British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

EXPLORING EFL LEARNERS' ENGAGEMENT IN VIRTUAL REALITY ENVIRONMENTS: A CASE STUDY ON KING ABDUL-AZIZ UNIVERSITY (KAU) LEARNERS IN KSA

Ehteram Ahmed Saleh Sharyan

ABSTRACT: *Virtual reality (VR) is used for many applications, ranging from medicine to space and from entertainment to training. In this research paper, VR is applied in EFL education in order to explore the effect of virtual reality environments (VR) on the engagement of female students of English as a Foreign Language (EFL) at King Abdul-Aziz University (KAU) in Saudi Arabia. A mixed-research approach was taken to examine the correlation between three major variables (students' engagement, students' cognitive load, students' behaviour in VR). The quantitative research made use of three research instruments: 1) a pre/post self-perceived engagement questionnaire, 2) the NASA TLX (Task Load Index), which was used to calculate the learners' cognitive load, and 3) a Students' Behaviour in VR rubric, which was used to analyse learners' video-recorded interactions in VR. I used the qualitative interview tool to collect data from each participant after the experiment in order to analyse students' perceptions and subjective assessment of their VR experience. The study includes (n=10) EFL female students from the Department of European Languages at King Abdul-Aziz University. I used various statistical treatments from IBM SPSS (Statistical Package for the Social Sciences) to investigate any significant differences or correlations between the three major variables in the quantitative data. The findings indicated that the level of engagement among students was significantly improved when using VR, while the results did not reveal any significant correlations between students' cognitive load and behaviour in VR that could affect their engagement.*

KEYWORDS: EFL, Learners' Engagement, Virtual Reality, Environments, King Abdul-Aziz University, KSA

INTRODUCTION

Since the 1960s, 'Virtual reality' (VR) is a term used to describe many technologies and software such as the Sensorama Simulator (Heilig, 1962) as well as virtual worlds such as Second Life, a three-dimensional online virtual world occupied by avatars representing real human beings who communicate with each other through voice and text chat (Jensen & Konradsen, 2018). Virtual technologies are 'at the forefront of technological developments' (Lange et al., 2010, p. 479), which can be attributed to their rapid technological advancement and universal interest. In recent years virtual reality (VR) and augmented reality (AR) systems have become a great investment, totaling a value of $15.9 billion by 2019; while in 2016 Digi-Capital predicted that this will rise to $120 billion by 2020 (Digi-Capital, 2016 cited in in Martín-Gutiérrez et al., 2017, p. 470). Additionally, about $800 million was invested in the first quarter in 2016 by Magic Leap, one of the main players in the VR/AR industry (alongside Oculus, Blippar, and Mindmaze) (cited in Martín-Gutiérrez et al., 2017, p. 470). Consequently, advancements in virtual technologies could positively contribute to the learning/teaching process. In general, the spread of digital technology and its integration into different aspects of our daily life has contributed to changing our perspectives towards several domains, one of which is education (Lange et al., 2010. p. 340). VR is considered one of the most modern and expanding technologies because it is a very close representation of reality, with the use of a Head-Mounted Display (HMD), glasses, gloves, sensors and other devices (Ibanez et al, 2013). It has been used for many years with flight simulators, but now VR is cheap, globally available for public use and portable. The more ways VR can be used, the more advanced it will become.

Although VR applications are well known in the field of entertainment (Martín-Gutiérrez et al., 2017), they are now used extensively for educational purposes (Jensen and Konradsen, 2018, p. 1516; Peixoto et al., 2019), but they are still limited within the classroom (Jensen and Konradsen, p. 1516) which might be attributed to the expensive systems that are required. However, as large company profits expand (e.g. Google, Apple, Microsoft, and Samsung), VR has become more readily available at a cheaper price, and accessible for the wider public, benefiting many fields, including education (Bonner, 2018, p. 36). In other words, the attention that VR received from popular companies along with media have encouraged the use of VR in different domains, education in particular (Makransky and Lilleholt, 2018). Studies have revealed that VR environments can foster students' motivation and engagement in many subjects such as medicine and space travel (Makransky and Lilleholt, 2018). This might extend the breadth of learning and streamline teaching to match technological advances (Freina and Ott, 2015).

As student expectations have evolved in recent years and instructors are increasingly under pressure to present their teaching in novel ways, VR has been deemed to be the next revolutionary phase in education (Freina and Ott, 2015). As early as the 1990s, scholars predicted VR to be at the forefront in both training and education. Regian and Shebilske argued that the move to more experiential learning would be progressive and beneficial (Freina and Ott, 2015). They also found in their study that students using VR found the experience more engaging and performed better than peers being taught in the more traditional and standard way. A question to ponder is how can VR be manipulated and utilized to suit learning in individual subjects?

Many different concepts of VR can aid in the education system. Virtual universities offer services to students as if they were physically on-site at the university. Google Cardboard headsets are very popular and cheap, allowing students to go on virtual field trips using 360-degree videos to fully immerse themselves in the environment. This approach can offer an excellent alternative to training in the traditional style, thus eliminating potential dangers. This may include scientific experiments or experiences that could be life-threatening. VR is also useful in exploring technical information with graphics and following a step-by-step procedure in construction or experiments. Numerous case studies have already been done exploring this field (Ibañez and Kloos, 2013).

As VR becomes more advanced and popular, it is also becoming more affordable. Due to its positive impact on education in general, VR has been utilized for language learning. This is not only to enhance language acquisition but also to better motivate and inspire English-language students' extra-curricular activities in Spanish and to help elementary children better empathize with classroom topics (Bonner et al., 2018). However, so far the use of VR in language settings has been limited (Peixoto et al., 2019 p. 582). Therefore, this study will explore VR in a limited environment, and I will explore the wealth of opportunities that VR can offer education through learning and mere participation.

The ease and interest in using VR might be the main motivation for its educational potential. The sense that you are actually immersed in the environment rather than merely being an observer could enhance the learning experience. Thus virtual technologies make students feel more committed and motivated, opening up new paths for teaching and learning (Aieillo et al., 2012). Another reason to use VR in classroom settings is its ability to reduce distractions. Immersion not only limits distractions but also makes learners more aware of the environment. This also has helped students to make real world connections between subject matters and their own experiences. This experience can be applied to any topic students are studying and the global experience (Bonner et al., 2018). Learners are no longer limited by their imagination, they are now able to see, witness and experience things beyond their imagination.

Furthermore, VR could give students the opportunity to live and experience situations that would otherwise be impossible in reality. According to Freina and Ott (2015), there are many factors that might limit students' experiences such as time (allowing for the experience of historical periods), place (space exploration), danger (flight simulation), and ethics (contentious surgery). Enabling learners to virtually experience such limited opportunities will bring subjects to life in terms of imagination and motivation. Periods of history can be re-lived and ground-breaking surgery can be carried out without risk (Freina and Ott, 2015).

In this research paper, VR is applied to EFL education in order to explore the level of engagement among EFL learners at KAU. The objective of the study was to examine the impact of the VR environment on the students' engagement in EFL classrooms. Engagement, motivation and enhanced learning are all linked within the classroom and this study will explore these links in light of adopting VR as an educational vehicle and tool.

As stated previously, VR is not yet widely used in many modes and areas of education, language learning in particular (Peixoto et al., 2019, p. 582). The impact of VR on language acquisition and its use in the language learning environment is of specific interest in this study. The main objective of the current research is to investigate how female Saudi EFL students' engagement in VR experiences might have been influenced by their cognitive load and how their levels of achievement relate to their behaviour and interactions inside the virtual reality world. This objective was broken down into three main questions:

1. What is the effect of using VR environments on learners' engagement in EFL learning at KAU?
2. Is there any correlation between students' cognitive load and engagement in virtual environments?
3. Is there any correlation between students' behaviour (interaction) and engagement in virtual reality environments?

This paper will be divided into five sections. The first section will review some of the previous studies related to the topic. The second section explains the methodological approach used in this study. The findings will be discussed in the third section, followed by a detailed discussion relating to the research questions. The final part presents a conclusion and the recommendations of this research paper.

LITERATURE REVIEW

Though the main focus of virtual reality (VR) is in the field of entertainment (Martín-Gutiérrez et al., 2017), it has recently become popular in the educational sector (Standen and Brown, 2006; Curcio et al., 2016; Peterson and Stone, 2019). One possible reason might be its ability to effectively break the limitations in formal education as well as its efficacy in making students more immersed in real-life learning experiences (J.O and Cortiz 2017 cited in Al-Azawi et al., 2019). There have been numerous studies that investigated utilizing VR in the field of education, in formal education in particular (Curcio et al., 2016; Freina and Ott, 2015). In this section, my literature review will be divided into four sub-sections. The first will set out a theoretical framework defining major concepts in this study as well as citing relevant educational research on them. The second is devoted to studies investigating the use of VR in higher education programs. The third section aims to cite and comment on research that has examined the relation between engagement and the use of VR in language instruction. Due to the scarcity of research investigating the relation between cognitive load and virtual reality in the context of language instruction, I devoted a fourth sub-section to exploring studies that investigate engagement and cognitive load in language instruction that utilized technology.

Theoretical framework

Defining virtual reality (VR)

Phillip Barker describes a VR system as a sophisticated multimedia environment in which users are exposed to, and can participate in, surrogate tacto-audio-visual experiences. These experiences are created by means of a computer system to which are attached special types of peripheral device, enabling users to interact with the real and artificial objects that exist within the interaction environment. (Barker, 1993, p. 16).

According to Hussein and Nätterdal, VR is a collection of hardware which includes PCs, mobile phones, head-mounted displays (HMDs), tracking sensors and software that give an assimilative experience (Hussein and Nätterdal, 2015). Similarly, George Coates states that virtual reality enables the customer to experience environmental electric simulations by using head-mounted eye goggles and wired clothing (cited in Steuer, 1992).

In the online Oxford dictionary, virtual reality (VR) is defined as 'images created by a computer that appear to surround the person looking at them and seem almost real.'

Defining student engagement

Since the mid-1990s, 'student engagement' has attracted extensive attention in the literature (Trowler, 2010, p. 2). It has been widely researched due to its an essential relation to students' learning outcomes (Kahu, 2013). In education, student engagement refers to the level of attention, curiosity, interest, optimism, and passion that they show in the process of education. This includes the students' depth of motivation towards their academic development.

Based on definitions provided by previous literature reviews, Trowler 2010 defined 'student engagement' as follows:

the interaction between the time, effort and other relevant resources invested by both students and their institutions intended to optimise the student experience and enhance the learning outcomes and development of students and the performance, and reputation of the institution.

Many researchers define student engagement as a broad construct that consists of three main elements: cognitive, emotional and behavioural engagement (Fredricks et al., 2004). Behavioural engagement denotes the extent to which a student complies with what is expected from them inside a classroom such as attendance, doing assignments, and participating (Cooper, 2014, p. 365). Emotional engagement indicates students' involvement with learning. For example, when students are emotionally engaged, they feel positively about their class and this will be reflected by their participation, enjoyment, interest and willingness to do well (Fredricks et al., 2004). Students are cognitively engaged when they make cognitive investments in learning by applying mental efforts, seeking hard tasks and challenges and finding out new things (Cooper, 2014, p. 365).

Defining cognitive load in educational research

Cognitive Load Theory (CLT) was developed by John Sweller. He published a paper on the subject in the journal *Cognitive Science* in 1988. According to him, 'cognitive load' is linked to the amount of information that working memory can contain at a given time. In other words, working memory can only maintain a limited amount, and instructional methods should avoid overloading it with extra activities that don't directly relate to learning (Sweller, 1988).

Gerjets et al (2009) agree that 'cognitive load theory is based on a number of widely accepted theories about how human brains process and store information'. These assumptions include: that human memory can be divided into working memory and long-term memory; that information is stored in the long-term memory in the form of schemas; and that processing new information results in a 'cognitive load' on working memory that can affect learning outcomes (Anderson 1977; Atkinson and Shiffrin 1968; Baddeley 1983).

Virtual reality (VR) applications as instructional tools in higher education settings

Generally speaking, the use of VR in teaching and learning scenarios had been the focus of many studies that consider the application of VR in applied fields or the humanities. In the field of medicine, for example, Huang et al. (2010) described two case studies wherein medical students used detailed VR applications to better understand human anatomy. The first case study found that student feedback concerning VR, gathered by questionnaires, suggested the feeling of immersion being the most important motivator for students to use VR – with other factors such as intuitive interactions and the sense of physical imagination ranking lower in importance. The second case study investigated VR's possible effects on student collaboration and found that collaborative learning possibilities were the intended use of VR for 59% of students involved in the second study.

In the field of finance, a study conducted by Lee et al. (2017) demonstrates the high potential of VR in increasing students' engagement and enhancing their learning experience. The authors used Google Cardboard VR (involving 3-dimensional images and 360-degree surroundings) in business classrooms and compared the content delivery platforms to a flat screen display. The participants were 44 undergraduate and graduate students from the business school of a large Canadian university. Participants watched an educational video about the landscapes of Nepal, with some watching via a head-mounted VR device and others viewing a flat screen display. Though participants rated the two formats similarly in terms of reliability and understandability, they gave the VR format higher ratings in term of their enjoyment and interest.

The application of VR in education was investigated in different case studies by Curcio et al. (2016). Based on their analysis, they concluded that there was evidence that VR can increase students' motivation and engagement and support students' critical thinking.

Alhalabi (2016) confirmed that VREs can lead to positive cognitive outcomes in many educational settings, one of which is engineering. He emphasizes that the more students are involved in the learning process, the more their learning outcomes improve. Based on the experiment conducted by Alhalabi (2016), a comparison was conducted between a traditional teaching design and three immersive learning environments (a Corner CAVE System (CCS), an HMD system with tracking, and an HMD system without tracking). The study examines the participants' learning performance in those different learning systems through quizzes on four different topics (astronomy, transportation, networking and inventors). The study includes 48 participants in which 12 participants were assigned to each specific group (4 groups in total). The results of his study indicated that high-immersion VREs could significantly lead engineering students to more positive learning outcomes and revealed that students prefer learning in virtual environments to non-virtual ones.

Repetto et al. (2015) have shown that it was possible for VR applications to improve language learning outcomes via embodied cognition. The results of their study on 24 participants suggested that performing a virtual movement with a limb (i.e., virtual run — performed with the legs/feet) could speed up the understanding of verbs that describe actions performed with the same limb (to kick, performed with legs/feet).

Peixoto et al. (2019) investigated teachers' perceptions of VR as an educational tool that could motivate students for language learning better than audio-only listening exercises. The data collection includes three questionnaires presented in the participants' first language. The first one is a sociodemographic questionnaire. The second one is a 6-item 7-point Likert scale questionnaire which aims to assess each participants' satisfaction with the VR experience. Finally, a 4-item Likert scale questionnaire focused on the participants' satisfaction with the VR as an educational tool. Based on their pilot study, the researchers found that the participants, who were seven Portuguese foreign language teachers for whom English is a second language, highly appreciate VR application, as they thought it attractive in terms of its potential to intensify students' learning and motivate them. A quasi-experimental design was chosen along with a cross-sectional study with a quantitative focus.

Virtual reality (VR) and students' engagement in language instruction

There have been numerous research projects investigating the implications of VR use in academic settings and its impact on learners' engagement. Most of this research has considered specializations which lend themselves easily to applied studies like medicine, workplace safety, science, business, etc., as VR provides access to tools and applications that encourage hands-on practice. These studies found that the VR treatment impact on learners' engagement was positive (Huang et al, 2010; Madathil et al., 2016; Lee et al., 2017; Allocoat and Muhlenen, 2018; Filer and Holmes, 2018).

