3	-0.06	0.02	0.02	**0.76**	-0.56	0.25	0.10	0.02	-0.05	-0.09
SMSE4	-0.05	-0.09	0.12	**0.33**	-0.31	0.13	0.06	0.01	0.04	0.06
SMSE5	-0.06	-0.01	-0.03	**0.77**	-0.50	0.24	0.18	0.13	0.09	0.05
SMSE6	-0.14	-0.23	0.06	**0.56**	-0.25	0.26	0.20	0.20	0.21	0.21
SMSE7	0.10	0.12	-0.04	**0.62**	-0.36	0.21	0.10	0.07	0.08	-0.01
SMSE8	-0.05	0.09	0.02	**0.80**	-0.43	0.27	0.15	0.15	0.13	0.09
SMSE9	0.00	-0.02	0.03	**0.77**	-0.37	0.36	0.27	0.26	0.18	0.08
SMSE10	-0.02	0.04	-0.01	**0.78**	-0.40	0.34	0.22	0.15	0.14	0.06

Table 25 shows that six of the ten items of the construct English Class Anxiety had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study. Items ECAN6, ECAN7, ECAN8, ECAN9, and ECAN10 (all of the negatively keyed items) were removed from the construct due to poor loading. There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 25: Factor Structure Matrix of Loadings and Cross-loadings for the English Class Anxiety Construct

	VGSE	VGAN	VGAF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
ECAN1	0.13	0.01	0.05	-0.39	**0.72**	-0.07	0.14	0.07	0.11	0.24
ECAN2	0.13	-0.10	0.05	-0.37	**0.74**	-0.25	-0.02	0.12	0.08	0.14
ECAN3	0.10	0.05	0.11	-0.38	**0.70**	-0.08	0.16	0.12	0.12	0.18
ECAN4	0.00	-0.06	0.08	-0.43	**0.74**	-0.18	0.01	0.13	0.17	0.25
ECAN5	0.22	-0.00	0.12	-0.51	**0.78**	-0.10	0.14	-0.01	0.06	0.19
ECAN6	-0.03	0.09	-0.05	0.22	**-0.24**	-0.04	-0.02	-0.06	-0.05	-0.11
ECAN7	-0.02	0.12	-0.08	0.58	**-0.78**	0.22	0.07	0.02	-0.10	-0.18
ECAN8	-0.05	0.17	-0.14	0.26	**-0.51**	0.12	0.09	-0.04	-0.09	-0.23
ECAN9	0.08	-0.02	-0.03	0.39	**-0.62**	0.10	0.08	0.03	0.03	-0.06
ECAN10	0.07	0.21	-0.06	0.26	**-0.47**	0.15	-0.08	-0.08	-0.10	-0.13

Table 26 shows that six of the nine items of the construct Attitude Towards Learning English had factor loadings of 0.70 or greater, indicating that they are

satisfactory for this study. Items ATLE5, ATLE6, ATLE7, ATLE8 and ATLE9 were removed from the construct due to poor loading. These items were negatively keyed. (It appears that some of the respondents found it difficult to understand negatively keyed items. Most of the items discovered to be problematic in measurement model analysis were negatively keyed items.) There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 26: Factor Structure Matrix of Loadings and Cross loadings for the Attitude Towards Learning English Construct

	VGSE	VGAN	VGAF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
ATLE1	0.14	-0.01	0.06	0.35	-0.08	**0.73**	0.44	0.38	0.37	0.36
ATLE2	0.05	0.01	0.09	0.38	-0.21	**0.84**	0.47	0.33	0.37	0.27
ATLE3	0.08	0.04	0.01	0.41	-0.28	**0.81**	0.45	0.35	0.33	0.22
ATLE4	0.06	-0.01	0.03	0.26	-0.09	**0.78**	0.55	0.33	0.35	0.31
ATLE5	-0.07	0.05	0.05	-0.20	0.10	**-0.61**	-0.40	-0.27	-0.31	-0.26
ATLE6	0.11	0.09	0.07	-0.19	0.31	**-0.55**	-0.22	-0.16	-0.17	-0.05
ATLE7	-0.02	0.12	-0.11	-0.18	0.15	**-0.72**	-0.42	-0.26	-0.32	-0.21
ATLE8	-0.13	0.08	-0.00	-0.07	0.00	**-0.63**	-0.50	-0.33	-0.37	-0.31
ATLE9	-0.03	0.09	-0.06	-0.17	0.12	**-0.74**	-0.48	-0.27	-0.31	-0.24

Table 27 shows that three of the nine items of the construct Desire to Learn English had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study. Items DLE1, DLE2, DLE6, DLE7, DLE8, and DLE9 were removed from the construct due to poor loading. Of the removed items, DLE7, DLE8, and DLE9 were negatively keyed, and DLE1 and DLE2 used more complex expressions than the items that were retained. There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 27: Factor Structure Matrix of Loadings and Cross-loadings for Desire to Learn English Construct

	VGSE	VGAN	VGAF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
DLE1	0.28	0.02	0.13	-0.04	0.24	0.24	**0.61**	0.23	0.29	0.31
DLE2	0.13	0.08	-0.09	0.24	-0.16	0.47	**0.56**	0.18	0.22	0.12
DLE3	0.07	-0.07	0.05	0.21	0.03	0.45	**0.77**	0.34	0.36	0.30
DLE4	0.07	0.00	0.02	0.24	-0.03	0.56	**0.78**	0.28	0.34	0.25
DLE5	0.14	-0.03	0.15	0.13	0.12	0.44	**0.72**	0.36	0.39	0.36
DLE6	0.07	0.09	-0.02	0.04	-0.13	-0.06	**-0.27**	-0.11	-0.08	-0.10
DLE7	-0.06	-0.06	0.02	-0.09	-0.02	-0.21	**-0.44**	-0.12	-0.09	-0.07

	VGSE	VGAN	VGAF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
DLE8	0.01	-0.13	0.15	-0.08	0.03	-0.24	**-0.35**	-0.05	-0.15	-0.07
DLE9	-0.05	0.01	0.02	-0.17	0.09	-0.36	**-0.54**	-0.24	-0.30	-0.22

Table 28 shows that all of the six items of the construct Perceived Ease of Use had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study. None of the items were removed from the construct. There was no evidence of cross loading for this construct— all items loaded the strongest to their own construct.

Table 28 : Factor Structure Matrix of Loadings and Cross-loadings for Perceived Ease of Use

	VGSE	**VGAN**	**VGAFF**	**SMSE**	**ECAN**	**ATLE**	**DLE**	**PEOU**	**PU**	**INT**
PEOU1	0.11	-0.03	0.18	0.16	-0.02	0.26	0.20	**0.71**	0.53	0.49
PEOU2	0.02	-0.19	0.20	0.07	0.11	0.22	0.21	**0.70**	0.41	0.37
PEOU3	0.14	-0.10	0.25	0.10	0.05	0.35	0.32	**0.86**	0.67	0.63
PEOU4	0.18	-0.05	0.20	0.06	0.14	0.39	0.37	**0.83**	0.68	0.64
PEOU5	0.09	-0.12	0.16	0.19	0.07	0.34	0.38	**0.81**	0.62	0.50
PEOU6	-0.03	-0.18	0.19	0.22	0.04	0.38	0.36	**0.75**	0.59	0.52

Table 29 shows that all of the six items of the construct Perceived Usefulness had factor loadings of 0.70 or greater, indicating that they are satisfactory for this study. None of the items were removed from the construct. There was no evidence of cross loading for this construct— all items loaded the strongest to their own construct.

Table 29 : Factor Structure Matrix of Loadings and Cross-loadings for Perceived Usefulness

	VGSE	**VGAN**	**VGAFF**	**SMSE**	**ECAN**	**ATLE**	**DLE**	**PEOU**	**PU**	**INT**
PU1	0.21	-0.08	0.22	0.12	0.14	0.42	0.43	0.66	**0.88**	0.72
PU2	0.01	-0.16	0.18	0.11	0.08	0.43	0.44	0.62	**0.88**	0.70
PU3	0.05	-0.12	0.14	0.12	0.08	0.41	0.36	0.64	**0.87**	0.70
PU4	0.18	-0.07	0.17	0.15	0.09	0.43	0.46	0.65	**0.86**	0.71
PU5	0.10	-0.16	0.19	0.08	0.10	0.35	0.25	0.66	**0.75**	0.58
PU6	0.15	-0.03	0.14	0.03	0.18	0.25	0.38	0.62	**0.80**	0.65

Table 30 shows that all of the five items of the Intention to Use construct had factor loadings of 0.70 or greater, indicating that they are satisfactory for this

study. None of the items were removed from the construct. There was no evidence of cross loading for this construct - all items loaded the strongest to their own construct.

Table 30 : Factor Structure Matrix of Loadings and Cross-loadings for Intention to Use

	VGSE	VGAN	VGAFF	SMSE	ECAN	ATLE	DLE	PEOU	PU	INT
INT1	0.08	-0.16	0.34	0.01	0.29	0.33	0.31	0.54	0.68	**0.87**
INT2	0.10	-0.10	0.17	0.03	0.28	0.26	0.28	0.60	0.69	**0.85**
INT3	0.12	-0.06	0.34	0.16	0.14	0.35	0.36	0.62	0.68	**0.85**
INT4	0.12	-0.17	0.21	0.03	0.17	0.29	0.36	0.47	0.63	**0.75**
INT5	0.04	-0.08	0.28	0.00	0.15	0.27	0.34	0.57	0.63	**0.76**

Therefore, for TAM-related constructs—Perceived Usefulness, Perceived Ease of Use and Intention to Use—all items exceeded the threshold set by (Hair et al., 1998; Hair, Black, Babin, Anderson, & Tatham, 2006). All of the constructs related to Subject Matter Self-efficacy and most of the constructs related to video-games self efficacy had some of the items removed. Although items used in this study were based on well established scales, scales performing well in certain circumstances often have items with poor loadings when used in a different context (Barclay et al., 1995).

When all items of measuring a construct are reliable, it is evidence in favour of convergent validity for the construct. Internal consistency reliability presents further evidence in this respect. Internal consistency reliability describes the extent to which all indicators of a construct change in a consistent manner. Internal consistency reliability can be assessed by using composite reliability calculated by the PLS algorithm (Werts, Linn, & Joreskog, 1974). Nunally (1978) and Fornell and Larcker (1981) suggested 0.7 as a cut-off point for internal consistency reliability, and this value is commonly used in the literature (Gefen & Straub, 2005).

Table 31 lists composite reliability values for all constructs (calculated after the unreliable items were removed). As seen from the table, all constructs were highly reliable.

Table 31: Internal Consistency Reliability Indices

Construct	Composite Reliability	Cronbach Alpha	AVE
Attitudes Towards Learning English	0.91	0.86	0.71
Desire to Learn English	0.85	0.78	0.54
English Class Anxiety	0.89	0.85	0.62
Intention to Use	0.91	0.88	0.68
Perceived Ease Of Use	0.89	0.86	0.59
Perceived Usefulness	0.93	0.91	0.7
Subject Matter Self-Efficacy	0.9	0.87	0.61
Video Games Affect	0.91	0.87	0.72
Video Games Anxiety	0.82	0.7	0.6
Video Games Self Efficacy	0.86	0.78	0.6

Cronbach's alpha is an alternative measure of internal consistency reliability that does not rely on PLS algorithm, but is also reported by SmartPLS software. It is common to use the same cut-off point of 0.7 for Cronbach's alpha as for composite reliability (Gefen & Straub, 2005), even though, as can be seen from Table 31, Cronbach's alpha tends to have lower values than composite reliability for the same data. All constructs were reliable according to Cronbach's alpha criterion. Video Games Anxiety was somewhat less reliable than the rest.

The average variance extracted (AVE) describes the amount of variance in the indicators described by the latent construct. AVE can be viewed as another measure of internal consistency reliability (Fornell & Larcker, 1981). A common cut-off values for AVE is 0.50 (Hair et al., 1998). As shown in Table 31, all latent variables complied with this criterion.

Item reliability and internal consistency reliability (assessed via composite reliability, Cronbach's alpha, and AVE) present strong evidence of convergent validity for the measures used in this study (after poorly loading indicators were removed).

Discriminant validity reflects the extent to which the indicators for each construct are distinctly different from indicators of other constructs. As demonstrated in the beginning of this section, all indicators loaded higher on their own constructs then

on other constructs in the model. This constitutes evidence in favour of internal consistency reliability.

At the construct level, discriminant validity is considered adequate when the variance shared between a construct and any other construct in the model is less than the variance that the constructs share with their measures (Fornell, Tellis, & Zinkhan, 1982). The variance shared by two constructs is obtained by calculating the square root of the correlation between the two constructs. The variance shared between a construct and its measures corresponds to AVE. Discriminant validity was assessed by comparing the square root of the AVE for a given construct with the correlations between that construct and all other constructs.

Table 32 shows the correlation matrix for the constructs. The diagonal elements were replaced by the square roots of the AVE. For discriminant validity to be judged as adequate, these diagonal elements should be greater than the off-diagonal elements in the corresponding rows and columns (Fornell & Larcker, 1981). Discriminant validity was satisfactory at the construct level for all constructs. This indicates that each construct shared more variance with its items than it did with other constructs. Having achieved discriminant validity at both the item and construct level, the construct measures in the proposed research model were deemed to be adequate.

Discriminant validity was also confirmed by examining correlations among the constructs. According to Gefen & Straub (2005), a 0.85 correlation or larger indicates poor discriminant validity in structural equation modelling. None of the correlations presented in Table 32 are above 0.85.

On the whole, acceptable measurement quality was achieved, as demonstrated by the convergent and discriminant validity tests. All construct measures in the proposed model satisfied the criteria of reliability and discriminant validity; therefore, there was no need to remove or to merge constructs in the research model. Achieving the reliability and validity criteria at this stage is a prerequisite for further analysis (DeVellis, 1991).

Table 32: Correlation of Latent Variables & Square Roots of Average Variance Extracted (AVE)

	ATLE	DLE	ECAN	INT	PEOU	PU	SMSE	VGAF	VGAN	VGSE
ATLE	**0.84**									
DLE	0.55	**0.73**								
ECAN	-0.14	0.15	**0.79**							
INT	0.33	0.38	0.23	**0.82**						
PEOU	0.4	0.39	0.1	0.68	**0.77**					
PU	0.41	0.44	0.13	0.8	0.76	**0.83**				
SMSE	0.41	0.19	-0.52	0.01	0.14	0.09	**0.78**			
VGAF	0.06	0.08	0.09	0.33	0.26	0.21	0	**0.84**		
VGAN	0.03	0.02	-0.04	-0.12	-0.1	-0.09	0.04	-0.19	**0.77**	
VGSE	0.02	0.16	0.18	0.07	0.07	0.07	-0.08	-0.01	0.37	**0.77**

5.3.1.2 Test of the Structural Model for MUVE Acceptance—Testing Hypotheses A1 to A14

Following confirmation of good psychometric properties in the measurement model, the structural model was tested. Since PLS does not generate an overall good-of-fit index, the primary assessment of validity was by examining R^2 and the structural paths (Chwelos, Benbasat, & Dexter, 2001). Bootstrapping re-sampling with 500 samples was performed to generate t-statistics to assess the significance of the model paths, as recommended by (Chin & Frye, 1998).

Figure 18 shows values of path coefficients and their *p* values obtained via bootstrapping. It also shows R^2 values for all dependent constructs, which indicates fraction of variance in these constructs predicted by the model.

The model explained 69% of the variance in Intention to Use, 60% of the variance in Perceived Usefulness, 13% of the variance in Video Games Anxiety, 27% of the variance in English Class Anxiety, and 16% of the variance in Attitude Towards Learning English.

Table 33 shows the results for hypotheses A1 to A14. Based on the analysis, constructs Perceived Usefulness, Video Games Affect, and English Class Anxiety affected the Intention to Use MUVE e-Learning. In addition, Subject Matter Self Efficacy affected English Class Anxiety and Attitude Towards Learning English, and Video Games Self-efficacy affected Video Games Anxiety. Perceived Ease of

Use affected Perceived Usefulness, but did not affect the Intention to Use directly. The meaning of these relationships and their effect sizes (as measured by the values of β) are discussed in section 6.3.2.

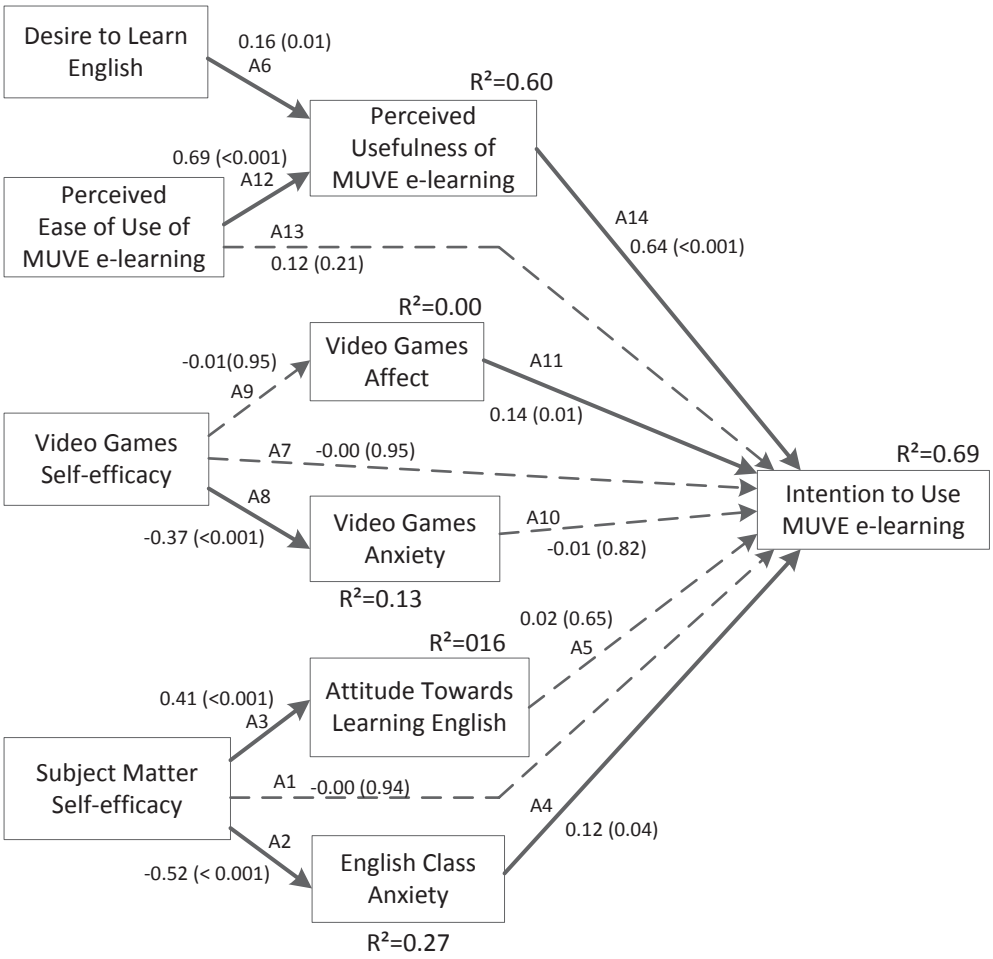

Figure 18: PLS results for the MUVE e-Learning acceptance model.

Table 33: Results of Hypothesis Testing

	Hypothesis	β	*p*	Outcome
A1	Subject Matter Self-efficacy affects Intention to Use	0.00	0.94	Not supported
A2	Subject Matter Self-efficacy affects English Class Anxiety	-0.52	< 0.001	Supported
A3	Subject Matter Self-efficacy affects Attitude Towards Learning English	0.41	< 0.001	Supported
A4	English Class Anxiety affects Intention to Use	0.12	0.04	Supported
A5	Attitude Towards Learning English affects Intention to Use	0.02	0.65	Not supported
A6	Desire to Learn English affects Perceived Usefulness	0.16	0.01	Supported
A7	Video Games Self-efficacy affects Intention to Use	0.00	0.95	Not Supported
A8	Video Games Self-efficacy affects Video Games Anxiety	-0.37	< 0.001	Supported
A9	Video Games Self-efficacy affects Video Games Affect	-0.01	0.95	Not supported
A10	Video Games Anxiety affects Intention to Use	-0.01	0.82	Not supported
A11	Video Games Affect affects Intention to Use	0.14	0.01	Supported

	Hypothesis	β	*p*	Outcome
A12	Perceived Ease of Use affects Perceived Usefulness	0.69	< 0.001	Supported
A13	Perceived Ease of Use affects Intention to Use	0.12	0.21	Not supported
A14	Perceived Usefulness affects Intention to Use	0.64	<0.001	Supported

Note. A relationship was deemed to be supported when the corresponding path was statistically significant at the level of 0.05.

5.3.2 Confirming Convergent and Discriminant Validity by Using Exploratory Factor Analysis

In an effort to strengthen the results for the convergent and discriminant validity analysed earlier by PLS (section 5.3.1.1), convergent and discriminant validity were also assessed by performing exploratory factor analysis using M-Plus software. This analysis identified the underlying factors or the dimensional composition of the indicators used to measure constructs of the MUVE e-Learning acceptance model. The aim of exploratory factor analysis was to reveal the latent variables that cause the indicators to co-vary (Costello & Osborne, 2005). It was expected that the latent variables thus discovered would match the latent constructs of the MUVE e-Learning acceptance model with sufficiently high EFA loadings, thus confirming the convergent and the discriminant validity of the constructs' measures.

Before running the sample data through factor analysis, the size of the sample needed to be considered. The subject to item ratio for the items of the MUVE e-Learning acceptance model was 2.01, and thus sufficiently high for EFA (Costello & Osborne, 2005). EFA analysis with a maximum likelihood (ML) estimator (using MPlus software) was used. ML estimator assumes multivariate normality in the indicators, although it is a common practice to use it with data that is slightly non-normal. To confirm that the data is close enough to normality, the criteria suggested by Kline (2005) were used. The criteria were met. In particular, skewness for all indicators was less than 3 and kurtosis was less than 10. The data met these criteria (for most indicators, both skewness and kurtosis were considerably less than 1). In the EFA analyses, the number of factors was set to 10, which is the number of constructs in the MUVE e-Learning acceptance model (Figure 16). The expectation was that all indicators of a construct would load the highest on the same factor (which, thus, would correspond to the construct).

For an item to be considered as associated with a given factor, it should load on it greater than any other factors, and the factor loading should be 0.4 or greater (Gefen & Straub, 2005). (However, it is noted that for PLS analysis the cut-off value for factor loading suggested in the same review article was 0.7—a more stringent requirement.)

The rest of this section presents a series of tables with EFA loadings for sets of indicators corresponding to different constructs in the MUVE e-Learning acceptance model. In these tables, the highest factor loading for each indicator is given in bold. Indicators that were dropped in PLS measurement model analysis are listed using a strike-through font. Each table is accompanied by a brief discussion of the extent of agreement between EFA results and PLS measurement model analysis results.

Table 34 shows that for Video Games Self-efficacy, the items retained in the PLS measurement model analysis all loaded the highest on the same factor. Moreover, the two items that loaded the highest on a different factor in EFA had loadings on that factor below 0.5. Overall, items that were removed in the PLS measurement model analysis also had relatively low loadings in EFA. Thus, EFA results for Video Games Self-efficacy were consistent with PLS results.

Table 34: Factor Loadings for Video Games Self-efficacy Indicators

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
~~VGSE1~~	~~0.004~~	~~0.11~~	~~0.09~~	~~0.01~~	~~0.19~~	~~0.05~~	~~0.10~~	~~**0.26**~~	~~0.10~~	~~0.09~~
~~VGSE2~~	~~0.004~~	~~0.03~~	~~0.02~~	~~0.05~~	~~0.27~~	~~0.04~~	~~0.04~~	~~**0.38**~~	~~0.20~~	~~0.09~~
~~VGSE3~~	~~0.10~~	~~0.10~~	~~0.02~~	~~0.02~~	~~0.08~~	~~0.09~~	~~**0.36**~~	~~0.16~~	~~0.01~~	~~0.14~~
~~VGSE4~~	~~0.02~~	~~0.20~~	~~0.02~~	~~0.02~~	~~0.09~~	~~0.02~~	~~**0.56**~~	~~0.03~~	~~0.08~~	~~0.03~~
VGSE5	0.02	0.23	0.09	0.05	0.05	0.01	**0.60**	0.05	0.05	0.00
VGSE6	0.02	0.01	0.01	0.20	0.03	0.10	**0.77**	0.04	0.10	0.12
~~VGSE7~~	~~0.03~~	~~0.03~~	~~0.03~~	~~0.06~~	~~0.04~~	~~0.16~~	~~**0.38**~~	~~0.14~~	~~0.08~~	~~0.18~~
~~VGSE8~~	~~0.11~~	~~0.05~~	~~0.24~~	~~0.003~~	~~0.02~~	~~0.05~~	~~**0.50**~~	~~0.00~~	~~0.14~~	~~0.03~~
VGSE9	0.13	0.01	0.04	0.32	0.04	0.03	**0.79**	0.03	0.04	0.02
~~VGSE10~~	~~0.02~~	~~0.11~~	~~0.14~~	~~0.13~~	~~0.07~~	~~0.05~~	~~**0.37**~~	~~0.06~~	~~0.11~~	~~0.01~~

Table 35 shows that for Video Games Anxiety, of the three items that were retained in the PLS measurement model analysis, two loaded the highest on the same factor (factor 9). The item that loaded the highest on a different factor

(VGAN1) still had its second highest loading on factor 9. Thus, overall, EFA results were consistent with PLS results.

Table 35: Factor Loadings for Video Games Anxiety Indicators

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
VGAN1	0.01	0.08	0.10	0.13	0.04	0.06	**0.43**	0.04	0.27	0.06
VGAN2	0.005	0.03	0.10	0.01	0.07	0.01	0.26	0.14	**0.52**	0.02
VGAN3	0.01	0.006	0.002	0.02	0.001	0.07	0.27	0.10	**0.51**	0.002
~~VGAN4~~	~~0.04~~	~~0.02~~	~~0.03~~	~~0.08~~	~~0.02~~	~~0.02~~	~~0.06~~	~~0.34~~	~~**0.61**~~	~~0.04~~

Table 36 shows that for Video Games Affect, all items loaded the highest on the same factor. For this construct, no items were removed in the PLS measurement model analysis. Thus, EFA results for Video Games Affect were consistent with PLS results.

Table 36: Factor Loadings for Video Games Affect Indicators

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
VGAff1	0.02	0.03	0.03	0.16	0.05	0.003	0.004	**0.85**	0.03	0.02
VGAff2	0.06	0.05	0.05	0.01	0.003	0.01	0.06	**0.86**	0.12	0.04
VGAff3	0.02	0.01	0.08	0.01	0.001	0.03	0.003	**0.85**	0.02	0.05
VGAff4	0.14	0.01	0.08	0.12	0.20	0.01	0.09	**0.57**	0.05	0.05
VGAff5	0.11	0.02	0.14	0.06	0.08	0.16	0.06	**0.61**	0.11	0.03

Table 37 shows that for Subject Matter Self-efficacy, three of the six items retained in the PLS measurement model analysis loaded on one factor (factor 1), and the remaining three items - on another factor (factor 4). Based on the magnitude of loadings, it was concluded that Subject Matter Self-efficacy corresponded to factor 1. Consistently with that view, all of the three items that loaded the highest on factor 4 had their second highest loadings on factor 1. Thus, EFA results for Subject Matter Self-efficacy were overall consistent with PLS results. The crossloadings to factor 4 are discussed further in connection with Table 38 listing factor loadings for English Class Anxiety.

Table 37: Factor loadings for Subject Matter Self-efficacy Indicators

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
SMSE1	0.36	0.04	0.09	**0.51**	0.08	0.03	0.05	0.01	0.04	0.05
~~SMSE2~~	~~0.27~~	~~0.14~~	~~0.06~~	~~0.25~~	~~**0.36**~~	~~0.02~~	~~0.06~~	~~0.20~~	~~0.10~~	~~0.04~~
SMSE3	0.40	0.04	0.11	**0.54**	0.07	0.02	0.01	0.06	0.01	0.04
~~SMSE4~~	~~0.01~~	~~0.07~~	~~0.05~~	~~**0.33**~~	~~0.13~~	~~0.12~~	~~0.05~~	~~0.19~~	~~0.14~~	~~0.07~~
SMSE5	0.46	0.01	0.13	**0.52**	0.01	0.05	0.01	0.01	0.09	0.10
~~SMSE6~~	~~**0.30**~~	~~0.02~~	~~0.21~~	~~0.28~~	~~0.01~~	~~0.02~~	~~0.11~~	~~0.03~~	~~0.13~~	~~0.05~~
~~SMSE7~~	~~**0.42**~~	~~0.04~~	~~0.02~~	~~0.32~~	~~0.01~~	~~0.15~~	~~0.16~~	~~0.05~~	~~0.03~~	~~0.04~~
SMSE8	**0.64**	0.00	0.10	0.41	0.16	0.03	0.34	0.02	0.18	0.01
SMSE9	**0.76**	0.01	0.01	0.29	0.02	0.08	0.05	0.03	0.12	0.11
SMSE10	**0.68**	0.09	0.02	0.31	0.04	0.02	0.01	0.01	0.04	0.05

Table 38 shows that for English Class Anxiety all of the six items that were retained in the PLS measurement model loaded on the same factor and had higher loadings than indicators that were removed in PLS analysis. Thus, EFA results for English Class Anxiety were consistent with PLS results.

Table 38: Factor Loadings for English Class Anxiety Indicators

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
ECAN1	0.005	0.03	0.03	**0.69**	0.19	0.00	0.03	0.003	0.11	0.06
ECAN2	0.14	0.15	0.04	**0.71**	0.02	0.05	0.01	0.02	0.02	0.05
ECAN3	0.01	0.02	0.003	**0.65**	0.23	0.03	0.05	0.03	0.13	0.09
ECAN4	0.05	0.06	0.17	**0.66**	0.11	0.11	0.17	0.03	0.09	0.08
ECAN5	0.01	0.06	0.08	**0.81**	0.10	0.03	0.12	0.06	0.01	0.09
~~ECAN 6~~	~~0.18~~	~~0.20~~	~~0.07~~	~~0.16~~	~~**0.28**~~	~~0.05~~	~~0.02~~	~~0.04~~	~~0.02~~	~~0.03~~
ECAN7	0.14	0.05	0.12	**0.69**	0.19	0.07	0.01	0.01	0.07	0.15
~~ECAN8~~	~~0.02~~	~~0.01~~	~~0.15~~	~~**0.39**~~	~~0.34~~	~~0.01~~	~~0.01~~	~~0.09~~	~~0.05~~	~~0.08~~
~~ECAN9~~	~~0.02~~	~~0.05~~	~~0.07~~	~~**0.54**~~	~~0.04~~	~~0.17~~	~~0.15~~	~~0.02~~	~~0.09~~	~~0.01~~
~~ECAN10~~	~~0.01~~	~~0.18~~	~~0.12~~	~~**0.33**~~	~~0.03~~	~~0.18~~	~~0.15~~	~~0.02~~	~~0.08~~	~~0.08~~

As noted earlier, some of the items for Subject Matter Self-efficacy also loaded the highest on factor 4. This is not surprising, because the two constructs—English Class Anxiety and Subject Matter Self-efficacy are related (and are directly connected in the MUVE e-Learning acceptance model in Figure 16). Thus, EFA analysis suggests that the discriminant validity between these two constructs is not strong. Indeed, in PLS analysis, strong negative correlation was found between these two constructs (see Figure 18), which, however, was not strong enough to fail the criteria for discriminant validity used in PLS analysis.

Table 39 shows that for Attitude Towards Learning English all of the items retained in the PLS measurement model analysis loaded the highest on the same factor, and their factor loadings on this factor were higher than for indicators that were not retained in PLS analysis. Thus, for Attitude Towards Learning English EFA results were consistent with PLS results.

Table 39: Factor Loadings for Attitude Towards Learning English Indicators

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
ATLE1	0.22	**0.53**	0.11	0.04	0.11	0.05	0.07	0.01	0.03	0.08
ATLE2	0.05	**0.83**	0.05	0.09	0.05	0.07	0.05	0.10	0.22	0.04
ATLE3	0.13	**0.84**	0.01	0.09	0.05	0.05	0.01	0.01	0.13	0.07
ATLE4	0.03	**0.70**	0.04	0.04	0.31	0.02	0.01	0.00	0.02	0.01
~~ATLE5~~	~~0.02~~	~~0.23~~	~~0.10~~	~~0.09~~	~~0.29~~	~~**0.32**~~	~~0.00~~	~~0.09~~	~~0.05~~	~~0.04~~
~~ATLE6~~	~~0.20~~	~~**0.39**~~	~~0.02~~	~~0.32~~	~~0.05~~	~~0.38~~	~~0.12~~	~~0.03~~	~~0.00~~	~~0.03~~
ATLE7	0.13	**0.50**	0.03	0.13	0.04	0.39	0.02	0.10	0.10	0.01
~~ATLE8~~	~~0.07~~	~~0.27~~	~~0.09~~	~~0.02~~	~~0.06~~	~~**0.56**~~	~~0.12~~	~~0.03~~	~~0.12~~	~~0.06~~
ATLE9	0.10	**0.49**	0.05	0.06	0.16	0.44	0.02	0.07	0.08	0.04

Table 40 shows that for Desire to Learn English, the indicators that were retained in the PLS measurement model analysis loaded the highest on the same factor (factor 5). However, the magnitudes of the factor loadings were relatively small (and below the cut-off value of 0.5).

Indeed, from EFA results only, one would conclude that Desire to Learn English corresponds to factor 6. Thus, EFA results were not entirely in agreement with PLS results, although there was a degree of consistency.

Table 40: Factor Loadings for Desire to Learn English Indicators

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
~~DLE 1~~	~~0.02~~	~~0.12~~	~~0.13~~	~~0.25~~	~~**0.41**~~	~~0.00~~	~~0.19~~	~~0.04~~	~~0.00~~	~~0.07~~
~~DLE 2~~	~~0.13~~	~~0.35~~	~~0.04~~	~~0.02~~	~~**0.38**~~	~~0.00~~	~~0.12~~	~~0.15~~	~~0.08~~	~~0.03~~
DLE 3	0.20	0.27	0.09	0.11	**0.38**	0.10	0.05	0.05	0.15	0.01
DLE 4	0.10	0.40	0.03	0.04	**0.47**	0.13	0.02	0.02	0.02	0.06
DLE 5	0.11	0.24	0.10	0.17	**0.35**	0.20	0.08	0.11	0.05	0.00
~~DLE6~~	~~0.15~~	~~0.12~~	~~0.01~~	~~0.15~~	~~0.16~~	~~**0.49**~~	~~0.07~~	~~0.02~~	~~0.01~~	~~0.07~~
~~DLE7~~	~~0.16~~	~~0.05~~	~~0.13~~	~~0.05~~	~~0.18~~	~~**0.52**~~	~~0.01~~	~~0.01~~	~~0.09~~	~~0.04~~
~~DLE8~~	~~0.12~~	~~0.08~~	~~0.02~~	~~0.03~~	~~0.23~~	~~**0.60**~~	~~0.00~~	~~0.10~~	~~0.19~~	~~0.13~~
~~DLE9~~	~~0.06~~	~~0.04~~	~~0.16~~	~~0.12~~	~~0.02~~	~~**0.53**~~	~~0.04~~	~~0.02~~	~~0.03~~	~~0.00~~

Table 41 shows that all indicators corresponding to TAM constructs loaded the highest on the same factor. This outcome was not entirely inconsistent with the PLS measurement model results, as these constructs correlated highly in TAM analysis (even though, they fulfilled the formal requirements for discriminant validity).

Table 41: Factor loadings for TAM indicators (Perceived Ease of Use, Perceived Usefulness, and Intention to Use)

	Factor									
Item	1	2	3	4	5	6	7	8	9	10
Perceived Ease of Use										
PEOU1	0.01	0.05	**0.58**	0.15	0.11	0.03	0.08	0.05	0.04	0.28
PEOU2	0.11	0.08	0.32	0.10	0.02	0.03	0.11	0.08	0.06	**0.34**
PEOU3	0.02	0.01	**0.63**	0.05	0.07	0.04	0.03	0.03	0.02	0.61
PEOU4	0.04	0.09	**0.60**	0.05	0.04	0.04	0.05	0.001	0.03	0.47
PEOU5	0.18	0.02	**0.50**	0.03	0.11	0.05	0.04	0.02	0.05	0.27
PEOU6	0.19	0.08	**0.48**	0.01	0.14	0.03	0.10	0.01	0.09	0.22
Perceived Usefulness										
PU1	0.09	0.09	**0.79**	0.06	0.01	0.05	0.18	0.01	0.09	0.04
PU2	0.04	0.13	**0.82**	0.04	0.14	0.06	0.06	0.05	0.06	0.12
PU3	0.05	0.02	**0.83**	0.08	0.06	0.09	0.05	0.06	0.03	0.03
PU4	0.10	0.05	**0.77**	0.01	0.07	0.11	0.13	0.03	0.04	0.01
PU5	0.01	0.10	**0.65**	0.03	0.09	0.04	0.05	0.02	0.03	0.24
PU6	0.03	0.13	**0.74**	0.03	0.14	0.08	0.05	0.02	0.09	0.11
Intention to Use										
INT1	0.07	0.06	**0.66**	0.21	0.05	0.01	0.002	0.16	0.01	0.06
INT2	0.14	0.02	**0.74**	0.22	0.15	0.003	0.03	0.01	0.11	0.02
INT3	0.03	0.04	**0.71**	0.02	0.19	0.003	0.03	0.20	0.15	0.03
INT4	0.09	0.08	**0.70**	0.01	0.19	0.02	0.08	0.06	0.04	0.14
INT5	0.18	0.01	**0.68**	0.03	0.26	0.02	0.05	0.11	0.08	0.12

A possible reason for a lack of discrimination between the three TAM constructs is the fact that there was a relatively low subject to item ratio. While the ratio of two can be seen as acceptable according to Costello and Osborne (2005), Fabrigar, Wegener, MacCallum and Strahan (1999) suggest a higher subject to item ratio for robust analysis (five or even ten). To explore this possibility, EFA analysis was conducted with TAM constructs' indicators only (so that subject to item ratio was close to nine). The results are shown in Table 42. As seen from the table, EFA at a higher subject to item ratio resulted in better discrimination between the TAM constructs, although cross-loading between Perceived Usefulness and Intention to Use was still considerable. This outcome may indicate a common method variance

issue between these two constructs, as suggested by Sharma, Yetton and Crawford (2009).