Research focusing on the use of VR and language instruction has highlighted numerous advantages of using VR as a learning/teaching method compared to traditional language learning strategies (Wilang and Soermphongsuwat, 2018). Previous educational research has suggested that implementing a variety of learning methods is very important. Gaytan and McEwen (2007) concluded that it is beneficial to use a variety of instructional methods to appeal to students' learning preferences. VR is considered '…incredibly helpful not only to bring language learners closer to the language culture but also to create realistic simulations that wouldn't even exist in the real world…' (Schwienhorst, 2002). Wilang and Soermphongsuwat (2018) however pointed out that the 'familiarity of the tool may also let the participants more focused on the language task.' In most of these studies, engagement was continuously investigated, indicating the significance of this variable in language instruction.

Integrating new technologies such as VR applications and tools into language learning could be useful for learners, as they increase motivation and engagement. Wilang and Soermphongsuwat (2018), for example, conducted a pilot study that investigated the impact of using a VR headset on language learner's engagement, motivation, and independent learning as well as on their ability to learn vocabulary and to follow instructions. The study sample includes eleven freshmen students registered in an introductory English course. Participants were asked to verbally identify vocabulary cards in English while wearing a VR headset, and data was collected through individual interviews, observation sheets to record the participants' reactions during the experiment, and finally a 5-point Likert-scale survey consisting of three items aiming to measure students' engagement, motivation, and independent learning. The findings of this research emphasized the benefits of using a VR headset in the language class, as the study reported positive scores for learners' engagement, motivation and independent learning. Moreover, students' scores in the vocabulary post-test were better than those of the pre-test.

Similarly, Gorham et al. (2019) conducted a pilot study in which three participants learned how to write Japanese kanji characters within the immersive environment of the Kingspray Graffiti Simulator on the Oculus Rift VR system. The study's results did not support one of the study's hypotheses: that learners using VR would be able to reproduce the script with more accuracy than those using pen-and-paper learning methods. However, the study's results did support their second hypothesis: that an experimental group would be more excited and engaged using VR to learn a new script than using pen and paper, though the inexperience of the participants with VR might have generated some of this excitement. The VR participants also expressed more discomfort and asked more questions, the study notes.

Chen et al. (2019) investigated the impact of using Google Earth VR on the writing skill and engagement of 22 English language students at a middle school in the Midwestern US. The participants were all non-native. The data was collected through a sequential explanatory design. The researchers documented a statistically significant increase in the students' expository writing skills, particularly in description, cause/effect, compare/contrast, and enumeration after incorporating Google Earth Virtual Reality (GEVR) into the English learners' expository writing experiences. The most common positive factors were attributed to the students' active engagement. The observed benefits of using Google Earth VR included enjoyment of the virtual learning experience, ease of use/control, lack of cost, provision of resourceful information, and improvement of scientific inquiry skills.

Cognitive load, engagement and the use of VR

The cognitive load construct or variable has been investigated in language instruction using technology. Moreover, there has been research on the relation between cognitive load and VR use in educational settings where other subjects were taught and engagement was considered as a variable. However, studying cognitive load in the dual context of virtual reality as an instructional technology and language learning has rarely been considered. The tools used to explore cognitive load changes varied from self-perceived instruments (questionnaires or interviews) to electroencephalography (EGG).

Research examining the link between learners' cognitive load and language instruction using technology demonstrated that the learner's cognitive load did not impact their performance or learning and decreased with the use of the chosen technology (Bernal, 2014; Yu et al. 2014; Wang and Wu, 2014; Lee, 2014).

Studies which examined the relation between cognitive load and VR use in academic settings where other specializations were taught demonstrated that there was no significant correlation between cognitive load and VR use. Moreover, they reported the decrease of cognitive load while using VR (Stone et al., 2011; Jiang and Laidlaw, n.d.). The relationship between engagement and cognitive load has not been confirmed in these studies either.

A study conducted by Jardina and Chaparro (2013) aims to understand the compatibility of two e-textbooks in regards to satisfaction, perceived workload, engagement and understanding of two separate e-textbook applications, namely Kindle and Inkling. This study was limited, as it only looked at the usability in terms of leisure reading and not as an e-textbook, and moreover it was conducted in a lab and not in real-world environment. It indicates that there were no important differences on the perceived workload, measured by scales (NASA-TLX), between the two-separate e-Textbook applications Kindle and Inkling. They found that the average workload scores were relatively low, indicating a low workload on both applications. Overall participants were happy using the e-textbook applications for the purpose of study and could understand the relevant material.

Men and women differ on many levels, biologically, physically and psychologically. Vail et al. (2015) proposed that the genders could possibly benefit differently from adaptive support. The purpose of this paper was to report on a study, which was conducted to compare four different types of intelligent tutoring systems to introduce computer programming, catering to different levels of cognitive and affective support to gain full potential benefits. The study used four versions of a Java tutor, the baseline version administering a problem-solving environment and the other three supplying progressive learning tasks with extra adaptive support. According to the results of this study, both genders gained the same amount from the cognitive support but the female students became a lot more frustrated and less motivated than their male counterparts when there was not affective support. The outcomes were determined by comparing gender and adaptive support conditions (Vail et al., 2015).

Mahendra et al. (2018) look at how individual and collaborative teams of web-based reviewers influence and justify the use of review structures, in getting better open-ended feedback in large classrooms, which is very important in learning to encourage motivation and engagement of students, thus enabling easier management. They compare positive outcomes of web-based team reviews and traditional individual peer reviews, as it was a topic they felt had not been explored. The main hypothesis is supported by the early work of Zhu et al. (2014). Although the study was conducted with authentic peer reviews, the quality was not controlled. The outcomes of this study could offer a basis for educators and researchers to train newcomers in gaining more positive, effective feedback, via structured peer reviews. The study indicated that students preferred collaborative web-based peer review teams whilst conducting the reviews, as they found it to be more engaging and believed it to be more rewarding.

However, other researchers have established a link between learners' cognitive load and engagement in educational or training settings that required the use of technology.

A study by Cheng-Hsuan Lan et al. (2019) looked at the differences between two different methods of collaborative learning, face-to-face vs. online. Although collaborative technologies are becoming quite significant, not many studies have looked closely at the effectiveness of f2f and online collaboration separately. According to the results of this study, visual learners seemed to have higher levels of engagement and less cognitive load with online collaborative learning compared to linguistic learners, who did not show significant differences.

Tan et al. (2014) looked in detail at relationships between mental workload and stress during collaborative work from both the instructor's and the workers' view. The method used for the study was interesting. The workers were not given instructions to build K'Nex models and had to rely on the instructor's directions. The results indicated that using biofeedback for video-mediated collaboration was quite significantly effective in lowering mental workload for both instructors and workers alike. According to Tan et al. (2014), these findings agree with research from other studies such as Knight et al. (2000, 142-150) and Bowman and Stern (1995, 294-303).

The above-mentioned articles have given me different perspectives on various aspects of my research. Vail et al's study.had some similarities, in the sense that I was conducting my study with female participants, whereas they were looking at the outcomes based on gender. The other studies were similar as they had technology as one of the focal points in relation to human attributes. They differed in techniques and had very different aims but mostly focused on the outcomes of collaboration of human abilities and technology.

METHODOLOGY

For the sake of answering the research questions (section 1) and exploring the significance of possible correlations between the research variables, a pre-experimental, on-group pre/post-test research design was adopted. For data collection, a mixed-research approach was implemented making use of quantitative and qualitative research techniques and tools. Three major variables were explored: students' engagement, students' cognitive load in VR, and students' behaviour in VR. Moreover, correlations were made between the major variables and (5) general variables: academic level, ICT skill, prior use of VR, prior use of virtual learning environments in KAU, and preferred learning style. This section deals with the characteristics of the participants.

Participants

A purposive sampling technique was adopted for the study, because the participants had to be mainly from the Department of European Languages at KAU in Saudi Arabia, and from a limited number of students taking summer classes in 2019. Therefore, ten participants (*n*=10) were chosen for this research study investigating Saudi EFL learners' engagement in virtual reality (VR) educational settings and how this engagement might be affected by their cognitive load and behaviour inside VR environments. For the purpose of knowing more about the participants, a learner profile short survey was conducted. Based on the learners' profile results, 50% of the participants have an intermediate level of ICT whereas 50% are considered advanced. In terms of having previous experience in using virtual reality applications, only 10% of them used the VR headset before, while 100% had an experience with virtual environments in general. The academic level of the participants, according to KAU system, was as follows: 5 were in level 8, 1 was in level 5, 1 was in level 6, and 3 were in level 7. The students' self-perceived evaluation of their learning styles revealed that 60% were visual learners, 30% were reading/writing learners, and 10% were auditory learners.

Experiment set-up

The Virtual Reality Experience, and its automated recording process, were run on a 44 LCD computer monitor with a Windows operating system (8.1) (Fig.1). The Oculus Rift headset, four (room-scale) sensors, and two touch controllers were attached to the computer (Fig. 2). The experience was run in a seated position to avoid any inconveniences, especially since most of the participants had not used a VR set before. The OBS screen-recording software was used to record the experiment and the Oculus Run environment was used to mirror what the students see through the headset so the researcher could monitor their interaction and provide support when needed.

Figure 1: Experiment Set-up

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Figure 2: Oculus Rift headset, Touch devices, and sensors

The VR experience

The Raven (2017) is a free single-user audio-visual experience for Oculus Rift based on Edgar Allan Poe's poem 'The Raven' read by Barry Carl with art by Thomas Pasieka, code by Eric Liga, and music by Jordan Rudess (Dream Theater) (Figs. 3 and 4) which is available via the Oculus store and categorized under entertainment and education. The 10-minute experience consists only of one module (interface) modelled in the shape of a room, which includes various artefacts representing the general atmosphere of the poem. Five teleportation pads (for movement inside the experience) are placed on the ground of the room. The poem's text is displayed on the wall in front of the user in sync with the audio voice-over. The user can interact with the experience using the right Touch device (pressing buttons to perform tasks) and head-gaze (looking at the pads on the ground activates them). There is no fixed scenario for this experience, and users can move around and explore their surroundings as they like while listening to the audio recitation.

Figure 3: The RAVEN VR experience (teleport pads)

Figure 4: The RAVEN VR experience (poem on the wall)

Data collection and analysis

Three quantitative research tools were prepared for this study, each of which collects data that was used to answer one of the research questions:

1- A pre/post engagement questionnaire (Appendix 1): Both the pre and post versions of the engagement questionnaires were designed to include 20 Likert-scale items investigating students' subjective (or self-perceived) sense of engagement in virtual environments; in other words, the impact of virtual environments on learners' engagement. Some of the questions, which are included in the pre/post engagement questionnaire, that I thought were important to measure students' engagement were created by me and the others were adopted from other research papers as they developed scales for measuring student's engagement in e-learning such as (Brockmyer et al., 2009 and Lee et al., 2019). A 5-Likert scale was used: Totally Agree, Agree, Neutral, Disagree, Totally Disagree, which were respectively coded as (5, 4, 3, 2, 1).

2-

In the pre-engagement questionnaire, students were supposed to rate their engagement with virtual learning environments used in KAU (particularly the Blackboard LMS), which is a learning management system used to support learning with several collaborative tools in which teachers and students can interact in virtual environments throughout sharing educational materials, providing assessments, and in the post-engagement questionnaire they were supposed to evaluate their engagement in the VR experience. To investigate the reliability of the questionnaire instrument, I piloted it on 10 students other than the actual participants in this research, then inputted their responses in SPSS (Statistical Package for the Social Sciences) and computed the Cronbach's alpha value, which is a measurement of the internal consistency of the instrument. High Cronbach's alpha values indicate the reliability of the tool and the relation between its items. According to research, a Cronbach's alpha of 0.70 or higher is considered 'acceptable' in educational and social research (Cortina, 1993). In our case, Cronbach's alpha was 0.779, which indicates the reliability of the instrument.

The pre-engagement questionnaire was prefaced by administering a learner's profile survey (Appendix 2). The latter tool included 5 items requesting information about students' academic level, ICT skills, prior use of VR, prior use of virtual learning environments in KAU, and preferred learning styles. The data from this short survey was used to investigate whether or not those general variables have an impact on students' engagement before and after the experiment.

3- The NASA TLX (Task Load Index): To study the impact of the students' cognitive load, i.e. the effort which they exerted to navigate the VR experience, on their engagement, I adopted the NASA TLX tool (Appendix 3) which is used by researchers in various fields to measure the cognitive load of technology users (NASA, 1986). The instrument is based on a subjective rating process and consists of two sections:

1) a 6-item survey representing the individual sub-scales of the NASA TLX, which include: mental demand, physical demand, temporal demand, performance, effort, frustration. The NASA TLX Manual explained these sub-scales as follows:

- Mental demand: how much thinking, deciding, or calculating was required to perform the task.
- Physical demand: the amount and intensity of physical activity required to complete the task.
- Temporal demand: the amount of time pressure involved in completing the task.
- Effort: how hard does the participant have to work to maintain their level of performance?
- Performance: the level of success in completing the task.
- Frustration level: how insecure, discouraged, or secure or content the participant felt during the task (17).

Before asking students to rate their cognitive load, we had first require them to perform a number of tasks inside the VR experience. Therefore, an instruction sheet (Appendix 4) was developed with 12 pre-defined tasks that students had to successfully accomplish in VR.

After accessing the VR experience and performing the assigned tasks, students had to rate each of the 6 sub-scales on a scale represented by a 12-cm line divided into 20 equal intervals ending on both sides with the descriptors (low/high). As reported on the NASA TLX Manual (1986): 'The 21 vertical tick marks on each scale divide the scale from 0 to100 in increments of 5. If subject marks between two ticks, the value of the right tick is used (i.e. round up)' (Fig. 5).

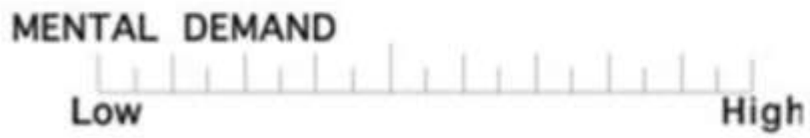


Figure 5: The rating mechanism for each of the NASA TLX sub-scales

2) According to the NASA TLX Manual (1986), the 15 pairings of 6 sub-scales described above are used determine which sub-scale of each pair is the most influential in the overall weighted load. These pairings are: Effort or Performance, Mental Demand or Physical Demand, Physical Demand or Temporal Demand, Effort or Physical Demand, Performance or Temporal Demand, Frustration or Mental Demand, Physical Demand or Performance, Performance or Mental Demand, Temporal Demand or Mental Demand, Mental Demand or Effort, Frustration or Effort, Performance or Frustration, Temporal Demand or Frustration, Temporal Demand or Effort, Physical Demand or Frustration (15-16).

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

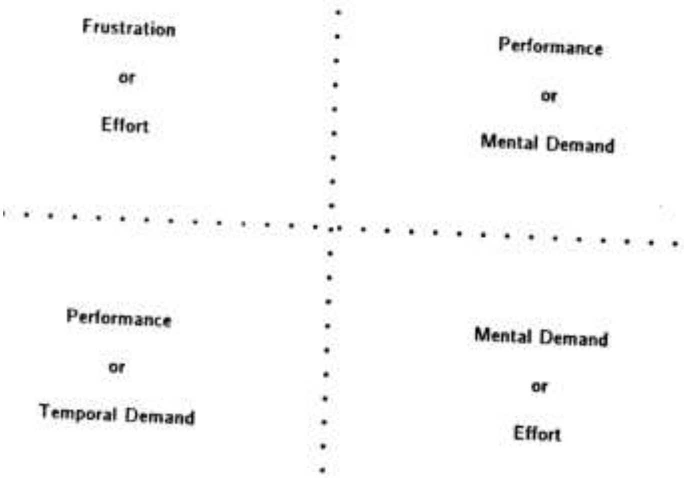


Figure 6: The rating mechanism for the pairs of the NASA TLX

The overall weighted average is then calculated based on the averages of the students' ratings for both the individual and paired sub-scales. This value represents the self-perceived cognitive load of the student in the VR experience.

To automate the process of collecting students' ratings of their cognitive load, I used an online version of the NASA TLX tool (Sharek 2009) (Fig. 7) instead of the pen and paper version, as it helped in setting up the experiment, entering the research participants IDs and doing the calculations, which were downloaded and processed later on in SPSS (Fig. 8). The participants were presented with two simple interfaces: One for the section of the NASA TLX, where they had to rate the sub-scales (Fig. 9), and the other with all the possible parings, which the students had to swipe though (Fig. 10). Directions and explanations were presented for each part of the tool.

Figure 7: The online version of the NASA TLX tool

Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

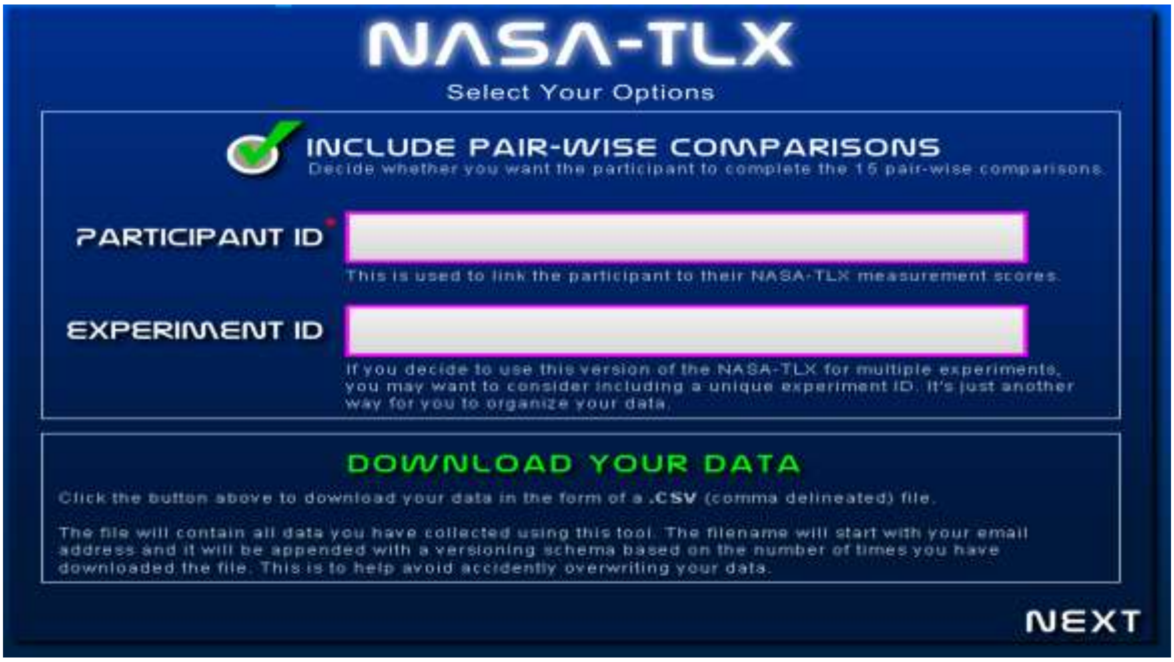


Figure 8: NASA TLX online tool: setting up the experiment

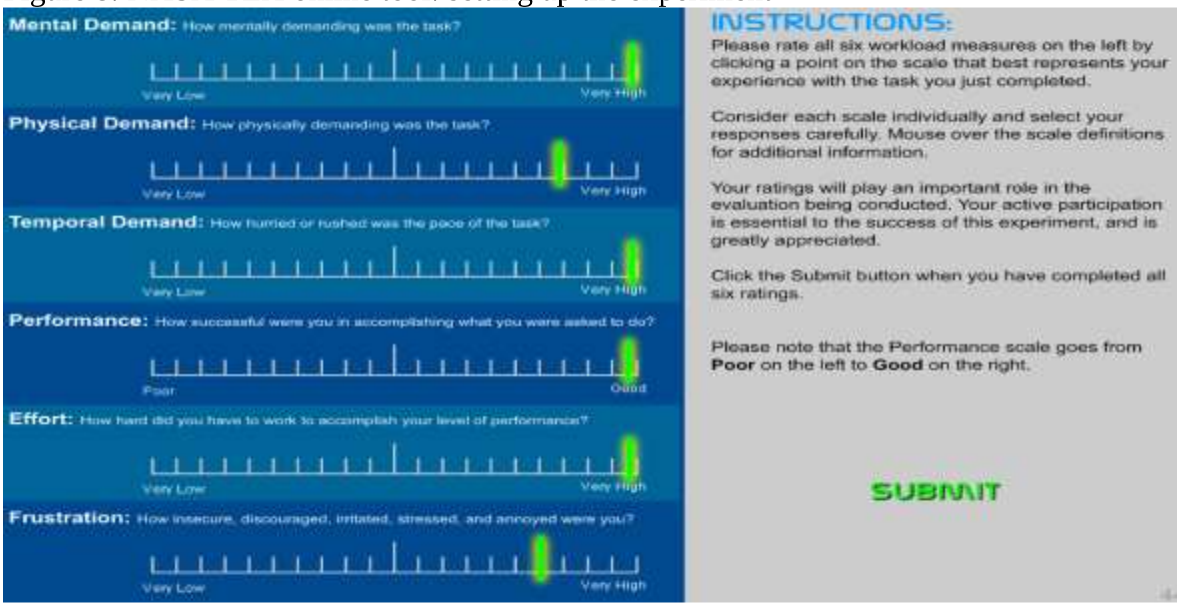


Figure 9: NASA TLX online tool: first interface

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Figure 10: NASA TLX online tool: second interface

2) Students' Behaviour Inside VR rubric (Appendix 5): This is a rating instrument used by the researcher to evaluate students' levels of task achievement in VR as demonstrated through the video recordings (for the video recordings see Appendix 11). The data for each student was entered into an individual rubric sheet, then aggregated using Excel. These responses were then processed in SPSS. The rubric's scale values ranged from 4 to 0, where 4 indicated that the student was able to 'successfully execute the task using the assigned path', 3 signified that the student was able to 'successfully execute the task using the assigned path but some sub-tasks cannot be fulfilled', 2 denotes that the learner managed to 'successfully execute task using an alternative path', and 1 implied that the learner was not able to 'execute task using the assigned path' (e.g. followed the assigned path but did not land at the assigned destination), and 0 meant that the student 'could not figure out how to access the assigned path'.

In addition to these three research instruments, I used a qualitative interview tool (Appendix 6) which consisted of 7 questions. The first three were primarily closed-type questions rating the quality of VR content (audio, visual, and readable). The rating evaluation was based on an Excellent, Good, Fair, and Poor scale that was numerically coded in SPSS from 4 to 1. The following two open-ended questions inquired about the most/least engaging aspects of the VR experience and asking for justification. The final two questions were open-ended ones as well inquiring about students' perceptions of the potential of VR in EFL education considering effort and tasks.

To compute the significant of differences between the various data types collected for this study I conducted the following statistical treatments using SPSS and Excel with the purpose of examining the correlations between the variables (engagement, cognitive load, and students' behaviour in VR):

- The Kolmogorov-Smirnov Test of Normality was used to determine whether or not the mean scores of the pre/post engagement questionnaires are normally distributed. This helped in determining the suitable type of statistical tests to use later on to analyse the data.

- The non-parametric Wilcoxon Matched-Pairs Signed-Ranks Test was used to compare the means of the pre/post engagement questionnaire.
- The ANOVA significance values were computed for each general variable to investigate the relation between the participants' general information and their level of engagement before and after the experiment.
- The Spearman's Rank Correlation Coefficient was used to compute any significant correlations between students' engagement and their cognitive load, and their engagement in relation to their behaviour in VR.
- Rosenthal's (1994) formula for *effect size* was used in Excel to calculate the *effect size* value (r) which represented the significant change between students' engagement before and after the VR experiment.

Procedure

In compliance with research ethics, two consent forms were distributed among the participants (Appendices 7 and 8), one before the pre-engagement questionnaire and the other one before the actual experiment in VR. The participants were informed that their participation in the study does not involve any physical or emotional risk to them beyond that of everyday life and also that the information collected from them will be confidential and the data will be used only for analysis and investigation of the topic under study. They were made fully aware that their participation in this study was voluntary and that they did not have to answer any question they did not want to answer.

The research process could be summarized as follows:

1. Students sign the first consent form and fill in the pre-engagement questionnaire under supervision.
2. Students are asked to read the instructions sheet, which include the tasks they have to perform in VR. I am there to answer any of their questions and clarify any points that they do not understand.
3. The Oculus Run environment is activated to mirror what the student sees via the headset, and the OBS recording software is opened to capture each student's interactions in VR. Then each student is required to wear the VR headset and perform the tasks one by one while I supervise and support them. After each student, I make sure to stop the recording and save it to the video database.
4. After exiting the VR experience, students are requested to fill in the NASA TLX cognitive load index via a PC device located at the lab where the experiment was conducted. Before and while they are filling in the TLX tool, I clarify concepts or requirements which they have to observe.
5. During the same setting, students are asked to fill in the 20-item post-engagement questionnaire
6. They are then asked to answer the interview questions about their experience in VR.

The experiment time takes about 25-35 minutes covering all the tasks mentioned above. All study procedures took place at the Deanship of E-learning and Distance Education (E-learning Innovation Lab) in King Abdul-Aziz University, Building 66/lab. The KAU was chosen for my study is because of their well-equipped lab and the accessibility of the various resources needed for this study.

RESULTS

The research experiment described in the previous sections was conducted during the 7th, 8th, and 10th of July 2019. The total number of participants was (*n*=10) EFL students from the Department of European Languages in King Abdul-Aziz University, who were enrolled in different academic levels. Basically, I used three research tools to collect quantitative data: 1) a pre/post self-perceived engagement questionnaire, 2) the NASA TLX (Task Load Index) which was used to calculate the learners' cognitive load, and 3) a Students' Behaviour in VR rubric, which was utilized to analyse learners' video-recorded interactions in VR. A fourth tool, an interview, was also used to collect qualitative data about students' experiences in VR. The data collected by the three quantitative research tools was cleaned and then pre-processed and analysed using IBM SPSS, which was used to investigate the correlations between the variables (engagement, cognitive load, and students' behaviour in VR). This section will categorize and present the results of the analyses, which were carried out in relation to the following main research questions:

1. What is the effect of using VR environments on learners' engagement in EFL learning at KAU?
2. Is there any correlation between students' cognitive load and engagement in virtual environments?
3. Is there any correlation between students' behaviour (interaction) and engagement in virtual reality environments?

Sub-section 4.1 compares the mean scores of the engagement questionnaire before and after the implementation of the VR experience. Sub-section 4.2 presents the results of the NASA TLX or cognitive load index and attempts to find correlations between students' engagement mean scores (post-VR experiment) and their overall cognitive load scores. Sub-section 4.3 then investigates whether or not correlations exist between students' engagement mean scores (post-VR experiment) and their behaviour inside the VR experience. Finally, 4.4 categorizes the students' interview responses that reflect on the VR experience.

The effect of VR environments on EFL learners' engagement

To examine the effect of the VR experience on students' level of engagement, I inputted the pre/post engagement questionnaires (Appendix 1) responses for all of the questionnaire 20 items in SPSS, and computed the means representing the overall engagement for each of the 10 students (the total number of research participants in this study).

To determine whether or not these mean scores are normally distributed, two methods were used: 1) a statistical test of data normality, and 2) a visual method. Basically, to say that a set of data is normally distributed means that the *p*-value (Sig.) extracted through a statistical test should be above .05 and visually a histogram plot of the data should demonstrate a bell-shaped curve (Das and Imon, 2016, 5-12). These tests of normality help in determining the type of statistical tests that can be used to examine the data.

Therefore, to determine whether or not these mean scores are normally distributed, the Kolmogorov-Smirnov test of normality was conducted using SPSS. Table 1 shows that the *p*-

Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

value is .02, which is below .05, hence one can safely state that the mean scores are not normally distributed.

Table 1: Kolmogorov-Smirnov

	Statistic	Df	Sig.
PRE_TEST	.153	10	.0200*
POST_TEST	.164	10	.0200*

The graphical histogram plots in Chart 1 below further emphasize the abnormal data distribution in our data set. The histograms demonstrate the frequency distribution in which the observed values are plotted against their ‘frequency’; thus, present a visual assessment of whether the distribution is ‘bell shaped”’ or not (Das and Imon, 2016), which is not the case for the pre/post engagement data in this research.

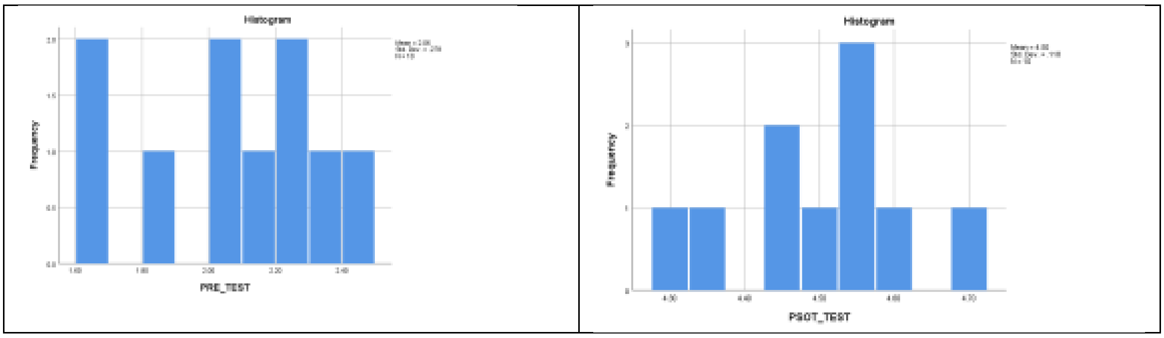


Chart 1: Test of normality histograms for pre-test and post-test

The primary aim of this section was to identify any variances in mean scores between the pre/post implementations of the engagement questionnaire. Therefore, a non-parametric Wilcoxon matched-pairs signed-ranks test was used to test for a difference in the mean of the paired observations (pre/post-engagement). Non-parametric tests are used when research samples are small, which might lead to a statistically significant abnormality in data distribution (as is the case with the responses of the engagement questionnaire in this study) (Conover, 1980, pp. 278-292).

The Wilcoxon matched-pairs signed-ranks test was implemented to investigate whether or not there is an increase in students’ subjective (or self-perceived) sense of engagement after conducting the experiment in VR. Tables 2, 3, and 4 below show the results of this test.