Table 42: Exploratory Factor Analysis for TAM Constructs' Items Only

	Factor		
Item	1	2	3
Perceived Ease of Use			
PEOU1	0.19	**0.48**	0.02
PEOU2	0.09	**0.51**	0.06
PEOU3	0.04	**0.91**	0.10
PEOU4	0.02	**0.79**	0.17
PEOU5	**0.42**	0.41	0.17
PEOU6	**0.39**	0.34	0.05
Perceived Usefulness			
PU1	**0.91**	0.01	0.06
PU2	**0.89**	0.08	0.05
PU3	**0.74**	0.06	0.10
PU4	**0.89**	0.01	0.11
PU5	**0.42**	0.37	0.01
PU6	**0.51**	0.22	0.14
Intention to Use			
INT1	**0.69**	0.002	0.11
INT2	**0.68**	0.08	0.01
INT3	0.35	0.24	**0.46**
INT4	0.42	0.01	**0.57**
INT5	0.001	0.44	**0.71**

5.3.3 Confirming PLS Results by Using Covariance-based SEM

5.3.3.1 Formulating a Reduced Model and Checking the Prerequisites for Using Covariance Based SEM

Covariance based SEM was used to confirm the results of PLS. Covariance based SEM requires high subject to item ratios and, therefore, fitting large models, such as the MUVE e-Learning acceptance model in Figure 16, requires unrealistically large data sets. To apply covariance based SEM, The original model was reduced to retain only the constructs that were found to be directly affecting the Intention to Use MUVE e-Learning (and also included the Desire to Learn English, as a non-TAM construct that affected Perceived Usefulness). The reduced model is given in Figure 19.

The analysis started with the measurement model used in the PLS analysis; a number of items in construct measures were dropped based on combined PLS and

EFA results. For Intention to Use, following items were retained: INT3, INT4, and INT5, the items that did not have discriminant validity problems in EFA analysis (see Table 42). For Perceived Usefulness, the following items were retained: PU1, PU2, PU3, and PU4; these items loaded considerably better than the rest in the EFA analysis (and somewhat better than the rest in the PLS analysis) (see Table 42 and Table 29). Similarly, for Video Games Affect, VGAff1, VGAff2 and VGAff3, which loaded considerably better than the rest in the EFA analysis (and somewhat better than the rest in the PLS analysis), were retained (see

Table 36 and Table 23). For English Class Anxiety, initially all items were retained, but it was found in initial CFA analysis with AMOS that ECAN1 and ECAN7 loaded in the AMOS CFA model considerably below 0.7. Therefore, these two items were dropped, so that only ECAN2, ECAN3, ECAN4, and ECAN5 were retained. For all constructs in the reduced model, the items retained were examined to ensure that the resulting measures had face validity (the wordings of all items are given in a series of tables in section 4.7.2). There were no significant shifts in the content of the measures. (It would be of value to confirm this view by conducting content analysis of the updated measures with independent experts, but this was not feasible.)

The reduced model had a roughly 4:1 subject to parameter ratio (estimated by using AMOS software). Although this was somewhat less than the recommended minimal 5:1 subject to parameter ratio (Kline, 2005), a view was adopted that, even for a smaller than ideal data set, if the results of covariance based analysis are consistent with PLS and EFA results, it constitutes additional evidence that the overall conclusions are correct.

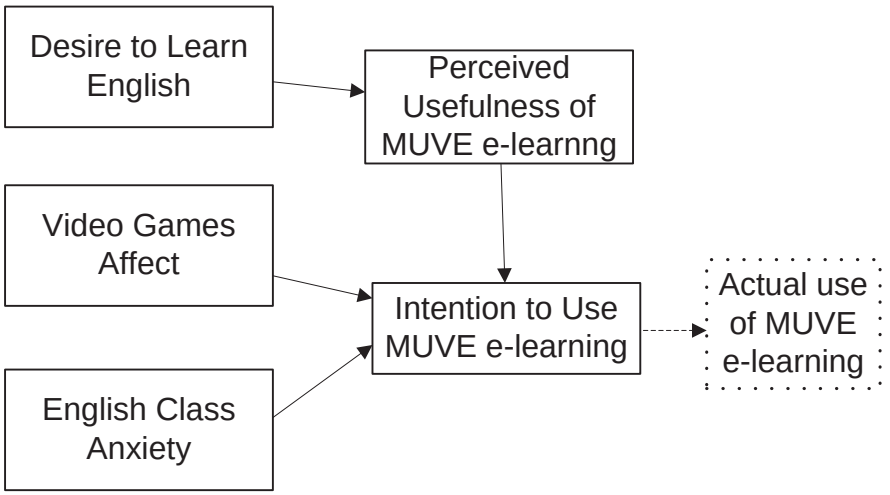

Figure 19. Reduced MUVE e-Learning acceptance model tested using covariance-based SEM, introduced in section 5.3.3.1.

Besides fulfilling the sample size requirement, the validity of maximum likelihood estimation used by AMOS software depends on whether the data meets the assumption of multivariate normality of the indicators. As discussed in section 5.3.2 in the context of fulfilling the conditions for EFA with a maximum likelihood estimator, to confirm that the data is close enough to normality, following the criteria suggested by Kline (2005) the researcher verified that skewness for all indicators was less than 3 and that kurtosis was less than 10. The data met these criteria. In addition, for indicators to be included in covariance analysis with AMOS, the Mardia's test of multivariate normality was conducted (Mardia, 1970; D' Agostino, 1986; Bollen, 1989). According to the Mardia's test, the data was not multivariate normal ($p < 0.001$ for both multivariate skewness and for multivariate kurtosis, which assumed the values of 64.37 and 368.88, respectively).

It is common to use statistical methods assuming normality for data that does not follow the normal distribution very closely, and in most cases, the results are satisfactory in terms of correctly capturing the substantive relationships (Micceri, 1989; Chou, Bentler, & Satorra, 1991). In particular, it is a common practice to use

covariance based SEM with data that fail multivariate normality tests, and as the data did fulfil the less stringent criteria for being close enough to multivariate normal suggested by Kline (2005), covariance based SEM analysis was executed.

5.3.3.2 Testing the Measurement Model

The confirmatory factor model in AMOS was defined as in Figure 20. The two-sided arrows in the figure represent Pearson's correlations between constructs that were evaluated in the model. A CFA model does not make any assumptions about the relationships between constructs so that its fit reflects the quality of the measurement
model.

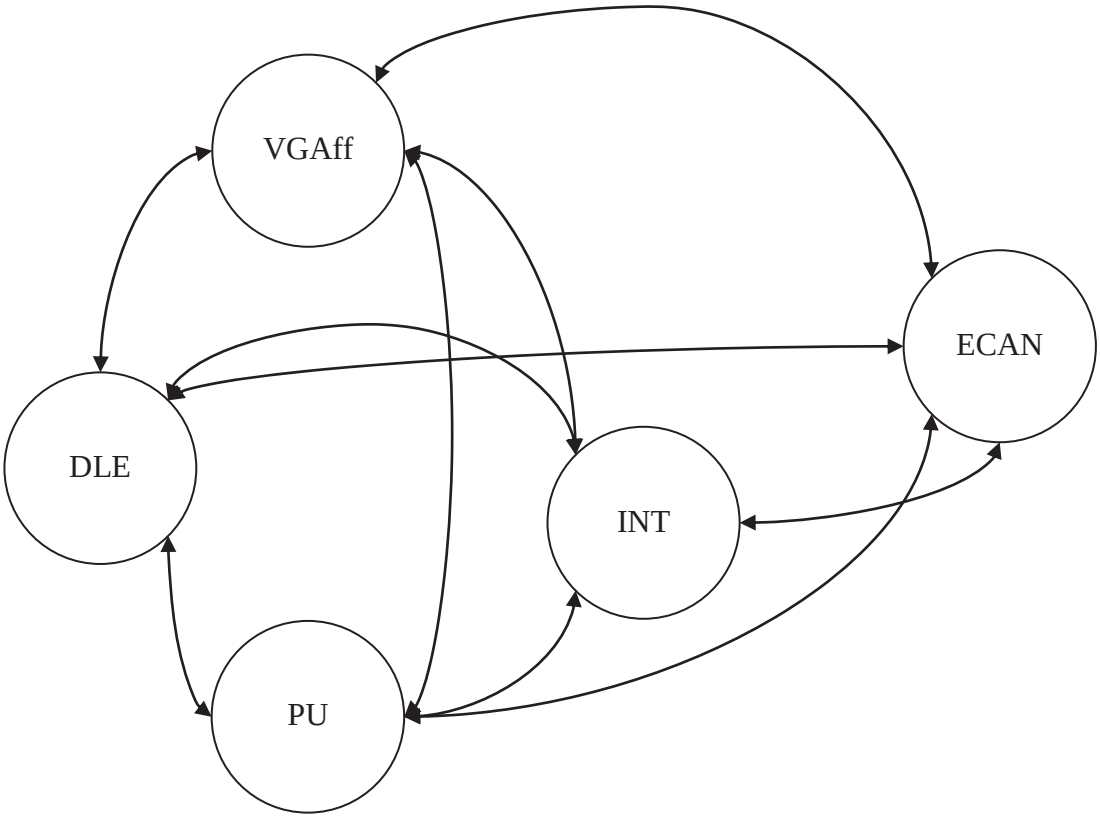

Figure 20: Specification of the CFA model for the reduced MUVE e-Learning acceptance model (introduced at the beginning of section 5.3.4).

Unlike PLS SEM, for which there are no recognised model fit indices (measures of the model fit as a whole), covariance based SEM has a variety of indices for measuring model fit (Kline, 2005). The values obtained for the CFA model are listed in Table 43 along with the benchmarks suggested by Gefen and Straub

(2005) for use in MIS research. As seen from the Table 43, the fit of the measurement model was acceptable.

Table 43: Global Fit Indices for The CFA Model (see Figure 20)

Fit index	Benchmark	Value
Chi-square / degrees of freedom	< 2	1.252
GFI (Goodness of fit Index)	> 0.9	0.907
AGFI (Adjusted goodness of fit Index)	> 0.80	0.869
NFI (Normed fit index)	> 0.90	0.912
CFI (Comparative fit index)	> 0.90	0.981
RMSEA (Root mean error of approximation)	< 0.08	0.041

Table 44: Factor Loadings from Confirmatory Factor Analysis with AMOS

Item	**Factor loading**
Video Games Affect	
VGAff1	0.88
VGAff2	0.85
VGAff3	0.83
English Class Anxiety	
ECAN2	0.69
ECAN3	0.73
ECAN4	0.77
ECAN5	0.76
Desire to Learn English	
DLE3	0.69
DLE4	0.78
DLE5	0.68
Perceived Usefulness	
PU1	0.86
PU2	0.89
PU3	0.85
PU4	0.81
Intention to Use	
INT3	0.85
INT4	0.82
INT5	0.83

The factor loadings are listed in Table 44. After the adjustments to the measurement model discussed in section 5.3.3.1, the factor loadings were all higher than the benchmark value of 0.7 suggested by Gefen and Straub (2005),

with the exception of ECAN2, DLE3 and DLE5, which at the factor loadings of 0.69, 0.69, and 0.68, respectively, were judged to be close enough to the benchmark value.

The CFA estimates of correlations between constructs are given in Table 45. None of the correlations were above the threshold of 0.85 (Gefen & Straub, 2005), suggesting that the constructs were correlationally distinct, and, thus, supporting discriminant validity.

Table 45: Correlations Between Constructs from CFA Analysis

Constructs	Correlation
VGAff ⟷ ECAN	0.15
VGAFF ⟷ DLE	0.10
ECAN ⟷ DLE	0.09
VGAFF ⟷ INT	0.33
PU ⟷ DLE	0.51
PU ⟷ INT	0.78
INT ⟷ DLE	0.45
PU ⟷ VGAFF	0.19
PU ⟷ ECAN	0.12
ECAN ⟷ INT	0.20

5.3.3.3 Testing the Structural Model

The SEM model in AMOS was defined as in Figure 21. Arrows in the figure correspond to the hypothesised cause-effect relationships (and match the arrows in the model in Figure 19).

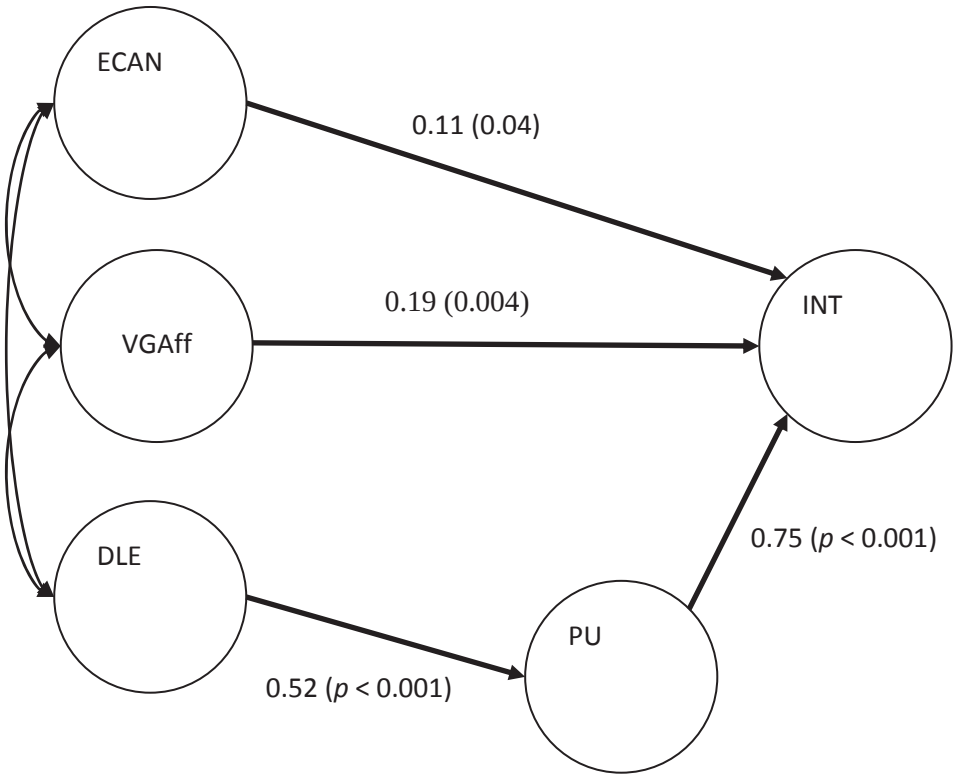

Figure 21. Specification of the SEM model for the reduced MUVE e-Learning acceptance model. Numbers show path coefficients obtained by fitting the data, with the respective p values given in parentheses.

The values obtained for the SEM model are listed in Table 46 along with the benchmarks suggested by Gefen and Straub (2005) for use in MIS research. As seen from the Table 46, the fit of the SEM model was acceptable. As common in covariance based modelling, the values of the fit indices for the SEM model were very close to the values in the corresponding CFA model.

Table 46: Global Fit Indices for The SEM Model for The Reduced MUVE e-Learning Acceptance Model (see Figure 19)

Fit index	Benchmark	Value
Chi-square / degrees of freedom	< 2	1.250
GFI (Goodness of fit Index)	> 0.9	0.905
AGFI (Adjusted goodness of fit Index)	> 0.80	0.870
NFI (Normed fit index)	> 0.90	0.910
CFI (Comparative fit index)	> 0.90	0.980
RMSEA (Root mean error of approximation)	< 0.08	0.041

The values of path coefficients and the corresponding *p* values are listed in Table 47. Each path coefficient corresponds to a hypothesis from section 3.3. Although only a small subset of the hypotheses are tested in the reduced model, in all cases, both the outcomes for the significance of paths, and the values of the coefficients are consistent with the PLS results (see Table 33).

Table 47: Path Coefficients in the Reduced MUVE e-Learning Acceptance Model

Path	Path coefficient	*p*	Hypothesis	Outcome
ECAN→INT	0.11	0.04	A4	supported
DLE→PU	0.52	$p < 0.001$	A6	supported
VGAFF→INT	0.19	0.004	A11	supported
PU→INT	0.75	$p < 0.01$	A14	supported

Note. A relationship was deemed to be supported when the corresponding path was statistically significant at a protection level of 0.05.

5.4 Summary

This chapter introduced the analysis of the data obtained in the experiment on MUVE e-Learning effectiveness. Then the chapter introduced the analysis of the data obtained in the cross-sectional survey of factors affecting MUVE acceptance by the learners. Reliability and validity issues for both analyses were addressed.

The results of survey data analysis using PLS were confirmed by using covariance-based techniques.

Chapter 6: Summary, Discussion and Conclusions

6.1 Introduction

This chapter starts with a summary of the study, which is followed by a discussion of the findings in view of the existing literature. This is followed by discussing the implications of the research for practice and of the limitations of the research, along with recommendations for further research. The chapter ends with stating the conclusions for the whole thesis.

6.2 Summary of the Study

The problem addressed in the present study was the lack of objective evidence of MUVE being effective as an environment supporting the teaching of English, as well as the lack of research on MUVE acceptance research conducted in context of using MUVE as an environment supporting the teaching of English as a second language. In particular, even though there are many experience reports, MUVE effectiveness was never demonstrated in a controlled experiment with a random assignment of participants to the experimental and to the control groups. Even though the study by Shen and Eder (2009) explored the factors affecting MUVE acceptance in the context of learning Management Information Systems, MUVE acceptance has never been studied in the context of learning English as a second language. Moreover, the present study was motivated by the need to improve business English proficiency by Malay learners, and none of the prior studies of MUVE effectiveness or acceptance were conducted with Malay participants.