The descriptive statistics in Table 2 display a recognizable variation in the overall engagement mean between the pre/post responses. The pre-test mean (2.06) compared to the post-test mean (4.5) shows an increase in the engagement score after the VR experience.

Table 2 Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
PRE_TEST	10	2.0600	.27769	1.65	2.45
POST_TEST	10	4.5000	.11785	4.30	4.70

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Additionally, Chart 2 offers a boxplot visualization of the variations in pre/post engagement questionnaire means. Of course, the most noticeable feature is the space separating the two boxplots, which reflects the mean difference reported in Table 2. One observable feature of the two boxplots in Chart 2 is that the one visualizing the pre-test is longer comparable to the one representing the post-test. This shows that there was more variation among students' responses to the engagement items in the pre-test questionnaire, whereas the shortness of the post-test boxplot indicates the students' overall agreement on their level of engagement after conducting the VR experiment.

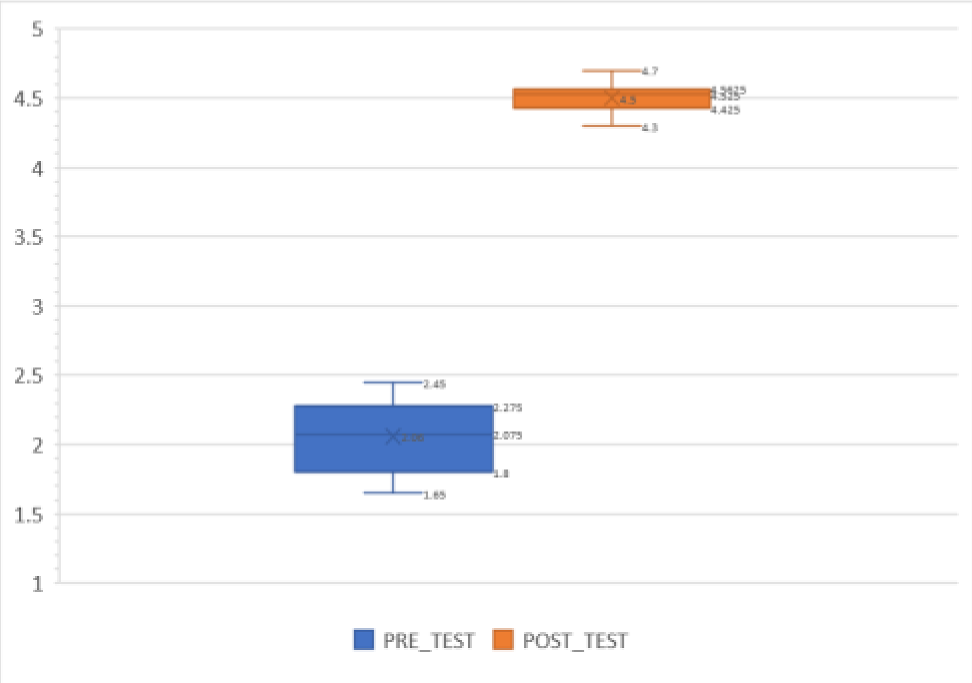


Chart 2: Boxplots of pre/post engagement questionnaire mean scores

The ranks in Table 3 provide some interesting data on the comparison of participants' engagement scores before (Pre) and after (Post) conducting the VR experiment. We can see from the table's legend that 0 participants had a lower engagement score after their VR experiment. All 10 participants constituting the sample of this research had a higher engagement score after the VR experiment, and (0) participants saw no change in their engagement score.

Table 3: Ranks

		N	Mean Rank	Sum of Ranks
POST_TEST PRE_TEST	-Negative Ranks	0[a]	.00	.00
	Positive Ranks	10[b]	5.50	55.00
	Ties	0[c]		
	Total	10		

a. POST_TEST < PRE_TEST
b. POST_TEST > PRE_TEST
c. POST_TEST = PRE_TEST

Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

By examining the final Test Statistics in Table 4, I investigated whether the changes which were due to learners' participation in the VR experiment led overall to a statistically significant difference in engagement scores. I was looking for the 'Asymp. Sig. (2-tailed)' value (alternatively, the *p*-value), which in this case is 0.005. This value proves that the difference between the pre/post questionnaire mean-scores is significant and that there is an increase in the means of engagement as reported in the post-engagement questionnaire responses. This positive result might be attributed to the use of VR, at least in this particular context.

Table 4: Test Statistics[a]

	POST_TEST - PRE_TEST
Z	-2.807[b]
Asymp. Sig. (2-tailed)	.005

a. Wilcoxon Signed Ranks Test
b. Based on negative ranks.

To further study the correlations between the learners' general information (academic level, ICTskill, prior use of VR, prior use of virtual learning environments in KAU, and preferred learning style) and their engagement scores before and after the experiment, the ANOVA significance values were computed for each variable. The results reported in Table 5 indicate that there was no observable significant impact of these factors on students' engagement, as the significance was above the 0.05 value.

Table 5: ANOVA significance values for the general variables

ITEM	Pre-test	Post-test
Academic level	0.300	0.953
ICT skill	0.754	0.806
Prior use of VR	0.835	0.681
Prior use of virtual environments in KAU	0.540	0.630
Preferred learning style	0.487	0.438

Thus, the results of comparing the means of the students' engagement scores before and after their interaction with the VR experience revealed an increase in engagement levels which on further investigation could not be attributed to students' academic level, ICT skills, prior use of VR, prior use of virtual environments in KAU, or preferred learning style.

The last step would be to compute the *effect size* (*r*) of the change in engagement mean scores identified through the 'Wilcoxon matched-pairs signed-ranks test', because this is the focus of the first research question of this study: 'What is the effect of using VR environments on learners' engagement in EFL learning at KAU?' According to the research, the Effect Size (*r*) is a statistical measure that determines the strength of a link between two variables in a dataset. It assists researchers in understanding the importance of a result, using a scale that is independent of the unit of measurement being used (Kelley and Preacher, 2012). This is particularly useful when comparing change for two different kinds of measures, as in this research between the mean scores representing students' engagement before and after accessing the VR experience.

Thus the determination that there is a positive and significant difference between students' pre/post engagement mean scores should be supported by calculating the effect size as well, which was done in Excel using the value reported in Table 4; i.e., the z-score (-2.807) and the *n=20* value (number of paired observations; in this case, 2 observations for each participant). The formula used for that was Rosenthal's (1994):

$$r = Z/\sqrt{N}$$

The calculated effect size (r) = (-0.627664281) which indicates that there is a moderate effect size. Cohen (1977) suggested that an effect size with a value = 0.2 is small, 0.5 moderate, and 0.8 large. Though the r is negative (-) in our results, the negative sign only indicates the direction of the effect. So the reported value (-0.6) still falls within the moderate effect size even when considering the negative direction.

Therefore, one can conclude that the first question of this study has been answered in two main steps: 1) identifying the difference in the mean scores between the pre/post implementations of the engagement questionnaire using the Wilcoxon matched-pairs signed-ranks test, then 2) measuring the effect size of this difference. As the results showed, the significance of the difference between the means falls below the 0.05 value, and the effect size of that difference is moderate (-0.6), showing that the use of VR has created a sense of engagement in students which outweighed the engagement levels reported while using other virtual environments in KAU.

Acknowledging that the use of VR positively impacts students' engagement should lead to considering what influences EFL students' engagement in VR, and specifically the role that the other two variables of this research, cognitive load and behaviour, play in improving EFL students' engagement in the virtual reality experience chosen for this study. The findings of correlations between students' engagement scores and their cognitive load in VR are to be discussed in the Section 4.2 to investigate whether or not cognitive load impacts engagement in EFL educational settings.

Correlations between EFL learners' engagement and cognitive load

This section reports both students' cognitive load results, which were assessed using the NASA TLX tool (Appendix 3), as well as the correlation between their engagement scores and the overall cognitive load score. As with the Engagement Questionnaire, students' ratings of the NASA TLX 6 individual and 8 paired sub-scale items were subjective (self-perceived) and recorded after the tasks in VR were completed according to the cognitive walkthrough instruction sheet (Appendix 4), which was distributed among participants prior to accessing the VR experience. Therefore, one can safely say that students' ratings of the NASA TLX were more or less their own estimation of the mental demand, physical demand, temporal demand, performance, effort, and frustration load rates which they experienced in VR.

The online tool developed by Sharek (2009) (Figure 7), which automated the process of the NASA TLX, was used to collect students' subjective assessment of the cognitive load they experienced after being in VR. Students rated the six sub-scales of the index (mental demand, physical demand, temporal demand, performance, effort, frustration) individually and in pairs using the online tool directly after exiting the VR experience. The NASA TLX CSV file input reported: 1) the weighted averages for each subscale (individually and in pairs) as well as 2)

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by *ECRTD-UK*
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

an overall cognitive load score based on these weighted averages, which is supposed to range from 1 to 100. Appendix (10) is a spreadsheet view of the results.

The (tlx_Score) column represents the overall cognitive load score, which ranges from 1.17% to 12.17%, thus indicating low cognitive load ratings for all participants (10). The higher the value of the NASA TLX score, the more demanding the technology. In this case, students gave low cognitive load scores, which means they did not rate the VR experience as demanding.

To identify which sub-scale scored the highest weighted mean, which can thus be considered influential in comparison to others, I computed the means for each sub-scale using the input ratings from the students. Chart 3 demonstrates that the temporal sub-scale scored the highest (11.9), followed by the mental (9.4). It seems that both the effort and frustration sub-scales were related as they scored 7.1 and 7.3, respectively. Performance came next at 5.4. The lowest mean was for the physical sub-scale (2.4). The physical as well as the performance sub-scales seem to be the lowest in comparison to others, hence of lower impact on the overall TLX score. This suggests that the students' self-reported evaluation of their VR interaction did not consider the physical demand a major hinderance. They also reported that they did not encounter a problem accomplishing the assigned tasks in VR (performance).

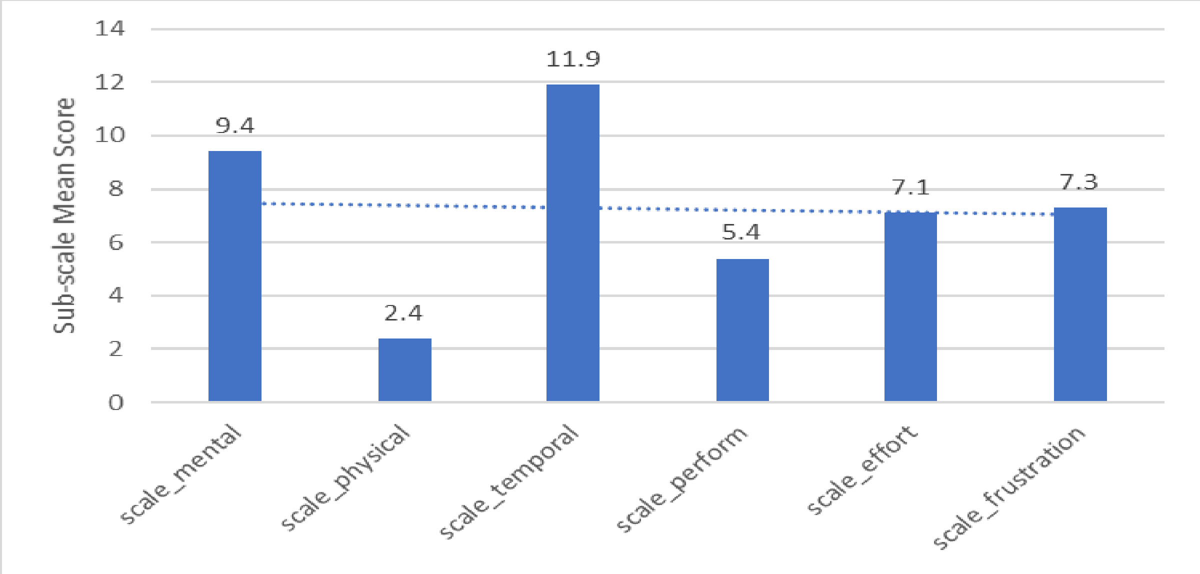


Chart 3: NASA TLX sub-scale individual scores

Chart 4 is a stacked bar visualization of the weighted means of the 8 pairs of sub-scales that constitute the second part of the NASA TLX: Effort or Performance, Temporal Demand or Effort, Performance or Frustration, Physical Demand or Performance, Temporal Demand or Frustration, Physical Demand or Frustration, Physical Demand or Temporal Demand, and Physical Demand or Mental Demand. For each of the 10 students participating in this study, the scores resulting from comparing the pairs were averaged and assigned to each of the sub-scales indicating which sub-scale was chosen the most in the pairing section of the test. As Chart 4 demonstrates, 'Frustration' was not chosen in any of the pairs which the 10 students rated. This finding clearly contrasts with their rating of the first part of the NASA TLX, where the weighted average of the 'Frustration' sub-scale scored 7.3.

Moreover, for all participants, the 'Physical Demand' of the VR experience was within the lowest ranges of all the tested pairs (0.1-1). This result is comparable to the sub-scale ratings in the first part of the NASA TLX. A noticeable variation in the range of scores for the 'Temporal' sub-scale (0-3.67) indicates that students differed in the evaluation of this dimension, with results that can be compared to the previously demonstrated scores for the individual sub-scales in Chart 3.

Chart 4: NASA TLX sub-scale paired scores

The following step was to identify any significant correlations between the overall NASA TLX scores and the post-test engagement mean scores. The researcher assumed that the higher the student's cognitive load score, the lower her engagement score and vice versa. In both cases, the statistical test should establish any correlation whether positive or negative.

Because I was analysing data that was not normally distributed (as indicated previously) I opted for using a non-parametric correlation test (the Spearman's Rank Correlation Coefficient) which was used to determine the strength of a link between two sets of data: In this case, the weighted average representing the overall TLX score, and the means (averages) of the post-implementation engagement questionnaire. This test gave a *p*-value indicating significance. If the *p*-value is less than .05, then one can assume that there is a statistically significant bivariate

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

association between the two variables. If the *p*-value is more than .05, then there is no statistically significant correlation between the variables (Dodge, 2010, p. 502).

The first step in this process was to normalize the two sets of previously computed scores, because they were not using the same scale. Normalizing the data in this context meant rescaling its values so that they can be comparable. The values for the NASA TLX scores ranged from 1 to 100, whereas the means of the engagement data ranged from 1 to 5. Therefore, the z-score values for each set of data were computed using SPSS, so that the values of both data sets be (0-1).

The Spearman's Rank Correlation Coefficient test was then applied on both sets of normalized data. This is a nonparametric test to identify rank correlation (statistically measured dependency between the ranks of two variables). It evaluates how clearly the link between two variables can be demonstrated and proven. Ideally, the Spearman correlation between two variables is reported as high when the tested observations have a similar (or identical for a correlation of 1) rank between the two variables, and low when observations have a unrelated (or fully opposite for a correlation of −1) rank between the two variables (Corder and Foreman, 2014).Table 6 below presents the results.

Table 6: Spearman's Rank Correlation Coefficient

			Z-score(TLXscore)	Z-score(PostEng)
Spearman's rho	Z-score(TLXscore)	Correlation Coefficient	1.000	.185
		Sig. (2-tailed)	.	.610
		N	10	10
	Z-score(PostEng)	Correlation Coefficient	.185	1.000
		Sig. (2-tailed)	.610	.
		N	10	10

As reported in Table 5, the significance value for the correlation between the post-engagement and TLX scores is 0.61, which is above the statistically significant value (0.05), hence this could be interpreted as there being no sufficient statistical evidence to support the researcher's assumption that there is a correlation between students' engagement and cognitive load, at least among the sample used in this study.

To identify if there were any correlations between the individual or paired NASA TLX sub-scales and the engagement mean scores, the Spearman's Rank Correlation Coefficient Test was also conducted on the normalized means of each sub-scale and paired item. Current research has encouraged investigating the NASA TLX sub-scales or paired items and focusing only the overall score of NASA TLX (Galy et al., 2018). Results (reported in Appendix 9 and 10) further support that there is no significant correlation between the students' engagement and cognitive load scores.

Therefore the results reported in this section answer the second research question, 'Is there any correlation between students' cognitive load and engagement in virtual environments?' Considering both the correlation between engagement and the overall NASA TLX scores as well as between engagement and the individual and paired sub-scales of NASA TLX, there was no significant relation or link between sets of data. So the amount of cognitive load which the students exerted did not affect their engagement inside the VR experience.

The last variable to be investigated is students' behaviour in VR and how it might correlate with their engagement. Section 4.3 presents the results for this dimension of the study.

Correlations between EFL learners' engagement and behavior in VR

Students' behaviour in the VR experience was recorded using a video recording software (OBS). The videos captured how students executed a set of predefined interactions or tasks, which were introduced to them through a cognitive walkthrough instruction sheet (Appendix 4). A Students' Behaviour Inside VR rubric (Appendix 5) was then used to assess learners' success levels in accomplishing the assigned tasks in VR based on the video output: The assessment values ranged from 4 to 0, where 4 indicated that the student was able to 'successfully execute task using the assigned path', 3 signified that the student was able to 'successfully execute the task using the assigned path but some sub-tasks cannot be fulfilled', denotes that the learner managed to "successfully execute task using an alternative path", 1 meant that the learner was not able to 'execute task using the assigned path' (e.g. followed the assigned path but did not land at the assigned destination), and (0) meant that the student 'could not figure out how to access assigned path'.

The raw data extracted from the rubric (based on the video output) was then processed in SPSS to compute the mean scores for the rubric tasks. Chart 5 demonstrates the mean values for each task, which denote the levels of students' success in accomplishing each task in VR.

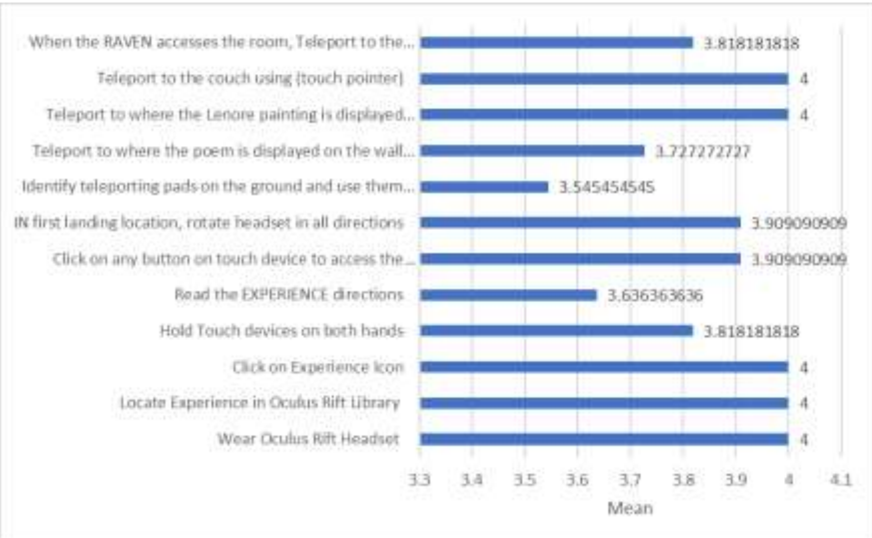


Chart 5: Mean scores for students' behaviour rubric based on video recordings

Overall students' performance rates on all tasks, as shown in the video recordings, was high, ranging from 3. 5 to 4, which indicated that the cognitive walkthrough tasks were either 'successfully' executed 'using the assigned path' or 'successfully' executed 'using the assigned path but some sub-tasks cannot be fulfilled'.

The lowest values were for identifying teleporting pads on the ground and using them to move around in the VR environment (3.5), and for reading the experience directions in VR (3.6). Both tasks required a number of sub-tasks like head rotation, focusing, and figuring out how the buttons on the Touch device might help in executing the task. Whether the interaction was with an object (the teleporting pads) or written text (the poem on the wall) students seem to need an alternative method of help or extra directions from the instructor to proceed. However, even with these low values for these two particular tasks, one can still consider that the students' behaviour inside VR led to successful and positive results.

One can further classify the results in Chart 5 as follows:

- In tasks which relate to the physical dimension of the VR experience — 1) Wear the Oculus Rift headset, 2) Hold Touch devices in both hands, 3) Click on any button on Touch device to access the VR experience, 4) Rotate the headset in all directions — students achieved the highest results, ranging from 3.8 to 4.
- In tasks in which students have to interact actively with the VR objects (accessing the VR steps, or actual teleporting from one place to another) the results were also high, ranging from 3.7 to 4.
- In tasks where students were supposed to read, either the directions of the VR experiences or the poem on the wall, results were lower, ranging from 3.6-3.7.
- In one task, students were asked to 'identify' objects in the VR experience (teleporting pads) and the overall mean is the lowest at 3.5.
-

So the more active or complex the interaction was, the more students struggled with executing it in VR.

The main aim of this section was to investigate whether or not there was a significant correlation between the students' performance as demonstrated in the Students' Behaviour Inside VR rubric data and the mean scores which they achieved on their post-engagement questionnaire. The data was tested for normality of distribution, and the results indicated otherwise. Therefore a non-parametric correlation test, the Spearman's Rank Correlation Coefficient, was used to decide if a relationship can be established between the two sets of data: in this particular context, the mean scores of the Students' Behaviour Inside VR rubric and the means of the engagement questionnaire (post-implementation). As indicated previously, if the p-value of the Spearman's Rank Correlation Coefficient Test is less than .05, then one can assume that there is a statistically significant bivariate association between the two variables. If the p-value is more than .05, then there is no statistically significant correlation between the variables.

The z-score for the mean scores of each dataset was then computed, because these sets were not on the same scale and hence needed to be normalized and rescaled. The values for the Students' Behaviour Inside VR rubric range from 0 to 4, whereas the means of the engagement data range from 1 to 5. After the normalization process, the z-score values for both datasets ranged from 0 to 1, and the Spearman's Rank Correlation Coefficient Test was applied, which resulted in the values reported in Table 7 below.

Table 7: Spearman's Rank Correlation Coefficient

			Z-score (PostEng)	Z-score: Student Behaviour
Spearman's rho	Z-score(PostEng)	Correlation Coefficient	1.000	-.261
		Sig. (2-tailed)	.	.467
		N	10	10
	Z-score: Student Behaviour	Correlation Coefficient	-.261	1.000
		Sig. (2-tailed)	.467	.
		N	10	10

As Table 7 demonstrates, the significance value for the correlation between the mean-scores of the Students' Behaviour Inside VR Rubric and those of the Post-Engagement Questionnaire was (0.467), which was above the statistically significant value (0.05) and therefore could be interpreted as there being no sufficient statistical evidence of a correlation between students' engagement and their behaviour in VR. Thus, the third research question, 'Is there any correlation between students' behaviour (interaction) and engagement in Virtual reality environments?', is answered indicating the non-existence of a link (at least, in the context of the current implementation) between the two research variables of engagement and students behaviour in VR.

It still remains to consider the students' reflective feedback on the whole experiment. Section 3.4 presents the students' qualitative input concerning their VR experience using an interview tool that focused on three dimensions: quality of VR content, students' reflection on their engagement, and students' reflection on the use of VR in EFL contexts.

Students' feedback: interview responses

The last research tool to which students contributed was an interview (Appendix 6) consisting of 7 questions in total; 3 of them were closed-type questions, and the remaining 4 were reflective. Students' responses to the interview questions were categorized under three sub-headings:

- **Quality of VR content**, where students were asked to rate the audio, visual, and readable content as delivered via the VR experience. The ratings available were Excellent (4), Good (3), Fair (2), Poor (1).
- **Students' reflections on engagement**, where students had to answer two questions about the most/least engaging aspects of the VR experience and provide reasons for the rating.
- **Students' reflections on the use of VR in EFL contexts**, where they had to answer two questions about the possibility that either effort or assigned tasks might prevent them from using VR to learn English. The rating units for the first of these two questions were No (3), Maybe (2), Yes (1). The reversed coding meant that a No response is positive; whereas a Yes response is negative. They were required to provide justifications for their answers.
-

Quality of VR content

The content of a VR experience is usually an essential component that should be evaluated to investigate whether or not it impacts learners' engagement. Primarily, students' responses were coded in SPSS, then simple statistics were computed, as shown in Table 8 and Chart 6.

Table 10 demonstrates that for all three content metrics (audio quality, visual quality, and readable content quality) students' mean score ratings ranged from Excellent (4) to Good (3), which could be considered a positive estimation of the quality of content in the virtual reality experience they were exposed to.

Table 8: Means and Std. Deviation of the VR Content Metrics

		AudioQ	Visual Q	ReadQ
N	Valid	10	10	10
	Missing	0	0	0
Mean		4.0000	3.5000	3.1000
Std. Deviation		.00000	.70711	.87560

Chart 6 below is a visualization of both the mean scores and the standard deviation values for all of the three VR content metrics. We notice an overlap between the error bars for all three columns representing VR content metrics. This overlap points to the absence of a significance correlation between these metrics; i.e. the students' positive evaluation of the VR audio quality did not impact, for example, their evaluation of the visual or readable content, which might be interpreted in different ways in the context of our research project. It could mean that each of these components (or instructional elements) stands out on its own but not in a comprehensive instructional model that encourages learners' engagement. Conversely, it could be taken to indicate that the sample size impacted the reliability of the mean for each of the components, thus affecting the significance of the difference between the means of these metrics. One might say that although the ratings for the instructional content components are high, the absence of significant difference between the various metrics leads us to believe that the mean scores are not reliable due to the size of the sample, hence a generalization about the quality of the VR content cannot be drawn.

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

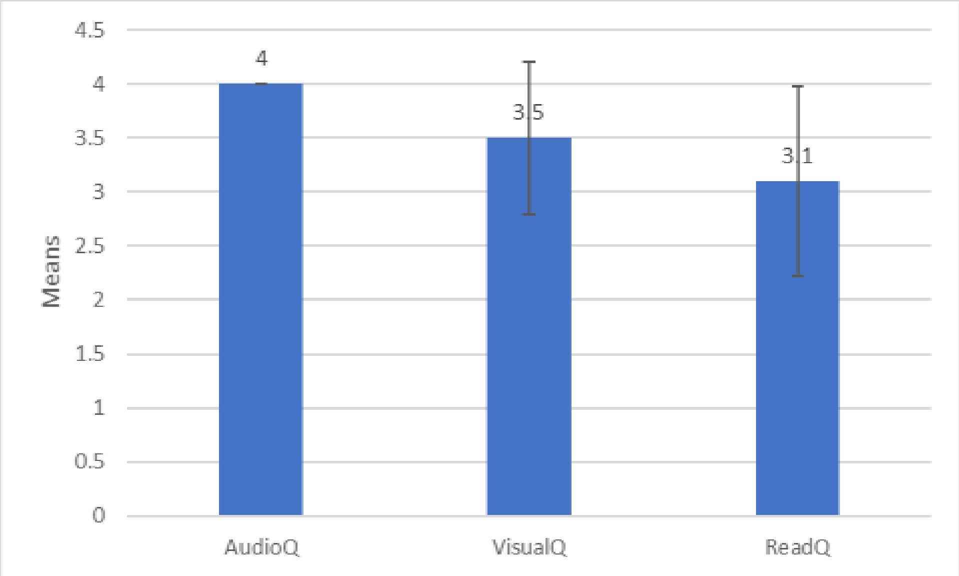


Chart 6: Mean-Scores Plotted Against (-/+ 1) Std. Deviation

The previous results are further substantiated by conducting the Spearman's Rank Correlation Coefficient Test to compare the post-engagement questionnaire mean scores and the mean scores of the content variables investigated in the previous paragraphs. Table 9 sums up the statistical test results indicating the absence of significance at the level of the 0.05 value, with an actual *p*-value of 0.194.

		Z score(ENG_Post)	Z-score(VR_Content_Q)
Pearson Correlation	Z-score(ENG_Post)	1.000	-.307
	Z-score(VR_Content_Q)	-.307	1.000
Sig. (1-tailed)	Z-score(ENG_Post)	.	.194
	Z-score(VR_Content_Q)	.194	.
N	Z-score(ENG_Post)	10	10
	Z-score(VR_Content_Q)	10	10

Students' reflections on engagement
The two interview questions about the most or least engaging aspects of the VR experience were purely qualitative in nature. Participants recorded their responses in Arabic, which then were thematically analysed to look for repeated patterns. The results are outlined below:

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ECRTD-UK
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

The most engaging aspects of the VR experience

- Four students reported that the most engaging VR feature for them was the ability to ‘see everything like I’m a player in a game’, ‘be inside the setting’ that the poem described, ‘feel like part of the place,’ or ‘get inside’ the world of the poem. All of them indicated that feeling as though they were inside the world of the poem helped them imagine the dark atmosphere of the poem, the storyline, and how the setting of the poem informs how one should read and interpret it. Two of the students, however, acknowledged the fact that they wanted to see the speaker of the poem and his emotions. This sense of immersion or presence in the VR experience seems to be very important for those learners, and they wanted more of it to be able to enjoy the experience and ‘learn’ from it.
-
- One student reported that the most engaging feature for her was the ability to ‘check all the objects in the game’. First, it is noted that most of participants thought of this VR experience as a game. Secondly, this student wanted to ‘play’ or ‘interact’ more with the ‘game’. The Raven VR allowed for limited interaction (moving from one place to another in the world of the poem).
- One student pointed to the visual content of the VR experience, what she called the ‘HD images and detailed furniture...’. This student justified that by saying that she likes to ‘imagine’ the world of literary works to be able to understand the ‘meanings and themes’ of these works. The Raven VR experience helped her understand the ‘theme’, though she ‘did not understand every word in the poem’.
- Three students wrote briefly about how this VR experience resembled ‘a movie’ which they can watch but be ‘like an actor’ in it as well. In a sense, their sense of engagement with this quality of the VR experience was modelled on an actual experience which they know already (watching movies) but they appreciated being participants.
- For one student, the most engaging feature was interestingly ‘the control of everything’, i.e. her ability to proceed as she likes in the world of the VR experience.
-

Least engaging aspects of the VR experience

- Three students, though differed in how they phrased it, considered the written poem on the wall to be the least engaging of all the aspects of the poem. For one of them, the ‘font’ in which the poem was written was distracting as well as the placement of the text once at the upper part of the wall and sometimes in the middle of it. The other two students considered the text of the poem ‘annoying”’ or ‘distracting’ because they could not place it in the context of the voice-over. On checking this feature in the VR experience, we noticed that the text and the voice-over were not misaligned. It seems that the students’ inability to listen and read at the same time was the reason behind their comments.
- Four students agreed that the least engaging feature of the Raven VR experience was the absence of other characters with whom they can interact. Again, the expected game context was brought up. On the one hand, one student wanted to see a character who can ‘explain the poem or its words’. Two students, on the other hand, wanted to see the speaker of the poem so they can ‘understand how he feels’, whereas one student wanted to see her own ‘avatar’. This last justification could be attributed to the need to feel presence in the VR experience.