Therefore, the research questions focused on MUVE effectiveness (ability to promote learning) and MUVE acceptance by the learners:

1. Is MUVE based learning effective in facilitating situated scenario-based learning of oral business English communication skills by Malay learners?
2. Which factors influence the acceptance of MUVE based learning of oral business English communication skills by Malay learners?

To address the research questions, a controlled experiment with random assignment of the participants and a survey were conducted. The experiment and the survey were combined to involve the same participants (the details of the research procedure are given Table 4).

The participants were university students at a university in Malaysia involved in a course devoted to professional communication in English, but instructional intervention used in the present study was not a part of the course. Rather, the research was conducted in a controlled environment, with volunteer participants.

The learning experience was designed based on communicative language learning, task based learning, and scenario based learning principles. The experience involved a scenario of selling a house. The participants, who learned in pairs, adopted in turn the roles of the buyer and of the seller. The researcher acted as an instructor and provided scaffolding and, when necessary, corrected mistakes and modelled correct language usage. The experience continued for 40 minutes, a relatively short duration that, however, was sufficient to improve the participants' performance in going through the scenario, because the scenario was very focused.

In the experiment, the participants were assigned at random in pairs to the experimental group and to the control group. Participants in the experimental group learned in MUVE using a highly realistic model of a house in the MUVE environment as a context for communication in English, discussing various aspects of the house relevant to the house's value. Participants in the control group learned face to face, using a plastic model of a house, which was less realistic than the house in MUVE.

Pre and post tests were conducted to assess learning performance, and involved audio recording the participants executing the scenario and rating the communication performance. The rating scheme focused explicitly on learners' use of English at performing the scenario roles, rather than on general language performance. Independent raters were used to rate the recordings; interrater reliability was at an acceptable level. Statistical analysis (*t*-tests) was used to determine if the improvement in communication performance differed between the groups. It was found that even though there was a statistically significant improvement between pre and post tests, there was no significant difference in performance gains between MUVE and face to face environments.

In the survey (which addressed the research question 2), the participants filled in a questionnaire, which was formulated based on measurement instruments validated in prior research. The questionnaire items measured the factors hypothesised to affect MUVE acceptance, the participants' Attitude Towards Learning English, Desire to Learn English, English Class Anxiety, Subject Matter (English) Self-efficacy, Video Games Affect, Video Games Anxiety, Video Games Self-efficacy, Perceived Ease of Use of MUVE for e-Learning, and Perceived Usefulness of MUVE as an environment for e-Learning, along with the main outcome variable, the Intention to Use MUVE for e-Learning in the future (the details of the research model, which was based on Technology Acceptance Model and Social Cognitive Theory, are given in Figure 16). Statistical analysis (using PLS SEM) was conducted to test the hypotheses suggested by the research model. According to the results of the SEM analysis, Perceived Usefulness, Video Games Affect, and English Class Anxiety affected the Intention to Use MUVE e-Learning. In addition, Subject Matter Self-efficacy affected English Class Anxiety, and Attitude Towards Learning English and Video Games Self-efficacy affected Video Games Anxiety. Perceived Ease of Use affected Perceived Usefulness, but did not affect the Intention to Use MUVE for e-Learning directly.

6.3 Discussion of Findings

Drawing on the information from data analysis, this section discusses the findings for the research questions put forward in Chapter 1, namely:

1. Is MUVE based learning effective in facilitating situated scenario-based learning of oral business English communication skills to Malay learners?
2. Which factors influence the acceptance of MUVE based learning of oral business English communication skills learning by Malay learners?

6.3.1 Is MUVE Based Learning Effective in Facilitating Situated Scenario-based Learning of Oral Business English Communication Skills by Malay Learners?

6.3.1.1 The Use of Scenario-based Approach in MUVE Is Effective to Teach Oral Business English Communication Skills.

The findings of the study provide evidence that MUVE is effective as a learning environment for scenario-based teaching of oral business English communication skills; the findings are based on a controlled experiment, therefore (unlike in studies relying on experience reports) the findings do not depend on the subjective judgement of the researcher.

The findings are consistent with the result by Cobb et al. (2009), who found, using a quasi-experiment, that MUVE is effective in promoting scenario based learning of lab procedures. The findings of the study also concur with the result by Lee, Wong, and Fung (2009), who also conducted a quasi-experiment comparing experiential learning of frog anatomy in a single user virtual environment and in a traditional classroom.

6.3.1.2 There Was No Significant Difference Between Learning Gains in MUVE and In The Classroom

It was found that the difference between learning gains in MUVE and in the classroom was not statistically significant, similar to the result by Wrzesien and Raya (2010), who compared the effectiveness of learning ecology and science in MUVE with face to face learning. The study involved relatively large numbers of participants and did detect learning gains in both environments at strong levels of statistical significance. The most likely reason for such a finding is that the difference in learning gains between MUVE and the classroom is small and, thus, difficult to detect experimentally. This outcome was consistent with the findings by Russell (1999), who reviewed a large number of studies comparing the effectiveness of using technology in teaching to more traditional approaches and concluded that most of the studies found no significant difference.

While Cobb et al. (2009) found scenario-based learning in MUVE to be more effective than classroom learning when learning gain is measured by the number of questions in a subsequent real lab session, the validity of their result is in doubt due to the approach they used to assign students to experimental groups (as discussed in section 2.11.1). It is more insightful to compare the result of the present study with Lee et al.'s (2009) result; Lee conducted a quasi-experiment to compare experiential learning of frog anatomy in a single user virtual environment and in a traditional classroom. With a very large sample (370 usable responses), Lee et al. (2009) found a significant difference, with the knowledge gain in the virtual environment being higher than in the classroom, but the effect size was rather small, so that the difference was, arguably, not practically significant. Similarly, if repeated with a very large sample, the experiment described in the present study is likely to find a difference between the learners' performance in MUVE and in the traditional classroom. However, as in the present study, with a middle-sized sample, there was no statistically significant difference; it is likely that the effect size found using a sample that is large enough to detect the difference would be small.

It should be noted that in an earlier study where learning in a single user virtual environment was compared with classroom learning (Crosier, Cobb, & Wilson, 2000), a statistically significant difference in favour of classroom learning was found, even though the sample was very small (29 participants). Crosier et al. (2000) explained the difference as being due to the poor design of the experience in the virtual environment; there were no scenarios or goals to pursue, and students were just exploring the environment with no clear direction. Crosier et al. (2000) reported introducing a number of changes based on their research, but did not publish any follow-up evaluation. Although in the present study the learning experience involved a well-defined scenario, with a design based on situated learning theory, there is a possibility that the learning experience could be improved, leading to a different outcome when the two environments were compared. The direction of an improvement is not clear, though, and further research is required to find if an improvement is possible.

Lim, Nonis and Hedberg (2006) and Wrzesien and Raya (2010) explained the less than expected learning gains in MUVE by students getting distracted by the

environment. In this study though, a pair of students worked under the supervision of the experimenter who kept them focused on the task, so that student distraction was not a factor.

6.3.2 Which Factors Influence the Acceptance of MUVE Based Learning of Oral Business English Communication Skills by Malay learners?

6.3.2.1 Factors Relating to the Technology Acceptance Model

In agreement with prior studies (Fetscherin & Lattemann, 2008; Saeed et al., 2008, 2009; Shen & Eder, 2009), the TAM model was confirmed to be valid in the context of the acceptance of MUVE based learning.

Perceived Usefulness. Perceived usefulness of MUVE as a learning environment was found to affect the Intention to Use MUVE for e-Learning. The effect size was strong, with β=0.64. This result was consistent with prior findings, which demonstrated that Perceived Usefulness plays an important role and affects the Intention to Use stronger than other constructs, such as Self-efficacy, Computer Playfulness, or Social Norms (Fetscherin & Lattemann, 2008; Saeed et al., 2008; Shen & Eder, 2009).

Perceived Ease of Use. The effect of Perceived Ease of Use on Intention to Use was found not to be significant. This is in agreement with Fetscherin and Lattemann (2008), Saeed et al. (2008), and Shen and Eder (2009).

Perceived Ease of Use did affect Perceived Usefulness, with β=0.69. This was in agreement with Fetscherin and Lattemann (2008), Saeed et al. (2008), Saeed et al. (2009), and Shen and Eder (2009).

6.3.2.2 Factors Relating to Social Cognitive Theory

Factors relating to social cognitive theory were of particular interest in view of the motivation behind pursuing the problem addressed by the present study (contributing to assessing MUVE as an environment suitable for Malay learners of English in Malaysia), because factors relating to social cognitive theory account for characteristics that are particularly relevant to Malay learners (such as shyness

and a negative attitude towards English, see section 1.1.3). However, overall, these factors were found to affect the Intention to Use relatively weakly, with some of the effects found not to be statistically significant and others relatively weak. (The two factors in this category that had statistically significant effects on Intention to Use MUVE for e-Learning were Video Games Affect and English Class Anxiety, with β=0.14 and β=0.12, respectively).

Video Games Related. Video Games Affect influenced the Intention to Use MUVE for e-Learning, but the effect size was small (β=0.14). The effect was smaller than the effect of affect on usage in the study by Compeau et al. (1999), who found β=0.25. The model by Compeau et al. (1999) did not include TAM variables, which might be the reason for a larger effect observed in their study.

The effect of Video Games Self-efficacy on Intention to Use was found to be not significant. At the same time, Video Games Self-efficacy negatively affected Video Games Anxiety, with a relatively strong effect size (β= -0.37). On the other hand, the effect of Video Games Self-efficacy on Video Games Affect was not found to be statistically significant. These results can be compared with the results of Compeau et al. (1999), who used a similar combination of constructs as independent variables in their model of computer usage. In the study of Compeau et al., the findings were as follows: self-efficacy had a significant, but relatively weak effect on usage (β=0.19), which is consistent with the present study result of no significant effect. In the study of Compeau et al., self-efficacy had a strong negative effect (β= -0.54) on anxiety (similar to the result of the present study). In the study of Compeau et al., self-efficacy was found to influence affect (β=0.39), which diverges with no significant effect found in the present study. The discrepancy may be due to the different nature of the relationship between self-efficacy and affect in a hedonistic context (one can love video games but be not good at playing them) and in a work-related context (it is unlikely that someone loves to work with computers and at the same time feels not confident about using them).

In the present study, the effect of Video Games Anxiety on Intention to Use was not found to be significant. This is consistent with the small effect size (β= -0.01) for a similar relationship in the study by Compeau et al. (1999)

Subject Matter Related. The effects of Attitude Towards Learning English and Subject Matter Self-efficacy on the Intention to Use MUVE for e-Learning were not found to be significant. On the other hand, the effects of Subject Matter Self-efficacy on Attitude Towards Learning English and English Class Anxiety were statistically significant and strong, with β= -0.52 and β=0.41, respectively. The latter finding was consistent with the findings by Compeau et al. (1999). However, the effects of Attitude Towards Learning English and Subject Matter Self-efficacy on Intention to Use in the present study should not be compared to Compeau et al. (1999), because Attitude Towards Learning English and Subject Matter Self-efficacy are not related to the use of computing technology.

The effect of English Class Anxiety on Intention to Use was found to be statistically significant, with β= 0.12. This finding was not consistent with the findings of Compeau et al. (1999), who found the effect of anxiety on usage to be not statistically significant. Again, the comparison is probably not appropriate, because English Class Anxiety is not directly related to the use of computing technology.

The impact of the Desire to Learn English on Perceived Usefulness of MUVE was found to be statistically significant, with β=0.16. The concept of Desire to Learn English can be compared to the concept of utility value in Chiu, Sun, Sun, and Ju (2007), which Chiu et al. found to affect technology use decisions.

6.4 Contribution to Knowledge

The study provided an empirical confirmation of the effectiveness of MUVE as an e-Learning platform. The study provided an objective empirical evidence of the effectiveness of MUVE as an e-Learning platform for situated scenario-based learning of oral business English communication skills by Malay learners. This finding contributes to the overall understanding of MUVE effectiveness as an e-Learning platform.

The finding suggests that situated scenario-based learning of oral business English communication skills in MUVE may be effective at improving the communication performance in the target scenario even when the duration of the instructional intervention is relatively brief, for as long as the intervention is focused on the

target scenario. Because in MUVE a target scenario can be implemented in a highly realistic simulated physical setting (thus, offering means for focusing on realistic settings), this finding is of particular importance.

Nonetheless, there was no statistically significant difference between learning gains in MUVE and learning gains in a face to face environment. This result, however, does not imply that MUVE based e-learning and face to face learning are equally effective for all possible scenarios. MUVE based learning may still be more effective than face to face learning in scenarios that depend on collaborative action in a particular physical environment more than the scenario used in the present study, such as scenarios around communication in emergency situations.

The study further validated technology acceptance theories in the context of MUVE. The study employed a research model based on existing technology acceptance theories, the Technology Acceptance Model and the Social Cognitive Theory, and thus validated these theories in the context of MUVE.

In particular, the study was the first to test a model based on Technology Acceptance Model and Social Cognitive Theory in the context of the acceptance of MUVE acceptance as an environment for learning English communication.

The study by Shen and Eder (2009) used a model based on Technology Acceptance Model and on elements of Social Cognitive Theory in the context of using MUVE as an environment for learning Management Information Systems. Unlike Management Information Systems, Business English is a non-technical subject; moreover, one may argue that unlike Management Information Systems, Business English mastery primarily depends on the development of skills, rather than on conceptual knowledge. Therefore, the context was different between the two studies. In particular, in an English as a second language class there is more emphasis on engaging in potentially embarrassing behaviour (in practising the skill) in front of other learners than in a Management Information Systems class. Therefore, constructs related to subject matter self-efficacy were included in the model used in the present study (these constructs were not included in the study by Shen and Eder), and one of these constructs, English Class Anxiety, was found to have effect on MUVE acceptance. In particular, students with higher levels of anxiety were more likely to accept MUVE. This was consistent with the view that

because of intermediation by avatars MUVE offers a safer environment for anxious learners.

The result of the present study is likely to be relevant to the use of MUVE for learning languages other than English, and even to using MUVE e-learning to learning other subjects focusing on learning skills, rather than on conceptual learning, such as midwifery or first aid courses implemented in MUVE environment.

6.5 Implications for Practice

This section discusses the implications of the results of the present study to practice.

6.5.1 Implications for Policy Makers

Even though there are many experience reports suggesting that MUVE e-learning is superior to face to face learning, the present study found no statistically significant difference between the two approaches. The results of the present study suggest that policy makers, when making decisions relating to funding projects involving the use of MUVE for e-Learning, should not make decisions solely based on subjective experience reports by MUVE e-learning innovators. Rather, the policy makers should assess the effectiveness of MUVE as an e-learning platform based on a range of sources of evidence, including studies relying on experimental or quasi-experimental designs.

The results of the present study do provide evidence that MUVE is effective in promoting communicative language learning. Therefore, policy makers should support further studies of MUVE effectiveness in supporting distance learning, in particular, distance second language learning.

6.5.2 Implications for Instructional Designers

The present study found no statistically significant difference between scenario based second language learning in MUVE and in a face to face environment. Therefore, instructional designers should be cautious about MUVE use for second

language teaching in settings where face to face learning is possible. For distance education, however, MUVE can be considered with more confidence.

In the present study, Video Games Affect was found to positively affect MUVE acceptance, suggesting that MUVE may be particularly suitable for students who like to play video games. Moreover, subject matter anxiety also had a positive effect on MUVE acceptance, suggesting that MUVE may be particularly suitable for students who are anxious about participating in classroom activities. This finding, however, is likely to be more relevant to teaching courses that emphasize skills, such as second language communication skills, than to courses focusing on conceptual knowledge.

6.5.3 Implications for Instructors

The present study found no statistical difference between scenario based second language learning in MUVE and in a face to face environment. Therefore, similar to instructional designers, instructors should be cautious about using MUVE when face to face learning is possible, but may use MUVE for distance learning with more confidence.

In the present study, subject matter anxiety was found to positively influence MUVE acceptance as and e-Learning platform by the learners. Therefore, when MUVE is available as an option, instructors should consider using MUVE with anxious learners and with learners fond of video games. Using MUVE may allow instructors to provide their learners with an environment where the learners can learn, interact, and support each other while feeling safe and unexposed because of the intermediation by avatars. The perception of safety may lead to greater learner involvement, even by learners who are reluctant to be active in conventional face to face environment.

In the present study, Video Games Affect was found to positively influence MUVE acceptance as and e-Learning platform by the learners. Therefore, in teaching learners who are fond of video games, the MUVE environment would be of benefit because such learners are already familiar with similar environments and are likely to have positive feelings about using MUVE for e-learning, possibly resulting in greater engagement and, ultimately, in greater progress in learning.

In the present study, the effect of Perceived Usefulness on Intention to Use was the strongest, suggesting that the best way for the instructors to convince learners to use MUVE is by demonstrating that MUVE is useful in helping the learners to achieve their learning goals.

6.5.4 Implications for Learners

The present study found no statistical difference between scenario based second language learning in MUVE and in a face to face environment. Therefore, learners may wish to refrain from enrolling into courses using MUVE for communicative language teaching when a face to face option is available, unless the learners feel that they are likely to feel more comfortable in MUVE environment because they feel anxious about participating in conventional face to face classroom activities or because they are used to virtual environments because of their background in playing video games. Students may enrol in distance courses using MUVE for communicative language teaching with more confidence.

6.5.5 Implications for the Use of MUVE in Malaysia to Improve English Language Proficiency of Malay Learners

Insufficient English language proficiency by Malaysian university graduates is seen as an obstacle in achieving the country's aim to become an industrialised nation by 2020 (Nunan, 2003; Ming et al., 2011). Although there is a shortage of qualified teachers in rural areas (Nunan, 2003), broadband Internet connectivity is gradually becoming available throughout the country (Razak & Malek, 2008; Zulkefli, Ainin, & Tengku, 2008; Lawas, 2010). At the same time, Malaysian students rate high on communication apprehension (Gillani, Sariff, Sulaiman, Halim, & Sari, 2010) and students' shyness makes it difficult to implement communicative English teaching practices in the classroom (Mustapha, Ismail, Singh, & Elias, 2010). The results of this study suggest that once connectivity is available, MUVE-based learning may become a viable option for teaching oral business English communication skills at a distance. At the same time, MUVE-based learning may lead to better participation of students with classroom anxiety problems.

6.6 Limitations of the Study

The problem addressed by the present study was motivated by the needs of Malay learners of English in Malaysia. Correspondingly, the study was conducted with Malay university students as participants. Therefore, the results of the study are particularly relevant to the possible use of MUVE to promote communicative English learning by Malay students in Malaysia. Nonetheless, with caution, the results may be generalizable to other cultural contexts, as well as to the use of MUVE to study subjects other than English or other languages (particularly, to courses emphasizing the learning of skills rather than the learning of concepts, particularly if situated, scenario based approach is used).

It has to be noted that statistical generalizability to the target population cannot be asserted because the participants were not drawn from the target population at random and the research relied on a convenience sample—participants that were available were used. Therefore, analytical generalization (see Yin, 2009), rather than statistical generalization to a well-defined population, applies—the readers of the present thesis will compare the context of the present study with their contexts of interest and make their own judgments regarding the extent to which the results apply to their particular circumstances. It has to be noted that statistical generalizability is very difficult to achieve in studies of e-Learning effectiveness and in studies of technology acceptance. In particular, all of such studies reviewed in the present thesis (see sections 2.6, 2.11, and 2.13) relied on convenience samples, and thus, shared this limitation with the present study.

Because the instructional intervention was conducted over a short period of time (about 40 minutes of teaching), the findings of the present study should be generalised to longer period of learning (e.g. over a full semester) only with caution. A short-term intervention might bias the results either in favour of MUVE (students more engaged in learning due to novelty effect) or against MUVE (as adjusting to the new environment takes cognitive effort, students might perform better after they have learned to use it). Moreover, the instructional intervention

was limited to a single scenario. There would have been more confidence in the outcomes if a broad range of scenarios could be covered.

It should be noted, though, that conducting an objective comparison over a long period of time and with a broad range of scenarios is fraught with considerable implementation difficulties. All of the objective comparisons between MUVE and face to face learning conducted in prior research and reviewed in section 2.11 involved a brief instructional intervention and thus shared the limitation in this respect with the present study.

Although the overwhelming majority of the students taking the English for Professional Communication course did volunteer to participate in the study, some of the students did not participate. If the non-participating students are particularly shy and thus have high levels of anxiety, it may have biased the conclusions relating to the effect of English Class Anxiety. If the anxious students who did not participate are the ones who would have found MUVE particularly attractive in helping them to overcome their anxiety, the effect of English Class Anxiety may have been underestimated.

The study relied on a cross sectional survey to study the factors affecting MUVE acceptance. Cross sectional surveys do not allow to test hypotheses regarding the direction of causation and are vulnerable to common method bias. These threats to validity were addressed by providing strong argument and relying on established theories to support the direction of causation suggested by the research model and by using the best practices, such as participant anonymity, in data collection. Nonetheless, infeasibility of using a more robust design, such a longitudinal study with independent and dependent variables measured by using different methods constitutes a limitation of the present study.

6.7 Suggestions for Future Research

It is desirable to replicate the present study using a broad range of instructional scenarios and using an instructional intervention of longer duration, approximating the duration of real university courses.

Adding a delayed post-test would also be of value as it would enable the assessment of student retention of the knowledge and skills learnt. As the

experiences in MUVE are more realistic because of the physical environment, the retention of knowledge gained in MUVE may be better compared to the retention of knowledge gained in the classroom.

MUVE-based learning environments are claimed to offer greater transferability of knowledge to real-life situations (Jones et al., 2005). For example, leaners that learned how to sell houses in MUVE may perform better as real estate agents compared with students who learned in the classroom because MUVE better represents the context of the real-world activity. Experimental study of transferability of knowledge gained in MUVE is a topic for further research.

It would be of value to replicate the study using participants from different populations. In particular, because the study implies that MUVE based learning is primarily appropriate for distance learning, it would be particularly interesting to test the validity of the results with typical distance students.

Future studies should explore instructor acceptance of MUVE, because the acceptance by the instructors is essential to ensure the success of MUVE e-Learning implementation.

6.8 Conclusion

The aim of the present study was to compare objectively the effectiveness of situated scenario based leaning of English in MUVE environment with face to face learning, as well as to explore the factors affecting the acceptance of MUVE by the learners in context of using MUVE as an environment for studying English as a second language. The study was motivated by the desire to contribute to helping Malay learners to achieve better proficiency at business English communication; therefore, the study was conducted with Malay participants and considered MUVE acceptance factors of particular relevance to Malay learners.

To compare the effectiveness of MUVE based language learning with face to face language learning, a controlled experiment was conducted. It was found that situated scenario based leaning of English in MUVE environment is as effective as face to face language learning, but not more effective than face to face language learning. There was no statistically significant difference between the learning gains in the two environments.

To explore the factors affecting the acceptance of MUVE by learners, a cross sectional survey was conducted, with the data fitted to a research model based on the Technology Acceptance Model and on the Social Cognitive Theory. It was found that Perceived Usefulness of MUVE e-Learning, Video Games Affect, and Subject Matter Anxiety affected the learners' Intention to Use MUVE for e-Learning.

Therefore, the study provided an objective empirical confirmation of the effectiveness of MUVE as an e-Learning platform and validated the Technology Acceptance Model and the Social Cognitive Theory in the context of predicting the acceptance of MUVE e-Learning. Further, the study has a number of implications for practice: the results suggest that MUVE is more likely to be suitable for distance learning than for learning on campus, and that MUVE is likely to be particularly suitable for students fond of video games and for students anxious about participating in face to face activities in the traditional classroom.

In future research, it is desirable to replicate the present study by using a longer instructional intervention covering a broad range of scenarios and to replicate the study with other populations of interest, such as with typical distance learners. It is also desirable to extend the study to explore instructor acceptance of MUVE based e-Learning.

References

Abdous, M.H., Facer, B. R., & Yen, C.-J. (2011). Academic effectiveness of podcasting: A comparative study of integrated versus supplemental use of podcasting in second language classes. *Computers & Education, 58*(1), 43-52.

Abdulla, D. (2012). Attitudes of college students enrolled in 2-year health care programs towards online learning. *Computers & Education*, *59*(4), 1215–1223.

Abrams, Z. I. (2003). The effect of Synchronous and Asynchronous CMC on Oral Performance in German. *The Modern Language Journal, 87*(2), 157-167.

Agarwal, R., & Karahanna, E. (2000). Time flies when you're having fun: Cognitive absorption and beliefs about information technology usage. *MIS Quarterly, 24*(4), 665-694.

Agarwal, R., & Prasad, J. (1999). Are individual differences germane to the acceptance of new information technologies? *Decision Sciences, 30*(2), 361-391.

Agostinha, S., Meek, J., & Herrington, J. (2005). Design methodology for the implementation and evaluation of a scenario-based online learning environment. *Journal of Interactive Learning Research, 16*(3), 229-242.

Ahn, Y. Y., Han, S., Kwak, H., Moon, S., & Jeong, H. (2007). Analysis of topological characteristics of huge online social networking services. In *Proceedings of the 16th international conference on World Wide Web* (pp. 835-844). ACM.

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.

Ajzen, I. (1991). The theory of planned behaviour. *Organizational Behaviour and Human Decision Processes, 50*(2), 179-211.

Alessi, S. M., & Trollip, S. R. (2001). *Multimedia for learning: Methods and development* (3 ed.). Boston: Allyn & Bacon.

Alshumaimeri, Y. (2011). The effects of wikis on foreign language students writing performance. *Procedia - Social and Behavioral Sciences*, 28, 755-763.

American Psychological Association (2010). *Publication manual of the American Psychological Association, 6th edition*. Washington, DC: American Psychological Association.

Amoako-Gyampah, K., & Salam, A. F. (2004). An extension of the technology acceptance model in an ERP implementatin environment. *Information & Management, 41*(6), 731-745.

Anderson, L. W., & Krathwohl, D. R. (Eds.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's Taxonomy of educational objectives*: Complete edition, New York : Longman.

Antonacci, D. M., & Modaress, N. (2008). Envisioning the educational possibilities of user-created virtual worlds. *AACE Journal 16*(2), 115-126.

Azuma, R. T. (1997). A survey of augmented reality. *Presence-Teleoperators and Virtual Environments*, 6(4), 355-385.

Babbie, E. (1990). *Survey research methods* (2nd ed.). Belmont,CA: Wadsworth.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. UK: Oxford University Press.

Bachman, L. F., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.

Bagozzi, R. P., Yi, Y., & Phillips, L. W. (1991). Assessing construct validity in organizational research. *Administrative science quarterly*, 36(3), 421-458.

Bahiyah, A. H. (1992). "Malu"-Shyness and shyness behaviour in the english language classroom: Who, what, where, when, how and why. *The English Teacher,* 11, 1-7.

Baker, S. C., Wentz, R. K., & Woods, M. M. (2009). Using virtual worlds in education: Second life as an educational tool. *Teaching of Psychology, 36*(1), 59-64.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory* Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organisational Behaviour and Human Decision Processes., 50*(2), 248-287.

Bandura, A. (1997). *Self-efficacy: The exercise of control.* New York: W.H. Freeman and Company.

Barclay, D., Higgins, C., & Thompson, R. (1995). The partial least squares (PLS) approach to causal modelling: Personal computer adoption and use as an illustration. *Technology Studies 2*(2), 285-309.

Berhannudin,M.S., Ahmad,E., Asri,S., Hussian,O.,Abdullah,S., &Khairul,A.S. (2010) Computer-mediated communication as a tool for improving the English language among adult learners in TVETprogramme. *Journal of Technical Education and Training 2*(1), 31-43.

Biasutti, M., & El-Deghaidy, H. (2012). Using wiki in teacher education: Impact on knowledge management processes and student satisfaction. *Computers & Education*, *59*(3), 861–872.

Blake, C. (2009). Potential of Text-Based Internet Chats for Improving Oral Fluency in a Second Language. *The Modern Language Journal*, *93*(2), 227-240.

Blake, R. J. (2007). New trends in using technology in the language curriculum. *Annual Review of Applied Linguistics*, 27, 76-97.

Bollen, K. A. (1989). *Structural equations with latent variables.* New York: John Wiley.

Boulos, M. N. K., Hetherington, L., & Wheeler, S. (2007). Second Life: An overview of the potential of 3-D virtual worlds in medical and health education. *Health Information and Libraries Journal, 24*(4), 233-245.

Bourgonjon, J., Grove, F. D., Smet, C. D., Van Looy, J., Soetaert, R., & Valcke, M. (2013). Acceptance of Game-Based Learning by Secondary School Teachers. *Computers & Education*, 67, 21-35.

Boyd, F. A. (1991). Business English and the case method: A reassessment. *TESOL Quarterly, 25*(4), 729-734.

Boyle, T. (2010). Layered learning design: Towards an integration of learning design and learning object perspectives. *Computers & Education*, 54(3), 661-668.

Bradley., T., & Lomicka., L. (2000). A Case study of learner interaction in technology-enhanced language learnng environments. *Journal of Educational Computing Research, 22*(3), 347-368.

Bronack, S., Cheney, A., Riedl, R., & Tasher, J. (2008). Designing virtual worlds to facilitate meaningful communication: Issues, considerations and lessons learned. *Journal of the Society for Technical Communication, 55*(3), 261-269.

Bronack, S., Riedl, R., & Tashner, J. (2006). Learning in the zone: A social constructivist framework for distance education in a 3-dimensional virtual world. *Interactive Learning Environments, 14*(3), 219-232.

Bronack, S., Sanders, R., Cheney, A., Riedl, R., Tashner, J., & Matzen, N. (2008). Presence pedagogy: Teaching and learning in a 3D virtual immersive world. *International Journal of Teaching and Learning in Higher Education*, 20(1), 59-69.

Brown, H. D. (2000). *Principles of language teaching and learning*. White Plains, NY: Longman.

Brown, H. D. (2001). *Teaching by principles: An interactive approach to language pedagogy* (2nd ed.). White Plains, NY.: Addison Wesley Longman.

Brown, J. D., & Bailey, K. M. (1984). A categorical instrument for scoring second language writing skills. *Language Learning, 34*(1), 21-42.

Brown, J. S., Collins, A., & Duguid, P. (1989). Situated cognition and the culture of learning. *Educational Researcher, 18*(1), 32-41.

Brown, S. A., Massey, A. P., Montoya-Weiss, M. M., & Burkman, J. R. (2002). Do I really have to? User acceptance of mandated technology. *European Journal of Information Systems, 11*(4), 283-295.

Bryman, A., & Bell, E. (2003). *Business research methods*. New York: Oxford University Press.

Burdea, G., & Coiffet, P. (2003). Virtual reality technology. *Presence: Teleoperators & Virtual Environments*, *12*(6), 663-664.

Burkhart, M., & Brass, D. (1990). Changing patterns of patterns of change: the effect of a change in technology on social network structure and power. *Administrative Science Quarterly, 35*(1), 104-127.

Burton-Jones, A., & Hubona, G. S. (2005). Individual differences and usage behaviour: Revisiting a technology acceptance model assumption. *ACM SIGMIS Database, 36*, 58-77.

Butler, Y. G. (2004). What level of English proficiency do elementary school teachers need to attain to teach EFL? Case studies from Korea, Taiwan and Japan. *TESOL Quarterly, 14*(2), 97-114.

Byrne, B. M. (2001). *Structural Equation Modeling with AMOS: Basic Concepts, Applications, and Programming.* Mahwah,N.J.: Lawrence Erlbaum Associates.

Caballé, S., Xhafa, F., & Barolli, L. (2010). Using mobile devices to support online collaborative learning. *Mobile Information Systems*, 6(1), 27-47.

Cabanero-Johnson, P. S., & Berge, Z. (2009). Digital natives: Back to the future of microworlds in a corporate learning organization. *The Learning Organization, 16*(4), 290-297.

Campbell, C. M., & Ortiz, J. (1991). Helping students overcome foreign language anxiety: A foreign language anxiety workshop. In E. K. Horwitz & D. J. Young (Eds.), *Language Anxiety: From theory and research to classroom implications* (pp. 153-168). Englewood Cliffs, N.J., Prentice Hall.

Campbell, D. E., Wells, J. D., & Valacich, J. S. (2012). Breaking the Ice in B2C Relationships: Understanding Pre-Adoption E-Commerce Attraction. *Information Systems Research*. Advance online publication. doi: 10.1287/isre.1120.0429

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1-47.

Cargill-Kipar, N. (2009). My dragonfly flies upside down! Using Second Life in multimedia design to teach students programming. *British Journal of Educational Technology, 40*(3), 539-542.

Chalhoub-Deville, M. (1997). Theoretical models, assessment frameworks and test construction. *Language Testing, 14*(1), 3-22.

Chang, S. C., & Tung, F. C. (2007). An empirical investigation of students' behavioral intentions to use the online learning course websites. *British Journal of Educational Technology, 39*(1), 71-83.

Charles, D., Charles, T., McNeill, M., Bustard, D., & Black, M. (2011). Game-based feedback for educational multi-user virtual environments. *British Journal of Educational Technology*, *42*(4), 638-654.

Chau, P. Y. K. (2001). Influence of computer attitude and self-efficacy on IT usage behaviour. *Journal of Eng-User Computing, 13*(1), 26-33.

Chau, P. Y. K., & Hu, P. J. H. (2002). Investigating healthcare professionals' decisions to accept telemedicine technology: An empirical test of competing theories. *Information & Management, 39*(4), 297-311.

Chee, Y. S. (2007). Embodiment, embeddedness, and experience: game-based learning and the construction of identity. *Research and Practice in Technology Enhanced Learning, 2*(1), 3-30.

Cheong, D. (2010). The effects of practice teaching sessions in second life on the change in pre-service teachers' teaching efficacy. *Computers &Education, 55*(1), 868-880.

Cheong, Y., Pajares, F., & Oberman, P. (2004). Motivation and Academic Help-Seeking in HIgh School Computer Science. *Computer Science Education, 14*(1), 3-19.

Cheung, L. S. (2006). A Constructuivist approach to designing computer supported concept-mapping environment. *International Journal of Instructional Media, 33*(2), 153-164.

Chin, W. W. (2003). A permutation procedure for multi-group comparison of PLS models *Paper prsented at the PLS '03 International Symposium*. Lisbon, Portugal.

Chin, W. W., & Frye, T. A. (1998). PLS-Graph (Version 2.91.03.04).

Chin, W. W., & Newsted, P. R. (1996). Structural equation modelling analysis with small samples: Using partial least squares. In R. H. Hoyle (Ed.), *Statistical strategies for small-sample research* (pp. 307-341). California, USA: Sage Publication, Inc.

Chin, W. W., & Todd, P. A. (1995). On the use, usefulness, and ease of use of structural equation modeling in MIS research: A note of caution. *MIS Quarterly*, 19(2), 237-246.

Chiu, C.-M., Sun, S.-Y., Sun, P.-C., & L.Ju, T. (2007). An empirical analysis of the antecedents of web-based learning continuance. *Computers & Education, 49*(4), 1224-1245.

Chou, C.-P., Bentler, P. M., & Satorra, A. (1991). Scaled test statistics and robust standard errors for non-normal data in covariance structure analysis: A Monte Carlo study. *British Journal of Mathematical and Statistical Psychology 44*(2), 347-357.

Chu, S. K., Chan, C. K., & Tiwari, A. F. (2012). Using blogs to support learning during internship. *Computers & Education*, *58*(3), 989-1000.

Chwelos, P., Benbasat, I., & Dexter, A. S. (2001). Research report: empirical test of an EDI adoption model. *Information Systems Research, 12(3)*, 304-321.

Ciekanski, M., & Chanier, T. (2008). Developing online multimodal verbal communication to enhance the writing process in an audio-graphic conferencing environment. *ReCALL, 20*(2), 162-182.

Clark, K. (2003). Impact of Technology on the Academic Self-Efficacy and Career Selection of African-American Students. *Information Technology in Childhood Education Annual, 79*(11), 79-89.

Cobb, S., Heaney, R., Corcoran, O., & Henderson-Begg, S. (2009). The learning gains and student perceptions of a Second Life Virtual lab. *Bioscience Education, 13*(5), 1-9.

Cochrane, K. (2006, December, 2006). *Case study: International spaceflight museum.* Paper presented at the Proceedings of the first Second Life education workshop.

Cochrane, T., & Bateman, R. (2010). Smartphones give you wings: Pedagogical affordances of mobile Web 2.0. *Australasian Journal of Educational Technology,* 26(1), 1-14.

Cohen, J. (1988). *Statistical analysis for the behavioral sciences* (2 ed.). New York: Lawrence Erlbaum.

Cole, M. (2009). Using Wiki technology to support student engagement: Lessons from the trenches. *Computers & Education, 52*(1), 141-146.

Compeau, D. R., & Higgins, C. A. (1995a). Application of social cognitive theory to training for computer skills. *Information Systems Research, 6*(2), 118-143.