Students' Reflections on the use of VR in EFL contexts

The last set of interview questions related to the potential of VR experiences in EFL settings from the students' perspectives. The amount of effort and the number of tasks were used as metrics. Table 9 demonstrates the means and standard deviations for each metric. Both means fall within the "Maybe" option. Overall, students reported that both the effort they exert in using the VR experience and the number of assigned tasks in VR might be influential in their decision to use VR in learning English as a Foreign Language.

Table 9: Mean scores and std. deviations for students' reflections on the use of VR in EFL contexts

Metric	Mean	Std. Deviation
Tasks	2.1000	.87560
Effort	2.5000	.84984

Chart 7 plots the means of these two questions against their standard deviation values to gain an initial look at the significance of the relation between the two metrics. Here we wish to investigate whether students' responses on one item could be influential in their rating of the other. It turns out, as the overlapping error bars in the chart indicate, that there is no visible evidence of significance.

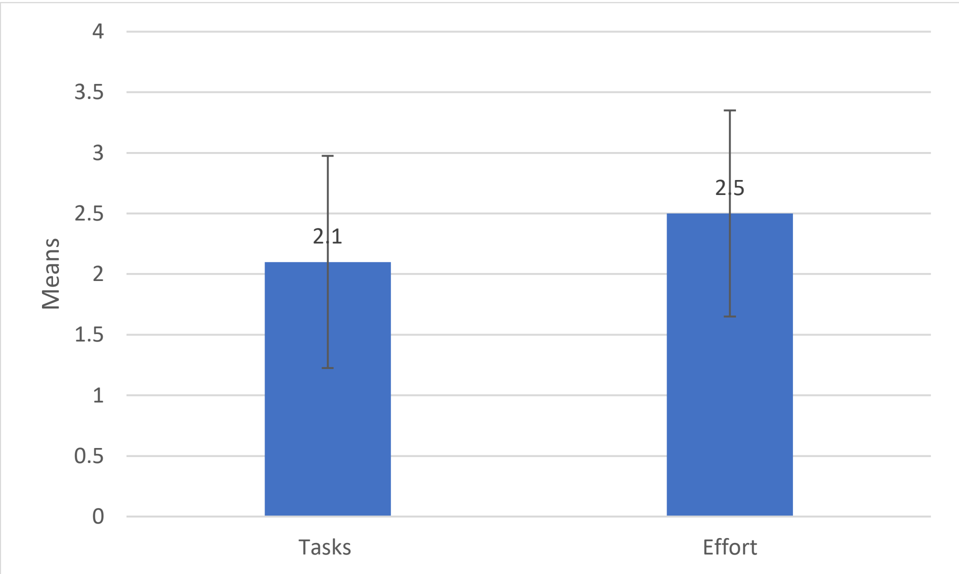


Chart 7: Mean-Scores Plotted Against (-/+ 1) Std. Deviation

To further examine the possibility of statistical correlation between the means of the post-engagement questionnaire and the means reported for the two interview questions investigated in this sub-section, the Spearman's Rank Correlation Coefficient Test was conducted and the results summarized in Table 10 below indicate that there is no significant correlation between the two sets of data, which could mean that there is no relation between students' self-perceived evaluation of their engagement in VR and their evaluation of the potential of VR in language

Vol.7, No.6, pp.18-74, December 2019
Published by ECRTD-UK
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

learning. At least for the participants in this study, positive engagement does not necessarily lead to positive views on VR adoption in language learning and vice versa.

			Z-score(ENG_Post)	Z-score(VR_EFL_Potential)
Spearman's rho	Z-score(ENG_Post)	Correlation Coefficient	1.000	.316
		Sig. (2-tailed)	.	.373
		N	10	10
	Z-score(VR_EFL_Potential)	Correlation Coefficient	.316	1.000
		Sig. (2-tailed)	.373	.
		N	10	10

Table 10: Results of Spearman's Rank Correlation Coefficient Test

Students' justifications for ratings of effort and tasks as determinants of using VR in EFL contexts referred to the Raven VR experience they were exposed to.

For example, one student pointed out the fact that the number of tasks were limited in the Raven VR; however, she felt focused on learning the tasks before getting into the VR experience. She wondered '…what if there were more tasks?'. Another student commented: 'more tasks means more effort' so if she has to learn so many 'technical' things before starting 'the actual learning' then she might get discouraged. Three students commented on time. They believed that more tasks or more effort meant more 'wasted time', which they cannot afford as advanced learners in the department. They would love to study using VR, but 'how much time do I need to learn English words…?' and would the time spent on VR tasks help them '…learn English faster'?

Both the quantitative and qualitative data analysed in this section indicated that though there is an increase in engagement levels when students used the VR experience, this increase could not be statistically linked to the other variables: Cognitive Load and Students' Behaviour. Moreover, students' evaluation of the VR content quality and the potential of using it in EFL settings could not be correlated to engagement. These findings will be further discussed in the following section 5.

DISCUSSION

The main objective of the current research was to investigate how female Saudi EFL students' engagement in VR experiences might be influenced by their cognitive load and the levels of achievement related to their behaviour and interactions inside the Virtual Reality world. This objective was broken down into three main questions:

- What is the effect of using VR environments on learners' engagement in EFL learning at KAU?
- Is there any correlation between students' cognitive load and engagement in virtual environments?

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

- Is there any correlation between students' behaviour (interaction) and engagement in virtual reality environments?

The findings discussed in the previous section (4) indicated that whereas the levels of students' engagement demonstrated an improvement when comparing the mean scores of the pre/post-engagement questionnaire, a relation could not be established between either students' cognitive load scores and engagement or their behavioural achievements in VR and their engagement. For the purpose of discussing the previously cited findings, this section will be divided into three sub-sections. Sub-section 5.1 examines the findings related to engagement in VR settings. The following sub-section, 5.2, discusses results demonstrating the potential links between engagement and cognitive load in VR. Sub-section 5.3 reflects on the significance of the findings that relate to correlations between engagement and students' interactions in VR.

Engagement in VR

In this study, the statistically documented *effect size* r= (-0.627664281) representing the moderate increase in students' engagement in the VR experience actually repeated the findings of previous research that has studied the relation between engagement and VR immersion (Madathil et al., 2016; Allocoat and Muhlenen, 2018; Filer and Holmes, 2018; Wilang and Soermphongsuwat, 2018; Gorham et al., 2019). This finding was further investigated through the interview questions related to the most/least engaging features of the VR experience. Therefore I have tried to link the increase in engagement scores to the VR aspects which students reflected on during the interview feedback.

For example, students seem to be highly engaged in VR when they feel a sense of 'presence' inside the poem's world and when they felt that the content of this world communicates something about the atmosphere of the literary work or its themes. The sense of 'play' and 'control' was appreciated by learners, as they found in the VR experience some aspects that made it familiar to them, like a game or movie. All of these aspects are more or less related to 'design', and in this particular case to instructional design, whereby the content of the experience should be organized in certain ways to allow for students' interaction and immersion and for them to be at the centre of the learning process and in total control of it.

Upon further investigating the qualitative feedback from the three first questions in the interview, we found that the students' rating of the quality of content (audio, visual, readable) in VR was considerably high, and though there was no significant correlation between the increase in engagement and the positive ratings of VR content, students seemed to appreciate the way content was presented.

I think the moderate *effect size* could be attributed to other design features that students identified in the same interview as the least engaging aspects of the VR experience. They acknowledged that the textual content as well as the absence of characters to interact with was something that they did not like.

The correlation between cognitive load and engagement

Our findings have demonstrated that the VR experience did not necessarily require a high cognitive load on the part of the students. Probably this could be attributed to factors such as the limited number of tasks that the students had to perform in VR and the simple and straightforward design of the VR experience, which was not demanding or stressful for students to navigate through. Though the temporal and mental dimensions of the cognitive scale scored the highest, this did not lead to any kind of an increased load or hinderance in performing the pre-defined tasks in VR. This decrease in cognitive load while using VR in educational settings was reported in previous research (Stone et al., 2011; Jiang, and Laidlaw, n.d.).

Moreover, the findings reported that there was no statistically significant correlation between engagement and cognitive load in section 3.2 coincide with previous research examining the potential links between the two variables in educational research that utilized technology (Jardina, and Chaparro, 2013; Vail et al., 2015; Mandala et al., 2018). Nonetheless, our results contradicted the findings of Lan et al. (2019) who reported an increase in engagement that correlated with an increase in cognitive load in collaborative learning environments. Most of the reviewed educational research did not establish a link between increased engagement and decreased cognitive load or vice versa. It did not discuss either a positive or negative change in both variables when compared to another. However, in the training and worker support fields, studies have demonstrated a correlation between lower cognitive loads and increased engagement (Tan et al., 2014).

When reflecting on students' feedback on the interview questions that related to the potential of using VR in EFL learning environments, we noticed that the participants believed that both the effort they exerted in accessing and navigating the VR experience and the number of tasks they have to execute (i.e. two representative dimensions of their cognitive load) might be significant in their decision to use VR in learning English as a Foreign Language.

The correlation between behaviour in VR and engagement

As for students' interaction inside the VR experience, which was measured using the Students' Behaviour Inside VR rubric, the findings outlined in the previous section showed that though the participants' performance of these tasks was successful, a correlation with engagement levels could not be established. Tasks which were based on physical interactions with the VR experience were done successfully with the higher mean (3.8) for task execution, which could be taken to indicate that the physical load of the VR was not high or did not affect students' performance in any way. More interactive tasks scored lesser minimum means that the previous ones (3.7), which could be taken to indicate that the more complex tasks demand more effort from students in VR, though their performance was not affected because of it. The reading tasks got a lower minimum score than either the physical or interactive ones (3.6) which points to both the complexity of the task and probably the quality of the readable content. Search or identify tasks achieved the lowest minimum mean (3.5), which sheds light on the potential of VR as a searchable browser that can be used in education. If students were not able to smoothly and quickly search and find objects in VR, then the demand of the experience can affect the performance or hinder learning behaviour.

CONCLUSION AND RECOMMENDATIONS

This case study mainly aims to investigate the effect of virtual reality environments on EFL students' level of engagement at KAU. To validate the research results, a mixed-research approach was implemented, i.e. using quantitative and qualitative research instruments. The quantitative data was collected throughout pre/post-engagement questionnaires and NASA Task Load Index survey whilst the qualitative data was gathered by interviewing each participant. Based on the research findings, it can be concluded that the levels of students' engagement demonstrated a significant improvement when comparing the mean scores of the pre/post-engagement questionnaires, as the significance of the difference of the mean scores between the pre/post-engagement questionnaire falls below the value of 0.05, and the effect size of that difference is moderate (-0.6). This indicated that unlike using any other virtual environments in KAU, using VR could have a positive role in increasing the engagement levels among EFL students at KAU. However, there was no clear evidence that the learners' general information, such as their academic level, ICT skill, prior use of VR, prior use of virtual learning environments in KAU, and preferred learning style, had a positive effect on their engagement, as significance was above the 0.05 value. Thus the results answer the first research question, 'What is the effect of using VR environments on learners' engagement in EFL learning at KAU?'.

Another important finding pertains to whether or not the cognitive load has an impact on students' engagement; the results reveal that the significance value for the correlation between the post-engagement and NASA TLX scores is 0.61, which is above the statistically significant value of 0.05. This could be interpreted as there being no sufficient statistical evidence to support my assumption that there is a correlation between students' engagement and cognitive load, at least among the sample used in this study. Therefore, it can be safely assumed that there is no significant correlation between the students' engagement and cognitive load scores (Appendix 9 and 10). In addition, considering both the correlation between engagement and the overall NASA TLX scores as well as between engagement and the individual and paired sub-scales of NASA TLX, there was no significant link between sets of data. Thus, the amount of cognitive load which the students exerted did not affect their engagement inside the VR experience, as students reported low cognitive load scores, which means they did not rate the VR experience as demanding. Therefore, the results answer the second research question, 'Is there any correlation between students' cognitive load and engagement in virtual environments?'.

In terms of the correlation between the Students' Behaviour Inside VR rubric data and the mean scores on their post-engagement questionnaire, there was no statistically sufficient evidence that could indicate a relation between students' engagement and their behaviour in VR. On this basis, the research findings indicated that there was no link (at least, in the context of the current implementation) between the two research variables of engagement and student's behaviour in VR even though the participants' performance of all assigned tasks was successful. This could be attributed to the limited number as well as the straightforward nature of the tasks. Thus the third research question, 'Is there any correlation between students' behaviour (interaction) and engagement in virtual reality environments?', is answered.

In other words, this research extends our current knowledge of the efficacy of VR environments on learners' engagement. It can be concluded that the levels of students' engagement demonstrated an improvement when using VR. However, a relation could not be established between either students' cognitive load scores and engagement or their behavioural achievements in VR and their engagement. Although the objective of this study has mainly investigated the effect of virtual reality environments on EFL learners' engagement on a small number of female EFL learners at King Abdul-Aziz University, its positive and significant results may reflect the potential of implementing VR systems in all EFL classes in Saudi Arabia due to its efficacy in raising students' engagement and motivation.

However, this study is not without limitations. It was conducted using a small number of students (n=10) enrolled in different academic levels from the Department of European Languages at King Abdul-Aziz University, Saudi Arabia. Thus additional studies are needed including a larger number of students as well as focusing on other colleges and other courses. Furthermore, the experiment was done using an enrichment course that is not included in the students' main curriculums. The reason for this was to validate the research results, assuming that if students had prior knowledge of the video used inside the VR, the findings might be affected. Therefore, further studies could be beneficial in investigating the effect of using course-related videos on students' engagement and learning outcomes.

REFERENCES

Ahmed Madini, A. and Alshaikhi, D. (2017). 'Virtual Reality for Teaching ESP Vocabulary: A Myth or A Possibility', *International Journal of English Language Education* [online] 5(2), p.111 –126. Available at: http://www.macrothink.org/journal/index.php/ijele/article/view/11993 (Accessed 8 Sep. 2019).

Al-Azawi, R., Albadi, A., Moghaddas, R. and Westlake, J. (2019). 'Exploring the Potential of Using Augmented Reality and Virtual Reality for STEM Education', *Communications in Computer and Information Science* [online], pp. 36–44. Available at: https://link.springer.com/chapter/10.1007/978-3-030-20798-4_4 (Accessed 4 Sep. 2019).

Alhalabi, W. (2016). 'Virtual reality systems enhance students' achievements in engineering education', *Behaviour & Information Technology* [online] 35(11), pp. 919–925. Available at: http://walhalabi.com/Papers/2016%20JUL%20Virtual%20reality%20systems%20enhance%20students%20achievements%20in%20engineering%20education.pdf(Accessed 4 Sep. 2019).

Allcoat, D. and von Mühlenen, A. (2018). 'Learning in virtual reality: Effects on performance, emotion and engagement', *Research in Learning Technology*, 26(0), pp.1–13.

Anderson, R.C., Spiro, R.J. and Montague, W.E. (1977). *Schooling and the acquisition of knowledge*. Hillsdale, N.J.: Lawrence Erlbaum Associates, New York.