Compeau, D. R., & Higgins, C. A. (1995b). Computer self-efficacy: Development of measure and initial test. *MIS Quarterly, 19*(2), 189-211.

Compeau, D. R., Higgins, C. A., & Huff, S. L. (1999). Social cognitive theory and individual reactions to computing technology: A longitudinal study. *MIS Quarterly, 23*(2), 145-158.

Cool, K., Dierickx.I, & Jemison, D. (1989). Business strategy, market structure and risk return relationship: A structural approach. *Strategic Management Journal, 10*(6), 507-522.

Corbit, M. (2002). Building virtual worlds for informal science learning (SciCentr and SciFair) in the Active Worlds Educational Universe (AWEDU). *Presence: Teleoperators Virtual Environments 11*(1), 55-67.

Costello, A. B., & Osborne, J. W. (2005). Best practices in exploratory factor analysis: Four recommendations for getting the most from your analysis. *Practical Assessment Research and Evaluation, 10*(7), 1-9.

Crosier, J. K., Cobb, S., & Wilson, J. R. (2000). Experimental comparison of virtual reality with traditional teaching methods for teaching radioactivity. *Education and Information Technologies, 5*(4), 329-343.

D' Agostino, R. B. (1986). Tests for normal distribution. In R. B. D"Agostino & M. A. Stephens (Eds.), *Goodness of fit techniques* (pp. 367-419). New York: Marcel Dekker.

Daft, R. L., & Lengel, R. H. (1986). Organizational information requirements, media richness and structural design. *Management Science, 32*(5), 554-571.

Dalgarno, B., Lee, M. J. W., Carlson, L., Gregory, S., & Tynan, B. (2011). An Australian and New Zealand scoping study on the use of 3D immersive virtual worlds in higher education. *Australian Journal of Educational Technology, 27*(1), 1-15.

Daniel, D. B. (2013). E-textbooks at what cost? Performance and use of electronic v. print texts. *Computers & Education*, *62*, 18-23.

Darus, S. (2010). The current situation and issues of the teaching of English in Malaysia. *立命館言語文化研究*, 22(1), 19-27.

David, M. K., & Govindasamy, S. (2005). Negotiating a language policy for Malaysia: Local demand for affirmative action versus challenges from globalization. In A. S. Canagarajah (Ed.), *Reclaiming the Local in Language Policy and Practice* (pp. 123-145). Mahwah, New Jersey: Lawrence Erlbaum Associates.

David, M. K., Cavallaro, F., & Coluzzi, P. (2009). Language policies – impact on language maintenance and teaching: Focus on Malaysia, Singapore, Brunei and the Philippines. *The Linguistics Journal*, Special Edition, September 2009, Language, Culture and Identity in Asia, 155-191.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319-340.

Davis, F. D., Gagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*(8), 982-1003.

Dede, C. (1995). The evolution of constructivist learning environments: Immersion in distributed, virtual worlds. *Educational Technology 35*(5), 46-52.

DeVellis, R. F. (1991). Scale development theory and applications. *Journal of Educational Measurement, 31*(1), 79-82.

Dewar, T., & Whittington, D. (2000). Online learners and their learning strategies. *Journal of Educational Computing Research, 23*(4), 385-403.

Di Serio, Á., Ibáñez, M. B., & Kloos, C. D. (2012). Impact of an augmented reality system on students' motivation for a visual art course. *Computers & Education*. Advance online publication. doi: 10.1016/j.chb.2013.03.006

Dickey, M. (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.

Doherty, P., & Rothfarb, R. (2006). *Building an interactive science museum in Second Life.* Paper presented at the Proceedings of the first Second Life education workshop.

Douglas. (2000). *Assessing languages for specific purposes*. Cambridge: Cambridge University Press.

Duffy, T. M., & Cunningham, D. J. (1996). Constructivism: Implications for the design and delivery of Instruction. In D. H. Jonassen (Ed.), *Handbook of Research for Educational Communications and Technology*. New York, USA.: Macmillan Library Reference.

Economic Planning Unit (2012). *The Malaysian economy in figures 2012.* Federal Government Administrative Centre, Putrajaya, Malaysia: Percetakan Nasional Malaysia Berhad, Report No JD004310.

Edirisingha, P., Nie, M., Pluciennik, M., & Young, R. (2009). Socialisation for learning at a distance in a 3-D multi-user virtual environment. *British Journal of Educational Technology 40*(3), 458-479.

El-Deghaidy, H., & Nouby, A. (2008). Effectiveness of a blended e-learning cooperative approach in an Egyptian teacher education programme. *Computers & Education 51*(3).

Emurian, H. (2004). A Programmed Instruction Tutoring System for Java: Consideration of Learning performance and Software Self-Efficacy. *Computers in Human Behaviour, 20*(3), 423-459.

Eschenbrenner, B., Nah, F. F. H., & Siau, K. (2008). 3-D virtual worlds in education: Applications, benefits, issues and opportunities. *Journal of Database Management, 19*(4), 91-110.

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods, 4*(3), 272-299.

Fageeh, A. I. (2011). EFL learners' use of blogging for developing writing skills and enhancing attitudes towards English learning: An exploratory study. *Journal of Language and Literature*, 2(1), 31-48.

Falk, R. F., & Miller, N. B. (1992). *A primer for soft modelling* (1st ed.). Ohio, USA: The University of Akron Press.

Faridah, C. A. (2005). *Development of Teaching and Learning Reading Comprehension Multimedia Courseware for UTMKL Diploma Students.* Master's Thesis, National University of Malaysia, Selangor.

Feldman, L. (2006). *Camp global kids: A case study of bringing a global youth development model into Teen Second Life.* Paper presented at the Proceedings of the first Second Life education workshop.

Felix, U. (2005a). Analysing recent CALL effectiveness research—towards a common agenda. *Computer Assisted Language Learning*, *18*(1-2), 1-32.

Felix, U. (2005b). What do meta-analyses tell us about CALL effectiveness? *ReCALL, 17*(2), 269-288.

Felix, U. (2008). The unreasonable effectiveness of CALL: What have we learned in two decades of research? *ReCall*, *20*(2), 141-161.

Fetscherin, M., & Lattemann, C. (2008). User acceptance of virtual worlds. *Journal of Electronic Commerce Research, 9*(3), 231-242.

Foo, B., & Richards, C. (2004). English in Malaysia. *RELC Journal, 35*(2), 229-240.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal in Marketing Research, 18*(3), 328-388.

Fornell, C., Tellis, G. J., & Zinkhan, G. M. (1982). *Validity assessment: A structural equations approach using partial least squares.* Paper presented at the AMA Educator's Proceedings, Chicago, USA.

Friedrich, H. F., & Hron, A. (2010). Factors influencing pupil's acceptance of an e-learnng system for secondary schools. *Journal of Educational Computing, 42*(1), 63-78.

Furió, D., GonzáLez-Gancedo, S., Juan, M., Seguí, I., & Costa, M. (2013). The effects of the size and weight of a mobile device on an educational game. *Computers & Education*, 64, 24-41.

Fusilier, M., & Durlabhji, S. (2005). An exploration of student internet use in India. *Campus-Wide Information Systems, 22*(4), 233-246.

Gamage, V., Tretiakov, A., & Crump, B. (2011). Teacher perceptions of learning affordances of multi-user virtual environments. *Computers & Education, 57*(4), 2406-2413.

Gardner, R. C. (1985). *Social psychology and second language learning: The role of attitudes and motivation.* London, UK: Edward Arnold.

Gardner, R. C. (2001). Integrative motivation and second language acquisition. In Z. Dornyei & R. Schmidt (Eds.), *Motivation and second language acquisition.* (pp. 1-19).

Honolulu, Hawai'i: University of Hawai'i, Second Language Teaching and Curriculum Centre.
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., Tremblay, P. F., & Masgoret, A. M. (1997). Towards a full model of second language learnng: An empirical investigation. *The Modern Language Journal, 81*(3), 344-362.
Gefen, D., & Straub, D. W. (2005). A practical guide to factorial validity using PLS-graph: Tutorial and annotated example. *Communications of the AIS, 16*(25), 91-109.
Gill, S. K. (1993). Standards and pedagogical norms for teaching English in Malaysia. *World Englishes, 12*(2), 223-238.
Gill, S. K. (2005). Language policy in Malaysia: Reversing direction. *Language Policy*, 4(3), 241-260.
Gill, S. K. (2012). The Complexities of Re-reversal of Language-in-Education Policy in Malaysia. *Multilingual Education*, 1, 45-61.
Gillani, S. W., Sariff, A., Sulaiman, S. A. S., Halim, N. A., & Sari, Y. O. (2010). Does communication apprehension reflect learning style; a population based survey among Malaysian students. *Technics Technologies Education Management, 5*(2), 334-341.
Gioia, D. A., & Pitre, E. (1990). Multiparadigm perspectives on theory building. *Academy of Management, 15*(4), 584-602.
Gist, M. E., Schwoerer, C., & Rosen, B. (1989). Effects of alternative training methods on self-efficacy and performance in computer software training. *Journal of Applied Psychology, 74*(6), 884-891.
Glaser, M. (2007). Your guide to social networking online. *PBS MediaShift*, August.
Golonka, E. M., Bowles, A. R., Frank, V. M., Richardson, D. L., & Freynik, S. (2012). Technologies for foreign language learning: a review of technology types and their effectiveness. *Computer Assisted Language Learning.* doi: 10.1080/09588221.2012.700315
Gong, M., & Xu, Y. (2004). An enhanced technology acceptance model for web-based learning. *Journal of Information Systems Education, 15*(4), 365-374.
Good, J., Howland, K., & Thackray, L. (2008). Problem-based learning spanning and virtual worlds: a case study in Second Life. *Association for Learning Technology Journal, 16*(3), 163-172.
Government of Malaysia.(1976). Third Malaysia Plan 1976-1980. Kuala Lumpur: Government Press
Graham, J., & Beardsley, R. S. (1986). English for Specific Purposes: Content, langauge and communication in a pharmacy course model. *TESOL Quarterly, 20*(1), 227-245.
Hackathorn, E. J. (2006). *Designing an educational island inside Second Life for the National Oceanic and Atmospheric Administration (NOAA) Earth System Research Laboratory (ESRL).* Paper presented at the Proceedings of the Second Life Education Workshop at the Second Life Community Convention, San Francisco, CA.
Hair, J. F., Anderson, R. E., Tatham, R. L., & Black, W. C. (1998). *Multivariate data analysis with readings.* (5th ed.). New York, USA: MacMillan.
Hair, J. F., Black, W. C., Babin, Anderson, R. E., & Tatham, R. L. (2006). *Multivariate data analysis.* (6th ed.). Upper Saddle River, NJ: Pearson Education.
Hamidah, Y. (2003). *Globalisation and diversity: The need for English language.* Paper presented at the Proceedings of the forty-eighth world assembly of the International Council Education on Teaching, Melbourne, Australia.
Hamp-Lyons, L. (1991). Scoring procedures for ESL contexts. In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts (pp. 241-276)*. Norwood.NJ: Ablex.

Henning, P. (1998). Everyday Cognition and Situated Learning. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology (pp. 143-168)* (2nd ed.). New York: Simon & Schuster.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. In R. R. Sinkovics & P. N. Ghauri (Eds.), *New challenges to international marketing (Advances in international marketing)* (Vol. 20, pp. 277-319): Emerald Group Publishing Limited.

Hew, K. F., & Cheung, W. S. (2010). Use of three-dimensional (3-D) immersive virtual worlds in K-12 and higher education settings: A review of the research. *British Journal of Educational Technology, 41*(1), 33-55.

Hill, T., Smith, N. D., & Mann, M. F. (1986). Communicating innovations: Convincing computer phobics to adopt innovative technologies. *Advances in Consumer Research, 13*(1), 419-422.

Hill, T., Smith, N. D., & Mann, M. F. (1987). Role of efficacy expectations in predicting the decision to use advance technologies: The case of computers. *Journal of Applied Psychology, 72*(2), 307-313.

Hincks, R. (2002). *Speech recognition for language teaching and evaluating: A study of existing commercial products.* Paper presented at the Proceedings from ICSLP 2002, Denver, Colorado,USA.

Hirschman, E. C., & Holbrook, M. B. (1982). Hedonic consumption:Emerging concepts, methods and propositions. *Journal of Marketing, 146*(3), 92-101.

Hislop, D. (2003). The complex relations between communities of practice and the implementation of technological innovations. *International Journal of Innovation Management, 7*(2), 163-188.

Hofstede, G. (1991). *Cultures and Organizations*. New York: McGraw-Hill.

Hong, W., Thong, J. Y. L., Wong, W.-M., & Tam, K.-Y. (2002). Determinants of user acceptance of digital libraries: An empirical examination of individual differences and system characteristics. *Journal of Management Information Systems, 18*(3), 97-124.

Horwitz, E. K., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal, 70*(2), 125-132.

Hu, P. J. H., Chau, P. Y. K., Sheng, O. R. L., & Tam, K. Y. (1999). Examining technology acceptance model using physician acceptance of telemedicine technology. *Journal of Management Information Systems, 16*(2), 91-112.

Hu, P. J. H., Lin, C., & Chen, H. (2005). User acceptance of intelligence and security informatics technology: A study of COPLINK. *Journal of the American Socirty for Information Science and Technology, 56*(3), 235-244.

Huang, H. M. (2002). Toward constructivism for adult learners in online learning environments. *British Journal of Educational Technology, 33*(1), 27-37.

Huang, H. M., Rauch, U., & Liaw, S. S. (2010). Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach. *Computers & Education, 55*(3), 1171-1182.

Hung, D., & Chen, D. (2002). Two kinds of scaffolding: The dialectical process within the authenticity-generalizability (A-G) continuum. *Education Technology & Society, 5*(4), 148-153.

Hussein, Z., Wahid, N. A., & Saad, N. (2009) *Behavioral study on Malaysian game player experiences: how the embedded information inside a computer game affect players' behaviour*. Paper presented at the 9th Global conference on business and economics,Cambridge University, UK.

Hutchinson, T., & Waters, A. (1987). *English for specific purposes: A learning-centered approach.* Cambridge: Cambridge University Press.

Hymes, D. (1974). *Foundations of Sociolinguistics: An ethnographic approach.* Philadelphia: University of Pennsylvania Press.

Ismail, N. W. (2008). The Effect of Language on Trade: The Malaysian Case. *Prosiding Persidangan Kebangsaan Ekonomi Malaysia ke-III (PERKEM III)*, 787 – 791.

Jakobson, M., & Taylor, T. L. (2003). *The Sopranos meets EverQuest: Social networking in massively multiplayer online games.* Paper presented at the Conference proceedings of the 2003 Digital Arts and Culture (DAC) conference, Melbourne, Australia.

Jin, L., Wen, Z., & Gough, N. (2010). Social virtual worlds for technology-enhanced learning on an augmented learning platform. *Learning, Media and Technology, 35*(2), 139-153.

Johnson, G. M., & Johnson, J. A. (2006). Computer technology and human learning: Review of recent quantitative syntheses. *Technology, Instruction, Cognition, and Learning*, 4(3), 287-301.

Jonassen, D. H. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (pp. 215-239). Mahwah,NJ: Lawrence Erlbaum Associates.

Jones, J. G., Morales, C., & Knezek, G. A. (2005). 3-Dimensional online learnng environments: Examining attitudes toward information technology between students in internet-based 3-dimensional and face-to-face classroom instruction. *Educational Media International, 42*(3), 219-236.

Kahai, S. S., & Cooper, R. B. (2003). Exploring the core concepts of media richness theory: the impact of cue multiplicity and feedback immediacy on decision quality. *Journal of Management Information Systems, 20*(1), 263-299.

Kalina, C., & Powell, K. C. (2009). Cognitive and social constructivism: Developing tools for an effective classroom. *Education*, *130*(2), 241-250.

Kam, H. W. (2002). English language teachng in East Asia today: An overview. *Asia-Pacific Journal of Education, 22*(2), 1-22.

Kamarainen, A. M., Metcalf, S., Grotzer, T., Browne, A., Mazzuca, D., Tutwiler, M. S., & Dede, C. (2013). EcoMOBILE: Integrating Augmented Reality and Probeware with Environmental Education Field Trips. *Computers & Education*. Advance online publication. doi: 0.1016/j.compedu.2013.02.018.

Kassim, H., & Ali, F. (2010). English communicative events and skills needed at the workplace: Feedback from the industry. *English for Specific Purposes, 29*(3), 168-182.

Ke, F. (2013). Computer-game-based tutoring of mathematics. *Computers & Education*, *60*, 448–457.

Kernan, M., & Howard, G. S. (1990). Computer anxiety and computer attitudes: An investigation of construct and predictive validity issues. *Educational and Psychological Measurement, 50*(3), 681-690.

Kim, L. S. (2003). Multiple identities in a multicultural world: A Malaysian perspective. *Journal of Language, Identity and Education, 2*(3), 137-158.

Kitchenham, A. B., & Pfleeger, L. S. (2002). Principles of survey research: Part 3: Constructing a survey instrument. *ACM SIGSOFT Software Engineeering Notes, 27*(2), 20-24.

Kline, R. B. (2005). *Principles and practice of structural equation modeling* (2 ed.). New York: Guilford Press.

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 application* (pp. 109-126). Englewood Cliffs, New Jersey, USA: Prentice-Hall.

Koufaris, M. (2002). Applying the Technology Acceptance Model and flow theory to online consumer behavior. *Information Systems Research, 13*(2), 205-233.

Krashen, S. D., & Terrell, T. D. (1983). *The natural approach: Language acquisition in the classroom*. Hayward, CA: The Alemany Press.

Kuo, T. C. T., Shadiev, R., Hwang, W. Y., & Chen, N. S. (2012). Effects of applying STR for group learning activities on learning performence in a synchronous cyber classroom. *Computers & Education, 58*(1), 600-608.

Lainema, T., & Nurmi, S. (2006). Applying authentic, dynamic learning environment in real world business. *Computers & Education, 47*(1), 94-115.

Lan, C. O. T., Khaun, A. L. C., & Singh, P. K. S. (2011). *Malaysian Journal of ELT Research*, 7(2), 82-103.

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge: Cambridge University Press.

Lawas, K. (2010, April 30, 2010). Broadband reduces rural-urban digital divide, *Borneo Post Online*. Retrieved from http://www.theborneopost.com/2010/04/30/broadband-reduces-rural-urban-digital-divide

Lazaro, D. C., & Medalla, E. M. (2004). *English as the Language of Trade, Finance and Technology in APEC: An East Asia Perspective.* Philippine Institute for Development Studies, Discussion Paper Series No 2004-36.

Lee, B. C., Yoon, J. O., & Lee, I. (2009). Learners' acceptance of e-learning in South Korea: Theories and results. *Computers & Education, 53*(4), 1320-1329.

Lee, E. A. L., Wong, K. W., & Fung, C. C. (2009). *Learning effectiveness in a desktop virtual reality-based learning environment.* Paper presented at the Proceedings of the 17th International Conference on Computers in Education: Asia-Pacific Society for Computers in Education, Hong Kong.

Leidner, D. E., & Jarvenpaa, S. L. (1995). The use of information technology to enhance management school education: A theoretical view. *MIS Quarterly, 19*(3), 265-291.

Levy, M. (1997). *Computer assisted language learning*. Oxford: Clarendon Press.

Lim, C. P., Nonis, D., & Hedberg, J. (2006). Gaming in a 3D multiuser virtual environment: Engaging students in science class. *British Journal of Educational Technology, 37*(2), 211-231.

Lin, P. C., Hou, H. T., Wang, S. M., & Chang, K. E. (2013). Analyzing knowledge dimensions and cognitive process of a project-based online discussion instructional activity using Facebook in an adult and continuing education course. *Computers & Education*, *60*(1), 110–121.

Liu, M., Moore, Z., Graham, L., & Lee, S. (2002). A look at the research on computer-based technology use in second language learnng: A review of the literature from 1990-2000. *Journal of Research on Technology in Education, 34*(3), 250-272.

Lu, M. (2008). Effectiveness of vocabulary learning via mobile phone. *Journal of computer assisted learning*, *24*(6), 515-525.

Luarn, P., & Lin, H. H. (2005). Toward an understanding of the behavioral intention to use mobile banking. *Computers in Human Behavior, 21*(6), 873-891.

Lucia, A. D., Francese, R., Passero, I., & Tortora, G. (2009). Development and evaluation of a virtual campus on Second Life: The case of Second DMI. *Computers & Education, 52*(1), 220-233.

Maarof, N., Osman, K., Yamat, H., & Yunus, M. (2003). *Keupayaan penguasaan kemahiran bahasa inggeris di kalangan pelajar melayu dalam arus globalisasi*. Paper presented at the Seminar Pembinaan Keupayaan Melayu dalam Arus Globalisasi, Universiti Kebangsaan Malaysia, Malaysia.

MacIntyre, P. D., & Gardner, R. C. (1989). Anxiety and second language learning: Toward a theoretical clarification. *Language Learning, 39*(2), 251-275.

MacIntyre, P. D., & Gardner, R. C. (1994a). The effects of induced anxiety on three stages of cognitive processing in computerised vocabulary learning. *Studies in Second Language Acquisition, 16*(1), 1-17.

Mackey, R., & Mountford, A. J. (1978). *English for Specific Purposes*. London: Longman.

Malaysia, D. O. S. (2007). *Population*. Retrieved 28 December, from http://www.statistics.gov.my

Mallan, K., Foth, M., Greenway, R., & Greg, T. (2010). Serious playground: using Second Life to engage high school students in urban planning. *Learning, Media and Technology, 35*(2), 203-225.

Malaysian Education Ministry, 1989. *Compedium: A Handbook for ELT teachers*.1/89

Malaysian Employers Federation (2004). Facing the realities of the world of work [PowerPoint slides]. Retrieved from: http://www.epu.gov.my/seminars.

Manan, A. A., & Shamsudin, S. (2012). Comparing form four Malay and Chinese students' spoken English. *The English Teacher*, XLI (1), 13-26.

Marcoulides, G. A., & Saunders, C. (2006). PLS: A silver bullet? *MIS Quarterly, 30*(2), 3-9.

Mardia, K. V. (1970). Measures of multivariate skewness and kurtosis with applications. *Biometrika, 57*(3), 519-530.

Martin, S., Diaz, G., Sancristobal, E., Gil, R., Castro, M., & Peire, J. (2011). New technology trends in education: Seven years of forecasts and convergence. *Computers & Education*, *57*(3), 1893-1906.

Mathieson, K. (1991). Predicting user intentions: Comparing the technology acceptance model with the theory of planned behaviour. *Information Systems Research, 2*(3), 173-191.

Mathieson, K., & Eileen, P. (2001). Extending the technology acceptance model: The influence of perceived user resources. *The Database for Advances in Information Systems, 32*(3), 86-112.

Maurer, M. M. (1994). Computer anxiety correlates and what they tell us: A literature review. *Computers in Human Behaviour, 10*(3), 369-376.

Mayer, R. E. (1987). *Educational Psychology*. New York, NY: Harper Collins.

Mayer, R. E. (1998). Cognitive theory for education: what teachers need to know. In N. M. Lambert & B. L. McCombs (Eds.), *How students learn: Reforming schools through learner-centered education* (pp. 353-377). Washington, DC.: American psychological association.

McCarthy, M. J. (1998). *Spoken language and applied linguistics*. Cambridge: Cambridge University Press.

McKay, S. L. (2002). *Teaching English as an international language*. Oxford: Oxford University Press.

McLellan, H. (1986). Situated learning: Multiple perspectives. In H. McLellan (Ed.), *Situated learning perspectives*. Englewood Cliffs, NJ: Educational technology publications.

Meluso, A., Zheng, M., Spires, H. A., & Lester, J. (2012). Enhancing 5th graders' science content knowledge and self-efficacy through game-based learning. *Computers & Education*, *59*(2), 497-504.

Met, M. (1998) Curriculum decision making in content-based language teaching. In J. Cenoz and F. Genesee (eds) *Beyond Bilingualism: Multilingualism and Multilingual Education* (pp. 35 – 63). Clevedon: Multilingual Matters.

Micceri, T. (1989). The unicorn, the normal curve and other improbable creatures. *Psychological Bulletin, 105*(1), 155-166.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research. *Computers & Education, 56*(3), 769-780.

Ming, T. S. (2004). Learning English in multicultural Malaysia: Are learners motivated? *Journal of Language and Learning, 2*(3), 142-153.

Ming, T. S., Ling, T. S., & Jaafar, N. H. (2011). Attitudes and motivation of Malaysian secondary students towards learning english as a second language: A case study. *The Southeast Asian Journal of English Language Studies, 17*(1), 40-54.

Minocha, S., & Reeves, A. J. (2010). Design of learning spaces in 3D virtual worlds: An empirical investigation of "Second Life". *Learning, Media and Technology, 35*(2), 111-137.

Mirjam, H., & Stella, H. (2005). Exploring the link between language anxiety and learner self-management in open language learning contexts. *European Journal of Open, Distance and E-learning., 2*, 1-12.

Mohamed, H., Jan, N. Y. C., & Daud, N. M. N. (2010). Exposure of computer games among IHL students in Malaysia: Case study of computer science students in UiTM Terengganu. *Computer and Information Science*, *3*(1), 144-151

Mohan, P., & Brooks, C. (2003). Learning objects on the semantic web. In *The 3rd IEEE International Conference on Advanced Learning Technologies* (pp. 195-199). IEEE

Monahan, T., McArdle, G., & Bertolotto, M. (2008). Virtual reality for collaborative e-learning. *Computers & Education, 50*(4), 1339-1353.

Moon, J. W., & Kim, Y. G. (2001). Extending the TAM for a World Wide Web context. *Information & Management, 38*(4), 217-230.

Moore, G. C., & Benbasat, I. (1991). Development of an instrument to measure the perceptions of adopting an information technology innovation. *Information Systems Research, 2*(3), 192-222.

Moreno, L., Gonzalez, C., Castilla, I., Gonzalez, E., & Sigut, J. (2007). Applying a constructivist and collaborative methodological approach in engineering education. *Computers & Education, 49*(3), 891-915.

Mumby, J. (1987). *Communicative syllabus design.* Cambridge.: Cambridge University Press.

Musa, N. C., Koo, Y. L., & Azman, H. (2012). Exploring English language learning and teaching in Malaysia. *GEMA: Online Journal of Language Studies*, 12(1), 35-51.

Mustapha, W. Z. W., Ismail, N., Singh, R. S. R., & Elias, S. (2010). ESL students communication apprehension and their choice of communicative activities. *ASEAN Journal of Teaching and Learning in Higher Education, 2*(1), 22-29.

Nadzri, S. (2005, 13 November). Preparing for competitive life, *The News Straits Times,* p.5.

Nair-Venugopal, S. (2006). An interactional model of English in Malaysia. *Journal of Asian Pacific Communication 16*(1), 51-75.

Ngai, E. W. T., Poon, J. K. L., & Chan, Y. H. C. (2007). Empirical examination of the adoption of WebCT using TAM. *Computers & Education, 48*(2), 250-267.

Nonaka, I. (2007). The knowledge-creating company. *Harvard Business Review, 85*(85), 162-171.

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in Asia-Pacific region. *TESOL Quarterly, 37*(4), 589-613.

Nunan, D. (2005). *Task-based language teaching.* Cambridge, UK: Cambridge University Press.

Nunnally, J. (1978). *Psychometric Theory.* (2nd ed.). New York, USA: McGraw- Hill.

Oliveira, C. P. D. (2004). Implementing task-based assessment in a TEFL environment. In L. L. Betty & R. W. Jane (Eds.), *Task-based instruction in foreign language education (*pp. 253-279*)*. Washington, D.C, USA: Georgetown University Press.

Ong, C. S., Lai, J. Y., & Wang, Y. S. (2004). Factors affecting engineers' acceptance of asynchronous e-learning systems in high-tech companies. *Information & Management, 41*(6), 795-804.

O'Reilly, T. (2007). What is Web 2.0: Design patterns and business models for the next generation of software. *Communications & Strategies*, 1, 17-27.

Oxford, R. L. (2006). Task-based language teaching and learning: An overview. *Asian EFL Journal*, 8(3), 94-121.

Pandian, A. (2002). English language teaching in Malaysia today. *Asia Pacific Journal of Education*, 22(2), 35-52.

Pang,S.G.S, & Heng, C.S. (1996) The use of English by Malaysian business executives in the commercial sector. *Pertanika J.Soc.Sci.& Hum* 4(2) 141-146.

Park, S. Y. (2009). An analysis of the technology acceptance model in understanding university students' behavioral intention to use e-learning. *Educational Technology & Society, 12*(3), 150-162.

Pessoa, S., Hendry, H., Donato, R., Tucker, G. R., & Lee, H. (2007). Content-Based Instruction in the Foreign Language Classroom: A Discourse Perspective. *Foreign Language Annals*, *40*(1), 102-121.

Peter, K. W. T. (2005). The medium of instruction debate in Malaysia: English as a Malaysian language? *Language Problems and Language Learning, 29*(1), 47-66.

Peterson, M. (2005). Learning interaction in an avatar and chat-based virtual world. *PacCALL, 1*(1), 29-40.

Peterson, M. (2006). Learner interaction management in an avatar and chat-based virtual world. *Computer Assisted Language Learning, 19*(1), 79-103.

Philips, D. C. (1995). The good, the bad, and the ugly: The many faces of constructivism. *Educational Researcher, 24*(7), 5-12.

Piaget, J. (1954). *The Construction of Reality in The Child*. New York, NY: Basic Books.

Piaget, J. (1980). The psychogenesis of knowledge and its epistemological significance. In M. Piattelli-Palmarini (Ed.), *Language and Learning*. Cambridge,MA: Harvard university press.

Pillay, H. (1998). *Issues in the teaching of English in Malaysia*. Retrieved from http://jalt-publications.org/old_tlt/files/98/nov/pillay.html

Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879-903.

Polites, G. L., & Karahanna, E. (2012). Shackled to the status quo: the inhibiting effects of incumbent system habit, switching costs, and inertia on new system acceptance. *MIS Quarterly*, *36*(1), 21-42.

Prasolova-Forland, E., & Divitini, M. (2003). *Collaborative virtual environments for supporting learnng communities: An experience of use.* Paper presented at the Proceedings of the International ACMSIGGROUP Conference on Supporting Group Work, Sanibel Island, Florida, USA.

Prator, C.H., & Celce-Murcia, M. (1979). An outline of language teaching approaches. In Celce-Murcia, M. and McIntosh, L. (Ed.), *Teaching English as a Second or Foreign Language*. New York: Newbury House.

Preshous, A. (2001). Where you going ah?: An account of the origin and development of Malaysian English. *English Today, 17*(1), 46-53.

Price, M. L. (1991). The subjective experience of foreign language anxiety: Interviews with highly anxious students. In E. K. Horwitz & D. J. Young (Eds.), *Language anxiety: From theory and research to classroom application* (pp. 101-108). Englewood Cliffs, New Jersey: Prentice-Hall.

Rankin, Y., Gold, R., & Gooch, B. (2006). *3D role-playing games as langauge learnng tools.* Paper presented at the Conference Proceedings of EUROGRAPHICS, Vienna, Austria.

Razak, N. A., & Malek, J. A. (2008). *Bridging digital divide in Malaysia:Cyber learnng for the marginalized community*. Paper presented at the Distance Learning and the Internet Conference, Waseda University, Tokyo, Japan.

Richards, K. (1989). Pride and prejudice: The relationship between ESP and training. *English for Specific Purposes, 8*(3), 207-222.

Ringle, C. M., Wende, S., & Will, A. (2005). *SmartPLS Release 2.0 (BETA).* Retrieved from http://www.smartpls.de

Roed, J. (2003). Language learner behaviour in a virtual environment. *Computer Assisted Language Learning, 16*(2), 155-172.

Rogers, E. M. (2003). *Diffusion of Innovations* (5 ed.). New York: Free Press.

Rosli, T., & Malachi, E. (1990). A comparative study of the achievement and the pfoficiency levels in English as a second language learners in selected rural and urban schools in Peninsular Malaysia. *The English Teacher, 19*, 48-57.

Rourke, L., Anderson, T., Garrision, D. R., & Archer, W. (2001). Methodological issues in the content analysis of computer conference transcripts. *International Journal of Artificial Intelligence in Education, 12*(1), 8-22.

Russell, T. L. (1999). *The no significant difference phenomenon: A comparative research annotated bibliography on technology for distance education: As reported in 355 research reports, summaries and papers.* Raleigh, NC: North Carolina State University.

Saeed, N., Yang, Y., & Sinnappan, S. (2008). *Media richness and user acceptance of Second Life.* Paper presented at the Proceedings of Ascilite 2008, Melbourne, Australia.

Saeed, N., Yang, Y., & Sinnappan, S. (2009). *User acceptance of Second Life: An extended TAM including hedonic consumption behaviours*. Paper presented at the 17th European conference on Information Systems, Verona, Italy.

Santos, O. A., & Ramos, F. (2004). Proposal of a framework for Internet based licensing of learning objects. *Computers & Education*, 42(3), 227-242.

Schibeci, R., Lake, D., Phillips, R., Lowe, K., Cummings, R., & Miller, E. (2008). Evaluating the use of learning objects in Australian and New Zealand schools. *Computers & Education*, 50(1), 271-283.

Schumaker, R. E., & Lomax, R. G. (2004). *A begineer's guide to structural equation modeling* (2 ed.). Mahwah, NJ: Lawrence Erlbaum.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26*(4), 207-231.

Seidlhofer, B. (2004). Research perspectives on teaching English as a lingua franca. *Annnual Review of Applied Linguistics, 24*, 209-239.

Sharma, R., Yetton, P., & Crawford, J. (2009). Estimating the effect of common method variance:The method-method pair technique with an illustration from TAM research. *MIS Quarterly, 33*(3), 473-490.

Shen, J., & Eder, L. B. (2009). Intentions to use virtual worlds for education. *Journal of Information Systems Education, 20*(2), 225-233.

Shi, L., Corcos, R., & Storey, A. (2001). Using student performance data to develop an English course for clinical training. *English for Specific Purposes, 20*(3), 267-291.

Shih, H. (2004). Extended technology acceptance model of internet utilization behaviour. *Information & Management, 41*(6), 719-729.

So, J. C. F., & Bolloju, N. (2005). Explaining the intentions to share and reuse knowledge in the context of IT service operations. *Journal of Knowledge Management, 9*(6), 30-41.

Spiegelman, M., & Glass, R. (2009). Games and Web 2.0: A winning combination for millenials. *Journal of Educational Technology Systems, 37*(3), 273-289.

Steinkuehler, C., & Williams, D. (2006). Where everybody knows your (screen) name: Online games as third places. *Journal of Computer-mediated Communication, 11*(4), 885-909.

Sternberg, R. J., Dietz-Uhler, B., & Leach, C. (1993). *The psychologist's companion.* Cambridge, UK: Cambridge University Press.

Stockwell, G. (2007). A review of technology choice for teaching language skills and areas in the CALL literature. *ReCall*, 19(2), 105-120.

Straub, D., Boudreau, M.-C., & Gefen, D. (2004). Validation guidelines for IS positivist research. *Communications of the Association for Information Systems, 13*(24), 380-427.

Stumpf, S. A., Brief, A. P., & Hartman, K. (1987). Self-efficacy expectations and coping with career-related events. *Journal of Vocational Behaviour, 13*(2), 91-108.

Suh, S., Kim, S. W., & Kim, N. J. (2010). Effectiveness of MMORPG-based instruction in elementary English education in Korea. *Journal of Computer Assisted Learning*, 26(5), 370-378.

Supyan, H. (2008). Creating a bigger Z.P.D. for ESL learners via online forum. *Journal of College Teaching and Learning, 4*(11), 1-9.

Sykes, J. M. (2005). Synchronous CMC and pragmatic development: Effects of oral and written chat. *CALICO Journal*, 22(3), 399-431.

Szajna, B. (1996). Empirical evaluation of the revised technology acceptance model. *Management Science, 42*(1), 85-92.

Tan, P. K. (2005). The medium-of-instruction debate in Malaysia: English as a Malaysian language? Language problems & language planning, 29(1), 47-66.

Tang, H., Mu-shang, Y., & Pan-ju, L. (2009, August). A meta-analytic review of current CALL research on second language learning. In *IT in Medicine & Education, 2009. ITIME'09. IEEE International Symposium on* (Vol. 1, pp. 677-684). IEEE.

Tenenhaus, M., Vinzi, V. E., Chatelin, Y.-M., & Lauro, C. (2005). PLS path modelling. *Computational Statistics & Data Analysis, 48*(1), 159-205.

Thang, S.M. (2004). Learning English in multicultural Malaysia: Are learners motivated? *Journal of Language and Learning, 2*(2), 142-153.

Thatcher, J. B., & Perrewe, P. L. (2002). An empirical examination of individual traits as antecedents to computer anxiety and computer self-efficacy. *MIS Quarterly, 26*(4), 381-396.

Thurlow, C., Lengeh, L., & Tomic, A. (2004). *Computer mediated communication: Social interaction and the Internet*. Thousand Oaks, CA: SAGE Publications Ltd.

Ting, S. H. (2003). Impact of language planning on language attitudes: A case study in Sarawak. *Journal of Multilingual and Multicultural Development, 24*(3), 195-210.