Atkinson, R.C. and Shiffrin, R.M. (1968). 'Human Memory: A Proposed System and its Control Processes', *Psychology of Learning and Motivation* [online] pp. 89–195. Available at: https://www.sciencedirect.com/science/article/pii/S0079742108604223 (Accessed 31 May 2019).

Baddeley, A.D. (1983). 'Working Memory', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 302(1110), pp. 311–324.

Barker, P. (1993). 'Virtual reality: Theoretical basis, practical applications', *ALT-J,* 1(1), pp. 15–25.

Brockmyer, J.H., Fox, C.M., Curtiss, K.A., McBroom, E., Burkhart, K.M. and Pidruzny, J.N. (2009). 'The development of the Game Engagement Questionnaire: A measure of engagement in video game-playing'. *Journal of Experimental Social Psychology*, *45*(4), pp.624-634.

Bonner, E. and Reinders, H. (2018). 'Augmented And Virtual Reality In The Language Classroom: Practical Ideas', *Teaching English with Technology*, [online] 18(3), pp. 33–53. Available at: https://files.eric.ed.gov/fulltext/EJ1186392.pdf (Accessed 9 Jun. 2019).

Chen, Y., Smith, T.J., York, C.S. and Mayall, H.J. (2019). 'Google Earth Virtual Reality and expository writing for young English Learners from a Funds of Knowledge perspective', *Computer Assisted Language Learning*, pp. 1–24.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences.* New York Academic pressRoutledge.

Cooper, K.S. (2014). 'Eliciting Engagement in the High School Classroom', *American Educational Research Journal*, 51(2), pp. 363–402.

Cortina, J.M. (1993). 'What is coefficient alpha? An examination of theory and applications', *Journal of Applied Psychology*, 78(1), pp.98–104.

Curcio, I.D.D., Dipace, A. and Norlund, A. (2016). 'Virtual realities and education', *Research on Education and Media*, 8(2), pp. 60–68.

Dawala|Soermphongsuwat, J. (2018). *Virtual Reality for Undergraduate English Language Learners: A Formative Study*. [online] Online Submission. Available at: https://eric.ed.gov/?id=ED585191 [Accessed 4 Sep. 2019].

Di Serio, Á., Ibáñez, M.B. and Kloos, C.D. (2013). 'Impact of an augmented reality system on students' motivation for a visual art course', *Computers & Education*, [online] 68, pp. 586–596. Available at: https://www.sciencedirect.com/science/article/pii/S0360131512000590 (Accessed 8 Sep. 2019).

Dickey, M.D. (2005). 'Three-dimensional virtual worlds and distance learning: two case studies of Active Worlds as a medium for distance education', *British Journal of Educational Technology*, 36(3), pp. 439–451.

Filer, T. and Holmes, M. (n.d.). Using VR To Increase Student Engagement And employability Skills When Teaching Taxation. [online] Available at: https://apiar.org.au/wp-content/uploads/2018/03/13_ICTP_Feb_2018_BRR750_Edu_v2_3-36.pdf.

Fredricks, J.A., Blumenfeld, P.C. and Paris, A.H. (2004). 'School Engagement: Potential of the Concept, State of the Evidence', *Review of Educational Research*, 74(1), pp. 59–109.

Freina, L. (2015). *A Literature Review On Immersive Virtual Reality In Education: State Of The Art And Perspectives*. [online] Academia.edu. Available at: https://www.academia.edu/12273022/A_Literature_Review_on _Immersive_Virtual_Reality_in_Education_State _of_the _Art_and _Perspectives.

Gere, A.R., Wixon, V. and Stone, P. (1978). 'Practice into Theory', *The English Journal*, 67(9), p. 74-76.

Gerjets, P., Scheiter, K. and Cierniak, G. (2008). 'The Scientific Value of Cognitive Load Theory: A Research Agenda Based on the Structuralist View of Theories', *Educational Psychology Review*, [online] 21(1), pp. 43–54. Available at: https://link.springer.com/article/10.1007%2Fs10648-008-9096-1 (Accessed 9 Sep. 2019).

Gorham, T., Jubaed, S., Sanyal, T. and Starr, E.L. (2019). 'Assessing the efficacy of VR for foreign language learning using multimodal learning analytics', *Professional development in CALL: a selection of papers*, [online] pp. 101–116. Available at: https://files.eric.ed.gov/fulltext/ED593932.pdf (Accessed 8 Sep. 2019).

Great Schools Partnership (2016). *Student Engagement Definition*. [online] The Glossary of Education Reform. Available at: https://www.edglossary.org/student-engagement/ (Accessed 26 Apr. 2019).

Hart, S. (1986). *Task Load Index (Nasa-Tlx) V 1.O*. [online] Available at: https://ntrs.nasa.gov/archive/nasa/casi.ntrs.nasa.gov/20000021488.pdf (Accessed 8 Sep. 2019).

Heilig, M.L. (1962). Sensorama simulator - Heilig, Morton L. *Freepatentsonline.com*. [online] Available at: http://www.freepatentsonline.com/3050870.html [Accessed 9 Sep. 2019].

Huang, H.-M., Rauch, U. and Liaw, S.-S. (2010). 'Investigating learners' attitudes toward virtual reality learning environments: Based on a constructivist approach', *Computers & Education*, 55(3), pp. 1171–1182.

Hussein, M. and Nätterdal, C. (2015). *The Benefits of Virtual Reality in Education A Comparison Study*. [online] Available at: https://gupea.ub.gu.se/bitstream/2077/39977/1/gupea_2077_39977_1.pdf (Accessed 4 Sep. 2019).

Jardina, J.R. and Chaparro, B.S. (2013). 'Usability, Engagement, and Satisfaction of Two e-Textbook Applications', *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 57(1), pp. 482–486.

Jensen, L. and Konradsen, F. (2017). 'A review of the use of virtual reality head-mounted displays in education and training', *Education and Information Technologies*, 23(4), pp. 1515–1529.

Jiang, E. and Laidlaw, D. (n.d.). *Practicing in Virtual Reality Improves Mental Rotation Ability: Lower Scorers Benefit More*. [online] Available at: http://cs.brown.edu/research/pubs/theses/ugrad/2019/jiang.elaine.pdf (Accessed 8 Sep. 2019).

Kahu, E.R. (2013). 'Framing Student Engagement in Higher Education', *Studies in Higher Education*, [online] 38(5), pp. 758–773. Available at: https://eric.ed.gov/?id=EJ1012777 (Accessed 9 Sep. 2019).

Kapil Chalil Chalil Madathil, Kristin Kelly Frady, Gramopadhye, A.K., Bertrand, J. and Hartley, R.S. (2016). *Integrating Case Studies in an Online Asynchronous Learning Environment: An Empirical Study to Evaluate the Effect on Community College Student Learning Outcomes and Engagement when Case Studies are Presented Using Virtual Reality*. [online] *Asee.org*. Available at: https://peer.asee.org/integrating-case-studies-in-an-online-asynchronous-learning-environment-an-empirical-study-to-evaluate-the-effect-on-community-college-student-learning-outcomes-and-engagement-when-case-studies-are-presented-using-virtual-reality.

Kavanagh, E.J., Litchfield, C. and Osborne, J. (2020). 'Bournemouth University Research Online [BURO] - Virtual technologies as tools of maltreatment: safeguarding in digital spaces', *Bournemouth.ac.uk*. [online] Available at: http://eprints.bournemouth.ac.uk/32435/.

Kelley, K. and Preacher, K.J. (2012). 'On effect size', *Psychological Methods*, [online] 17(2), pp. 137–152. Available at:

https://www3.nd.edu/~kkelley/publications/articles/Kelley_and_Preacher_Psychological_Methods_2012.pdf (Accessed 9 Sep. 2019).

Lan, C., Sheng, M., Hsu, Y. and Shuie, Y. (2019). 'Influence of Online and Face-to-face Collaboration and Learning Style on Cognitive Load and Engagement in a Health Introductory Course', *Proceedings of the third International Conference on Medical and Health Informatics 2019* pp. ACM142-148.

Lange, B.S., Requejo, P., Flynn, S.M., Rizzo, A.A., Valero-Cuevas, F.J., Baker, L. and Winstein, C. (2010). 'The Potential of Virtual Reality and Gaming to Assist Successful Aging with Disability', *Physical Medicine and Rehabilitation Clinics of North America*, [online] 21(2), pp. 339–356. Available at: https://sites.temple.edu/walkerbq/files/2015/12/The-Potential-of-Virtual-Reality-and-Gaming-to-Assist-Successful-Aging-with-Disability.pdf (Accessed 8 Sep. 2019).

Learning Technology for Education Challenges (2019). *Learning Technology for Education Challenges - 8th International Workshop, LTEC 2019,* Zamora, Spain, July 15–18, 2019. *Proceedings | Springer*. [online] Springer.com. Available at: https://www.springer.com/gp/book/9783030207977 (Accessed 4 Sep. 2019).

Lee, H. (2013). 'Measuring cognitive load with electroencephalography and self-report: focus on the effect of English-medium learning for Korean students', *Educational Psychology*, 34(7), pp. 838–848.

Lee, J., Song, H.D. and Hong, A.J. (2019). 'Exploring factors, and indicators for measuring students' sustainable engagement in e-learning. *Sustainability*, *11*(4), p.985.

Lee, S.H. (Mark), Sergueeva, K., Catangui, M. and Kandaurova, M. (2017). 'Assessing Google Cardboard virtual reality as a content delivery system in business classrooms', *Journal of Education for Business*, 92(4), pp. 153–160.

Mahender Mandala, Schunn, C.D., Dow, S., Goldberg, M., Pearlman, J., Clark, W. and Mena, I. (2018). 'Impact of collaborative team peer review on the quality of feedback in engineering design projects', *The International journal of engineering education*, [online] 34(4), pp. 1299–1313. Available at: https://dialnet.unirioja.es/servlet/articulo?codigo=6878629 (Accessed 8 Sep. 2019).

Makransky, G. and Lilleholt, L. (2018). 'A structural equation modeling investigation of the emotional value of immersive virtual reality in education', *Educational Technology Research and Development*, [online] 66(5), pp. 1141–1164. Available at: https://link.springer.com/article/10.1007%2Fs11423-018-9581-2 (Accessed 22 Jul. 2019).

Maria, A. (1346). *Effects of Text, Audio and Learner Control on Text-Sound Association and Cognitive Load of EFL Learners*. [online] Proquest Llc. Website: Http://www.Proquest.Com/En-Us/Products/Dissertations/Individuals.Shtml -00-00. Available at: https://eric.ed.gov/?id=ED569094 (Accessed 4 Sep. 2019).

Maria, A., Bernal, E., Nelson, B., Savenye, W. and Atkinson, R. (2014). *Effects of text, audio and learner control on text-sound association and cognitive load of EFL learners*. [online] Available at: https://repository.asu.edu/attachments/135019/content/EncisoBernal_asu_0010E_13964.pdf (Accessed 4 Sep. 2019).

Martín-Gutiérrez, J. (2016). 'Virtual Technologies Trends in Education', *EURASIA Journal of Mathematics, Science and Technology Education*, 13(1). 469–486

Mikropoulos, T.A., Chalkidis, A., Katsikis, A. and Emvalotis, A. (1996). 'Students' Attitudes Towards Educational Virtual Environments', *Education and Information Technologies*, [online] 3(2), pp. 137–148. Available at: https://link.springer.com/article/10.1023/A:1009687025419 (Accessed 8 Sep. 2019).

Oxfordlearnersdictionaries.com. (2019). *virtual-reality noun - Definition, pictures, pronunciation and usage notes | Oxford Advanced American Dictionary at OxfordLearnersDictionaries.com*. [online] Available at: https://www.oxfordlearnersdictionaries.com/definition/american_english/virtual-reality (Accessed 4 Sep. 2019).

Pacampara, M. and Castaneda, L. (2017). *Real World Research: Equity with Virtual Reality in Classrooms*. [online] Available at: https://conference.iste.org/uploads/ISTE2018/HANDOUTS/KEY_110798600/EquitywithVirtualRealityinClassrooms_RP.pdf (Accessed 8 Sep. 2019).

Pantelidis, V.S. (2009). 'Reasons to Use Virtual Reality in Education and Training Courses and a Model to Determine When to Use Virtual Reality', *Themes in Science and Technology Education*, [online] 2(1–2), pp. 59–70. Available at: http://earthlab.uoi.gr/theste/index.php/theste/article/view/22.

Peixoto, B., Pinto, D., Krassman, A., Melo, M., Cabral, L. and Bessa, M. (2019). 'Using Virtual Reality Tools for Teaching Foreign Languages' *In World Conference on Information Systems and Technologies*. *Springer Cham* pp. 581-588.

Peterson, K. and Stone, B.. (2019). 'From Theory to Practice', International Journal of the Whole Child, 4(1), pp. 67-74.

Repetto, C., Cipresso, P. and Riva, G. (2015). 'Virtual action and real action have different impacts on comprehension of concrete verbs', *Frontiers in Psychology*, 6. 1–9

Schwienhorst, K. (2002). 'The State of VR: A Meta-Analysis of Virtual Reality Tools in Second Language Acquisition', *Computer Assisted Language Learning*, 15(3), pp. 221–239.

Seng, C., Tan, S., Schöning, J., Luyten, K. and Coninx, K. (2014). One of a CHInd. [online] Available at: https://doclib.uhasselt.be/dspace/bitstream/1942/17019/1/CHI%202014%20p71.pdf (Accessed 8 Sep. 2019).

Sibilio, M, Aiello, P. (2012). 'A Constructivist Approach to Virtual Reality for Experiential Learning', *E-Learning and Digital Media*, [online] 9(3), pp. 317–324. Available at: https://eric.ed.gov/?id=EJ986859 (Accessed 8 Sep. 2019).

Standen, P.J. and Brown, D.J. (2006). 'Virtual reality and its role in removing the barriers that turn cognitive impairments into intellectual disability', *Virtual Reality*, 10(3–4), pp. 241–252.

Steuer, J. (1992). 'Defining Virtual Reality: Dimensions Determining Telepresence', *Journal of Communication*, 42(4), pp. 73–93.

Stone, R.T., Watts, K.P., Zhong, P. and Wei, C.-S. (2011). 'Physical and Cognitive Effects of Virtual Reality Integrated Training', *Human Factors: The Journal of the Human Factors and Ergonomics Society*, [online] 53(5), pp. 558–572. Available at: http://www.lincolnelectric.com/en-us/equipment/training-equipment/Documents/Physiological_and_Cognitive_Effects_of_Virtual_Reality_Integrated_Training_13July2011.pdf (Accessed 4 Sep. 2019).

Sweller, J. (1988). 'Cognitive Load During Problem Solving: Effects on Learning', *Cognitive Science*, 12(2), pp. 257–285.

Trowler, V. (2010). *Student engagement literature review*. [online] Available at: https://www.heacademy.ac.uk/system/files/studentengagementliteraturereview_1.pdf.

Tzone-I Wang and Wu, Y.H. (2014). *The effects of using synonyms annotation for English video subtitles on content understanding, vocabulary learning, and cognitive loading*. [online] 國立成功大學. Available at: https://researchoutput.ncku.edu.tw/zh/publications/the-effects-of-using-synonyms-annotation-for-english-video-subtit (Accessed 4 Sep. 2019).