Torkzadeh, R., Pflughoeft, K., & Hall, L. (1999). Computer self-efficacy, training effectiveness, and user attitudes: An empirical study. *Behaviour & Information Technology, 18*(4), 299-309.

Underhill, N. (1987). *Testing spoken language: A handbook of oral testing techniques*. Great Britain: Cambridge University Press.

Vanderplank, R. (2010). Déjà vu? A decade of research on language laboratories, television and video in language learning. *Language Teaching*, 43(1), 1-37.

Venkatesh, V. (1999). Creation of favourable user perceptions: Exploring the role of intrinsic motivation. *MIS Quarterly, 23*(2), 239-260.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science, 46*(2), 186-204.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425-478.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, Massachusetts: Harvard University Press.

Wagner, D., Reitmayr, G., Mulloni, A., Drummond, T., & Schmalstieg, D. (2010). Real-time detection and tracking for augmented reality on mobile phones. *Visualization and Computer Graphics*, IEEE Transactions on, 16(3), 355-368.

Wang, C., Song, H., Xia, F., & Yan, Q. (2009). Integrating Second Life into an EFL program: Students' perspectives. *Journal of Educational Technology Development and Exchange, 2*(1), 1-16.

Wang, R., Wiesemes, R., & Gibbons, C. (2012). Developing digital fluency through ubiquitous mobile devices: Findings from a small-scale study. *Computers & Education*, 58(1), 570-578.

Wang, Y. S., Lin, H. H., & Luarn, P. (2006). Predicting consumer intention to use mobile service. *Information Systems Journal, 16*(2), 157-179.
Weathington, B. L., Cunningham, C. J. L., & Pittenger, D. J. (2010). *Research methods for the behavioral and social sciences*. Hoboken, New Jersey: John Wiley & Sons, Inc.
Webster, J., & Martocchio, J. J. (1993). Turning work into play: Implication for microcomputer software training. *Journal of Management, 19*(1), 127-146.
Weir, C. J. (1988). *Communicative language testing with special reference to English as a foreign language*. Exeter: University of Exeter.
Werts, C. E., Linn, R. L., & Joreskog, K. G. (1974). Interclass reliability estimates: Testing structural assumptions. *Educational and Psychological Measurement, 34*(1), 25-33.
White, C. (2003). *Language learning in distance education*. Cambridge University Press.
Whitman, M. E., & Woszczynski, A. B. (2004). *The handbook of information systems research*. USA: Idea Group Publishing.
Wiersma, W., & Jurs, S. G. (2009). *Research methods in education: An introduction* (9th ed.). USA: Pearson.
Wilson, B. G., & Cole, P. (1996). Cognitive teaching models. In D.H. Jonassen (Ed.), *Handbook of research in instructional technology* (pp. 601-621). New York: MacMillan.
Winn, W. D. (1993). *A conceptual basis for educational applications of virtual reality.* (No. R-93-9). Seattle,WA: University of Washington.
Winn, W. D., Hoffman, H., & Osberg, K. (1995). *Semiotics and the design of objects, actions and interactions in virtual environments.* Paper presented at the Proceedings of the Annual Meeting of the American Educational Research Association, San Francisco,CA.
Wojciechowski, R., & Cellary, W. (2013). Evaluation of learners' attitude toward learning in ARIES augmented reality environments. *Computers & Education.* Advance online publication. doi: 10.1016/j.compedu.2013.02.014
Wong, L. H., & Looi, C. K. (2011). What seams do we remove in mobile-assisted seamless learning? A critical review of the literature. *Computers & Education*, 57(4), 2364-2381.
Wrzesien, M., & Raya, M. A. (2010). Learnng in serious virtual worlds: Evaluation of learning effectiveness and appeal to students in the E-Junior project. *Computers & Education, 55*(1), 178-187.
Wu, H. K., Wen-Yu Lee, S., Chang, H. Y., & Liang, J. C. (2013). Current status, opportunities and challenges of augmented reality in education. *Computers & Education*. 62, 41–49.
Wu, W. H., Jim Wu, Y. C., Chen, C. Y., Kao, H. Y., Lin, C. H., & Huang, S. H. (2012). Review of trends from mobile learning studies: A meta-analysis. *Computers & Education*. 59(2), 817-827.
Yan, S. (2006). Improvement of EFL learners' speaking and writing through reflective scenario-based learning. *Sino-US English teaching, 3*(5), 23-27.
Yi, M. Y., & Hwang, Y. (2003). Predicting the use of web-based information systems: Self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International Journal of Human-computer Studies, 59*(4), 431-449.
Yin, R.K. (2009).*Case study research: Design and Methods.* Fourth Edition, Sage Publications.
Yong, F. L. (2010). A study on the cultural values, perceptual learning styles, and attitudes toward oracy skills of Malaysian tertiary students. *European Journal of Social Sciences, 13*(3), 479-493.
Yoo, Y., & Alavi, M. (2001). Media and group cohesion:Relative influences on social presence, task participation, and group consensus. *MIS Quarterly, 25*(3), 371-390.
Young, D. J. (1990). An investigation of students' perspectives on anxiety and speaking. *Foreign Language Annals, 23*(6), 539-553.

Zaaba, Z., Ramadan, F. I., Anning, I. N. A., Gunggut, H., & Umemoto, K. (2011). Language-in-education policy: A study of policy adjustment strategy in Malaysia. *International Journal of Education and Information Technologies*, *2*(5), 157-165.

Zikmund, W. G., Carr, J. C., & Griffin, M. (2012). *Business Research Methods*. Mason, Ohio: South-Western.

Zulkefli, I., Ainin, S., & Tengku, M. F. (2008). The roles of community based telecenters in bridging the digital divide in rural Malaysia. *World Academy of Science, Engineering and Technology, 46,*352-364.

Zuwati, H. (2006). Open and Distance learning: The effectiveness of Online Discussion Forums in promoting the use of English for communication. *AsiaCall Online Journal, 1*(1), 22-33.

Appendix A Information Sheet Distributed to Participants

Teaching business English in Multiuser Virtual Environments: Effectiveness and Acceptance

INFORMATION SHEET

Researcher(s) Introduction

This research is conducted by Harmi Izzuan Bin Baharum, a PhD student in Information Systems at School of Management of Massey University (New Zealand). His supervisor is Dr. Alexei Tretiakov (School of Management, Massey University, New Zealand).

Project Procedures

Over about two and a half hours, we will teach you business conversation skills in English - in the classroom and by using a computer. To compare the two teaching methods, we will digitally record audio during parts of the lesson.

On two occasions during the session, we will ask you to complete a simple questionnaire, asking about your attitudes towards your study of English and towards using computers.

We will not record your name.

We will store all data securely under lock and key and ensure that no-one apart from the researcher involved in the project can access it. After the research is completed, we will dispose of questionnaires by following standard procedures for treating confidential material, and delete all copies of digital audio recordings.

The results of the research will be published as part of a Ph.D. thesis, which can be obtained via library interloan from Massey University (New Zealand) library. Only summary statistical results will be included in the publication. If you wish to receive a digital copy of the thesis, please advise your email address to harmiizzuan@gmail.com.

Participant's Rights

You are under no obligation to accept this invitation. If you decide to participate, you have the right to:

- *withdraw from the study;*
- *ask any questions about the study at any time during participation;*
- *be given access to a summary of the project findings when it is concluded;*
- *ask for the audio recorder to be turned off at any time during the lesson.*

Completion and return of a questionnaire implies consent. You have the right to decline to answer any particular question.

This project has been evaluated by peer review and judged to be low risk. Consequently, it has not been reviewed by one of the University's Human Ethics Committees. The researcher(s) named above are responsible for the ethical conduct of this research.

If you have any concerns about the conduct of this research that you wish to raise with someone other than the researcher(s), please contact Professor Sylvia Rumball, Assistant to the Vice-Chancellor (Ethics & Equity), telephone 06 350 5249, email humanethics@massey.ac.nz".

Appendix B Participant Consent Form

Teaching business English in Multiuser Virtual Environments: Effectiveness and Acceptance

PARTICIPANT CONSENT FORM

This consent form will be held for a period of five (5) years.

I have read the Information Sheet and have had the details of the study explained to me. My questions have been answered to my satisfaction, and I understand that I may ask further questions at any time.

I agree to parts of the lesson being recorded by using a digital audio recorder.

I agree to participate in this study under the conditions set out in the Information Sheet.

Signature: ..... **Date:** .....

Full Name - printed .....

Appendix C Learning Scenario

Pair up and participate

Scenario

You are a real-estate agent of Company A. You have received a phone call from a potential buyer who is interested in having a look at the house being advertised in the newspaper last week. The advertisement is as follows:

House for Sale

Jalan Semarak. Nice neighbourhood. Close to town and facilities. Large comfortable house. Fully furnished.

Tel : 03-3371122

There are two roles in the scenario which are a real estate agent and a house buyer. Student A will take up the role of a real-estate agent and student B will take up the role of a potential buyer. As the real-estate agent, you are going to do a short presentation about the house and answer questions (if necessary) from the potential buyers regarding the house being shown. As the potential buyer, you are going to listen to the demonstration and praise about the condition of the house. You may ask questions if necessary. Once the task is completed, the student will switch role.

Practice this dialogue

Dialogue	
House Sales Agent	Good Morning Mr Farid
Mr Farid	Hi, Good Morning
House Sales Agent	Thank you for coming to look at the house. My name is Tan and this is my business card.
Mr Farid	Thank you.
House Sales Agent	Let me give you a tour of the house. It has a nice big area. The house is surrounded by a few trees and it is very cool in the hot afternoon.
Mr Farid	Yes, I can see that. It is also nice for my children to play.

House Sales Agent	Let us go inside now. This is the living-room.
Mr Farid	Very nice
House Sales Agent	Over here, we have three large bedrooms
Mr Farid	The bedrooms are comfortable.
House Sales Agent	This is the kitchen , the dining-room and the bathroom
Mr Farid	Fantastic
House Sales Agent	Okay, that's the end of my tour of the house. Please call me if you are interested with the house.
Mr Farid	Thank you and I will call you soon.

Appendix D Proficiency Rating Scale

Student Number	

		4	3	2	1
	Organising ideas	Presents information on nearly everything and replies appropriately, with little need for adjustment or repetition	Presents information fairly well and usually replies appropriately, needing some repetition and/or adjustment	Presents little information and needs constant adjustment and repetitions	Presents very little information ; needs constant adjustment and repetitions
1	Introductions				
2	Presenting the house from outside				
3	Presenting the house from inside				
4	Parting				

		4	3	2	1
	Vocabulary	Produces the vocabulary required to carry out the task successfully	Sometimes lacks the right words, but manages to accomplish the task reasonably well	Lacks many of the right words, but manages to accomplish the task	Lacks most of the vocabulary required, which seriously harms his/her performance in the task
5	Names of rooms				
6	Price of the house				
7	Nature of room (large, small etc.)				
8	Size of the house				

		4	3	2	1
		Speech shows very few problems with word order and agreement	Speech shows few problems with word order and agreement	Speech shows several problems with word order and agreement	Speech shows constant problems with word order and agreement
9	*Accuracy*				

		4	3	2	1
		Student hardly ever halts or hesitates	Student is occasionally hesitant but manages to accomplish the task	Student frequently halts and hesitates, compromising the accomplishment of the task	Student halts and hesitates, seriously compromising of the task.
10	*Fluency*				

Appendix E Survey Questionnaire Part A – Statements on Video Games Background

Directions

This questionnaire contains statements your Video Games background. There are no 'right' or 'wrong' answers. Your opinion is what is required.

Be sure to give an answer for all statements. If you change your mind about an answer just cross it out and circle another.

Video Games Self-efficacy

Please circle numbers in the following statements based on the following Likert scale

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	I am confident of using video games if there was no one around to tell me what to do as I go	1	2	3	4	5
2	I am confident of using video games if I had never used a package like it before	1	2	3	4	5
3	I am confident of using video games if I had only the software manuals for reference	1	2	3	4	5
4	I am confident of using video games if I had seen someone else using it before trying it myself	1	2	3	4	5
5	I am confident of using video games if I could call someone for help if I got stuck	1	2	3	4	5
6	I am confident of using video games if someone else had helped me get started	1	2	3	4	5
7	I am confident of using video games if I had just the built-in help facility for assistance	1	2	3	4	5
8	I am confident of using video games if someone showed me how to do it first	1	2	3	4	5
9	I am confident of using video games if I had used similar packages before this one to do the same job	1	2	3	4	5

Video Games Anxiety

Please circle numbers in the following statements based on the following Likert scale

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	Given the opportunity to use video game, I am afraid that I might have trouble in navigating through it.	1	2	3	4	5
2	I hesitate to use video game in case I look stupid.	1	2	3	4	5
3	I feel uneasy about using video game.	1	2	3	4	5
4	Playing with video game scare me a lot.	1	2	3	4	5
5	I hesitate to use video game as I'm afraid of making mistakes I can't correct.	1	2	3	4	5
6	Video game makes me feel uncomfortable.	1	2	3	4	5

Video Games Affect

Please circle numbers in the following statements based on the following Likert scale

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	I like working with video games.	1	2	3	4	5
2	I look forward to those aspects of my task job that require me to use video games.	1	2	3	4	5
3	Once I start working on the video games, I find it hard to stop.	1	2	3	4	5
4	Working with video games is fun.	1	2	3	4	5
5	Video games make work more interesting.	1	2	3	4	5
6	I don't feel apprehensive about using video games.	1	2	3	4	5

Appendix F Survey Questionnaire Part B – Biodata and Questions on Your English Learning Background and Technology Acceptance

Directions

This questionnaire contains statements about your English learning background and technology acceptance. There are no 'right' or 'wrong' answers. Your opinion is what is required.

Be sure to give an answer for all statements. If you change your mind about an answer just cross it out and circle another

Student's Biodata

1. Gender Male ☐ Female ☐

2. Faculty _____

3. Age 18-20 ☐ 21-23 ☐ 24-26 ☐

4. Experience (number of months or years) using Second Life.

Put 0 if no experience _____

Attitudes Towards Learning English

Please circle the following statements based on the Likert scale below

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	English language is really great.	1	2	3	4	5
2	I really enjoy learning English.	1	2	3	4	5
3	I love learning English.	1	2	3	4	5
4	I plan to learn as much English as possible.	1	2	3	4	5
5	I hate English.	1	2	3	4	5
6	I would rather spend my time on courses other than English.	1	2	3	4	5
7	I find the study of English very boring.	1	2	3	4	5
8	Learning English is a waste of time.	1	2	3	4	5
9	When I finish this course, I shall give up the study of English entirely because I am not interested in it.	1	2	3	4	5

Subject Matter Self-efficacy

Please circle the following statements based on the Likert scale below

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	I'm sure I could speak English well in almost any circumstances.	1	2	3	4	5
2	When the English language is spoken to me, I feel I can understand practically everything.	1	2	3	4	5
3	I feel comfortable conducting myself in English almost any time and any place.	1	2	3	4	5
4	I believe that I can competently read and understand most books and articles written in English.	1	2	3	4	5
5	I may not be completely fluent in English, but I feel confident speaking it.	1	2	3	4	5
6	Despite the fact that I may not be completely proficient in English, I am self-assured conducting myself in English.	1	2	3	4	5
7	Even when I make mistakes speaking English, I still feel sure of myself while trying to communicate.	1	2	3	4	5
8	I am confident when having conversation with English-speaking people despite any errors I may make.	1	2	3	4	5
9	Regardless of how much English I know, I feel confident about using it.	1	2	3	4	5
10	I feel confident using English regardless of my ability.	1	2	3	4	5

English Class Anxiety

Please circle the following statements based on the Likert scale below

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	I never feel quite sure of myself when I am speaking in our English class.	1	2	3	4	5
2	It embarrasses me to volunteer answers in our English class.	1	2	3	4	5
3	It worries me that other students in my class seem to speak English better than I do.	1	2	3	4	5
4	I get nervous and confused when I am speaking in my English class.	1	2	3	4	5
5	I am sometimes afraid the other students will laugh at me when I speak English.	1	2	3	4	5
6	I don't usually get anxious when I have to respond to a question in my English class.	1	2	3	4	5
7	I feel confident when asked to participate in my English class.	1	2	3	4	5
8	I do not get anxious when I am asked for information in my English class.	1	2	3	4	5
9	I don't understand why other students feel nervous about using English in class.	1	2	3	4	5
10	Students who claim they get nervous in English class are just making excuses.	1	2	3	4	5

Desire to Learn English

Please circle the following statements based on the Likert scale below

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	I wish I had begun studying English at an early age.	1	2	3	4	5
2	If it were up to me, I would spend all my time learning English.	1	2	3	4	5
3	I want to learn English so well that it will become second nature to me.	1	2	3	4	5
4	I would like to learn as much English as possible.	1	2	3	4	5
5	I wish I were fluent in English.	1	2	3	4	5
6	Knowing English isn't really an important goal in my life.	1	2	3	4	5
7	I find I'm losing any desire I ever had to know English.	1	2	3	4	5
8	To be honest, I really have little desire to learn English.	1	2	3	4	5
9	I haven't any great wish to learn more than the basics of English.	1	2	3	4	5

Perceived Usefulness

Please circle the following statements based on the Likert scale below

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	Using interactive visual computer environment improves my learning efficiency.	1	2	3	4	5
2	Using interactive visual computer environment will increase my learning performance.	1	2	3	4	5
3	Using interactive visual computer environment increases my learning output.	1	2	3	4	5
4	Using interactive visual computer environment is useful for my learning.	1	2	3	4	5
5	Using interactive visual computer environment will allow me to accomplish learning tasks more quickly.	1	2	3	4	5
6	Using interactive visual computer will make it easier to learn course content.	1	2	3	4	5

Perceived Ease of Use

Please circle the following statements based on the Likert scale below

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	My intention with interactive visual computer environment is clear and understandable.	1	2	3	4	5
2	Interacting using interactive visual computer environment does not require a lot of mental effort.	1	2	3	4	5
3	I find visual computer environment easy to use.	1	2	3	4	5
4	I found it easy to get interactive visual computer environment to do what I want it to do.	1	2	3	4	5
5	Learning to operate interactive visual computer environment is easy for me.	1	2	3	4	5
6	It is easy for me to become skilful at using the interactive visual computer environment.	1	2	3	4	5

Intention to Use

Please circle the following statements based on the Likert scale below

No.	Statements	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	I intend to use interactive visual computer environment as often as needed.	1	2	3	4	5
2	Whenever possible, I intend to use visual computer environment for learning.	1	2	3	4	5
3	I plan to use visual computer environment in the future.	1	2	3	4	5
4	I intend to continue using visual computer environment in the future.	1	2	3	4	5
5	I expect my use of visual computer environment to continue in the future.	1	2	3	4	5

Thank you for your time and cooperation.