Vail, A.K., Boyer, K.E., Wiebe, E.N. and Lester, J.C. (2015). 'The Mars and Venus Effect: The Influence of User Gender on the Effectiveness of Adaptive Task Support', *Lecture Notes in Computer Science*, [online] pp. 265–276. Available at: https://link.springer.com/chapter/10.1007/978-3-319-20267-9_22 (Accessed 8 Sep. 2019).

Youngblut, C. (1998). *INSTITUTE FOR DEFENSE ANALYSES IDA Educational Uses of Virtual Reality Technology*. [online] Available at: http://papers.cumincad.org/data/works/att/94ea.content.pdf.

Yu, Z., Chen, W., Kong, Y., Sun, X.L. and Zheng, J. (2014). 'The Impact of Clickers Instruction on Cognitive Loads and Listening and Speaking Skills in College English Class', *PLoS ONE*, 9(9), p.1–8.

Appendices

Appendix 1: Engagement questionnaire

1. I follow the instructor's rules when I am studying in virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

2. I put forth effort when I study in virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

3. I carefully listen to/read content delivered via virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

4. I do not want to log out when I join a virtual learning environment.
- totally agree
- agree

- neutral
- disagree
- totally disagree

5. I request assistance in case I need help in virtual learning environments.

- totally agree
- agree
- neutral
- disagree
- totally disagree

6. I respond to questions and requests in virtual learning environments.

- totally agree
- agree
- neutral
- disagree
- totally disagree

7. I perform assigned activities in virtual learning environments.

- totally agree
- agree
- neutral
- disagree
- totally disagree

8. Being in virtual learning environments enhances my interest in learning.

- totally agree
- agree
- neutral
- disagree
- totally disagree

9. I am highly motivated in virtual learning environments.

- totally agree
- agree
- neutral
- disagree
- totally disagree

10. It is useful to study in virtual learning environments.

- totally agree
- agree
- neutral
- disagree
- totally disagree

11. It is very interesting to be in virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

12. I look forward to educational experiences in virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

13. I am satisfied with the virtual learning environment I have experienced.
- totally agree
- agree
- neutral
- disagree
- totally disagree

14. The virtual learning experience seems real.
- totally agree
- agree
- neutral
- disagree
- totally disagree

15. I can derive new ideas and interpretations from the knowledge delivered via virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

16. I have the ability to analyse ideas, experiences, and theories about the knowledge presented in virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

17. I can evaluate the knowledge presented in virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

18. I can apply the knowledge I have learned virtual environments to real problems or new situations.
- totally agree
- agree
- neutral
- disagree
- totally disagree

19. I approach the knowledge delivered via virtual learning environments with a new perspective.
- totally agree
- agree
- neutral
- disagree
- totally disagree

20. I can easily remember the information presented in virtual learning environments.
- totally agree
- agree
- neutral
- disagree
- totally disagree

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by *ECRTD-UK*
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Appendix 2: Learners' profile

1. Level
- 1st
- 2nd
- 3rd
- 4th
- 5th
- 6th
- 7th
- 8th

2. ICT SKILL
- Advanced
- Intermediate
- Beginner

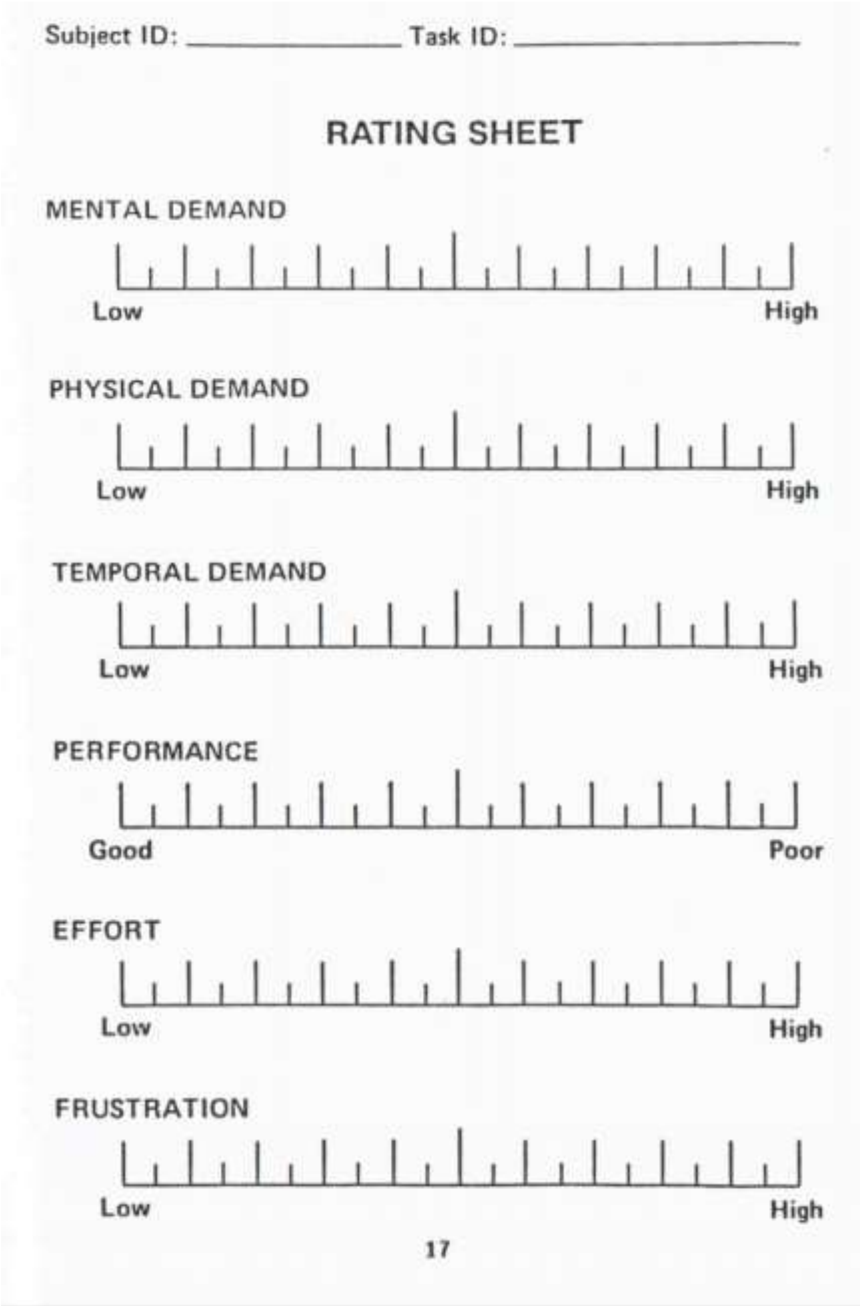
3. Have you used VR headsets before?
- Yes
- No

4. Have you used online learning environments in your courses?
- Yes
- No

5. Your preferred learning style is …
- Visual: by seeing and observing things, including pictures, diagrams, written directions
- Reading/Writing: you are drawn to expression through writing, reading articles on the internet, writing in diaries, looking up words in the dictionary and searching the internet
- Auditory: you learn better through lectures than reading written notes, and often use your own voice to reinforce new concepts and ideas
- Kinaesthetic: you learn through experiencing or doing things

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by *ECRTD-UK*
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Appendix 3: NASA TLX tool



Subject ID: ______________ Task ID: ____________________

RATING SHEET

MENTAL DEMAND

Low High

PHYSICAL DEMAND

Low High

TEMPORAL DEMAND

Low High

PERFORMANCE

Good Poor

EFFORT

Low High

FRUSTRATION

Low High

17

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by ***ECRTD-UK***
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Effort or Performance	Temporal Demand or Frustration
Temporal Demand or Effort	Physical Demand or Frustration
Performance or Frustration	Physical Demand or Temporal Demand
Physical Demand or Performance	Temporal Demand or Mental Demand

15

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by *ECRTD-UK*
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Frustration or Effort	Performance or Mental Demand
Performance or Temporal Demand	Mental Demand or Effort
Mental Demand or Physical Demand	Effort or Physical Demand
Frustration or Mental Demand	

16

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by *ECRTD-UK*
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Appendix 4: Instruction sheet

Goal	Task
Access the RAVEN EXPERINCE	Locate Experience in Oculus Rift Library
	Click on Experience Icon
	Wear Oculus Rift Headset
	Hold Touch devices on both hands
	Read the EXPERIENCE directions
	Click on any button on touch device to access the RAVEN EXPERIENCE
Interact with the RAVEN EXPERIENCE	In first landing location, rotate headset in all directions
	Identify teleporting pads on the ground and use them to move from one place to another using (touch pointer)
	Teleport to where the poem is displayed on the wall using (touch pointer)
	Teleport to where the Lenore painting is displayed using (touch pointer)
	Teleport to the couch using (touch pointer)
	When the RAVEN accesses the room, Teleport to the door using (touch pointer)

Appendix 5: Students' Behaviour in VR rubric

STUDENT ID	**( )**						
Goal	Task	Success Indicator 4 Successfully execute task using assigned path	Less Probable Action 3 Successfully execute task using assigned path but some sub-tasks cannot be fulfilled	Less Probable Action 2 Successfully execute task using alternative path	Less Probable Action 1 Could not execute task using the assigned path	Failure Indicator 0 Could not figure out how to access assigned path	Notes
Access the RAVEN EXPERINCE	Locate Experience in Oculus Rift Library						
	Click on Experience Icon						-
	Wear Oculus Rift Headset						
	Hold Touch devices on both hands						
	Read the EXPERIENCE directions						
	Click on any button on touch device to access the RAVEN EXPERIENCE						
Interact with the RAVEN EXPERINCE	In first landing location, rotate headset in all directions						
	Identify teleporting pads on the ground and use them to move from one place to another using (touch pointer)						
	Teleport to where the poem is displayed on the wall using (touch pointer)						
	Teleport to where the Lenore painting is displayed using (touch pointer)						
	Teleport to the couch using (touch pointer)						
	When the RAVEN accesses the room, teleport to the door using (touch pointer)						

Appendix 6: Interview instrument

Rate the following VR components as you experienced them:

1. Quality of audio in VR
- Excellent
- Good
- Fair
- Poor

2. Quality of visual content in VR
- Excellent
- Good
- Fair
- Poor

3. Quality of readable content in VR
- Excellent
- Good
- Fair
- Poor

Answer the following questions

4. What was the most engaging element in the VR experience? Why?

……………………………………………………………………………………………
…………………………………………………………………………………………

5. What was the least engaging element in the VR experience? Why?

……………………………………………………………………………………………
…………………………………………………………………………………………

6. Do you think that the effort you exert while using VR can prevent you from using it to learn English as a foreign language? Why?

……………………………………………………………………………………………
…………………………………………………………………………………………

7. Do you think that the number of tasks you have to do in VR can prevent you from using it to learn English as a foreign language? Why?

……………………………………………………………………………………………
…………………………………………………………………………………………

Appendix 7: First consent form

Study Title: Exploring English as Foreign Language (EFL) Learners' Engagement in Virtual Reality Environments: A Case Study on King Abdul-Aziz University (KAU) Learners in Saudi Arabia

Researcher: Ehtiram Sharyan
Supervisor: Martine Van Driel

I am a researcher at the University of Birmingham, enrolled in the Applied Linguistics MA program. I am planning to conduct a research study, which I invite you to take part in. This form has important information about the reasons for doing this study, what we will ask you to do if you decide to be in this study, and the way we would like to use information about you if you choose to be in the study.

You are being asked to participate in a research study that investigates Saudi EFL learners' engagement in Virtual Reality (VR) educational settings and how this engagement might be affected by their cognitive load and behaviour inside VR environments.
Your participation in the study involves the following:

1. Filling in a 20-item engagement questionnaire (pretest) which will take about 3-5 minutes.

Study time: Participation will take approximately 5 minutes covering all the tasks mentioned above.

Study location: All study procedures will take place at the Deanship of E-learning and Distance Education (E-learning Innovation Lab) in King Abdul-Aziz University, Building 66/lab

Please note the following:
1. Your participation in this study does not involve any physical or emotional risk to you beyond that of everyday life.
2. The confidentiality of the information we collect from you will be maintained and the data will be used only for analysis and investigation of the topic under study. Names, and any general information related to you are not going to be disclosed.
3. Participation in this study is voluntary. You do not have to answer any question you do not want to answer. If at any time and for any reason you would prefer not to participate in this study, please feel free not to. If at any time you would like to stop participating, please tell me. We can take a break, stop and continue at a later date, or stop altogether. You may withdraw from this study at any time, and you will not be penalized in any way for deciding to stop participation.
4. If you decide to withdraw from this study, the researcher will ask you if the information already collected from you can be used in the current study.
5. If you have questions, you are free to ask them now. If you have questions later, you may contact the researcher via email:

ehteram.sharyan@gmail.com

Consent
I have read this form and the research study has been explained to me. I have been given the opportunity to ask questions and my questions have been answered. If I have additional questions, I have been told whom to contact. I agree to participate in the research study described above.

____
Participant's Name (printed)

____

Participant's Signature Date

Appendix 8: Second consent form

Study Title: Exploring English as Foreign Language (EFL) Learners' Engagement in Virtual Reality Environments: A Case Study on King Abdul-Aziz University (KAU) Learners in Saudi Arabia

Researcher: Ehtiram Sharyan
Supervisor: Martine Van Driel

I am a researcher at the University of Birmingham, enrolled in the Applied Linguistics MA program. I am planning to conduct a research study, which I invite you to take part in. This form has important information about the reasons for doing this study, what we will ask you to do if you decide to be in this study, and the way we would like to use information about you if you choose to be in the study.

You are being asked to participate in a research study that investigates Saudi EFL learners' engagement in Virtual Reality (VR) educational settings and how this engagement might be affected by their cognitive load and behaviour inside VR environments.

Your participation in the study involves the following:

2. Filling in a 20-item engagement questionnaire (pretest) which will take about 3-5 minutes.
3. Participating in the VR experience which involves you wearing a VR headset, accessing a specific VR experience, and performing a number of predefined tasks which are included in the experiment's rubric. Your interaction inside the VR experiment will be video-recorded.
4. Filling in the NASA TLX cognitive load index

5. Filling in a 20-item engagement questionnaire (post-test) which will take about 3-5 minutes.
6. Participating in an interview with the researcher.

Study time: Participation will take approximately 25-35 minutes covering all the tasks mentioned above.

Study location: All study procedures will take place at the Deanship of E-learning and Distance Education (E-learning Innovation Lab) in King Abdul-Aziz University, Building 66/lab

Please note the following:

6. Your participation in this study does not involve any physical or emotional risk to you beyond that of everyday life.
7. The confidentiality of the information we collect from you will be maintained and the data will be used only for analysis and investigation of the topic under study. Names and any general information related to you are not going to be disclosed.
8. Participation in this study is voluntary. You do not have to answer any question you do not want to answer. If at any time and for any reason, you would prefer not to participate in this study, please feel free not to. If at any time you would like to stop participating, please tell me. We can take a break, stop and continue at a later date, or stop altogether. You may withdraw from this study at any time, and you will not be penalized in any way for deciding to stop participation.
9. If you decide to withdraw from this study, the researcher will ask you if the information already collected from you can be used in the current study.
10. If you have questions, you are free to ask them now. If you have questions later, you may contact the researcher via email:
ehteram.sharyan@gmail.com

Consent

I have read this form and the research study has been explained to me. I have been given the opportunity to ask questions and my questions have been answered. If I have additional questions, I have been told whom to contact. I agree to participate in the research study described above.

I consent that:

- My VR interaction is video-recorded
- I participate in the research interview

____

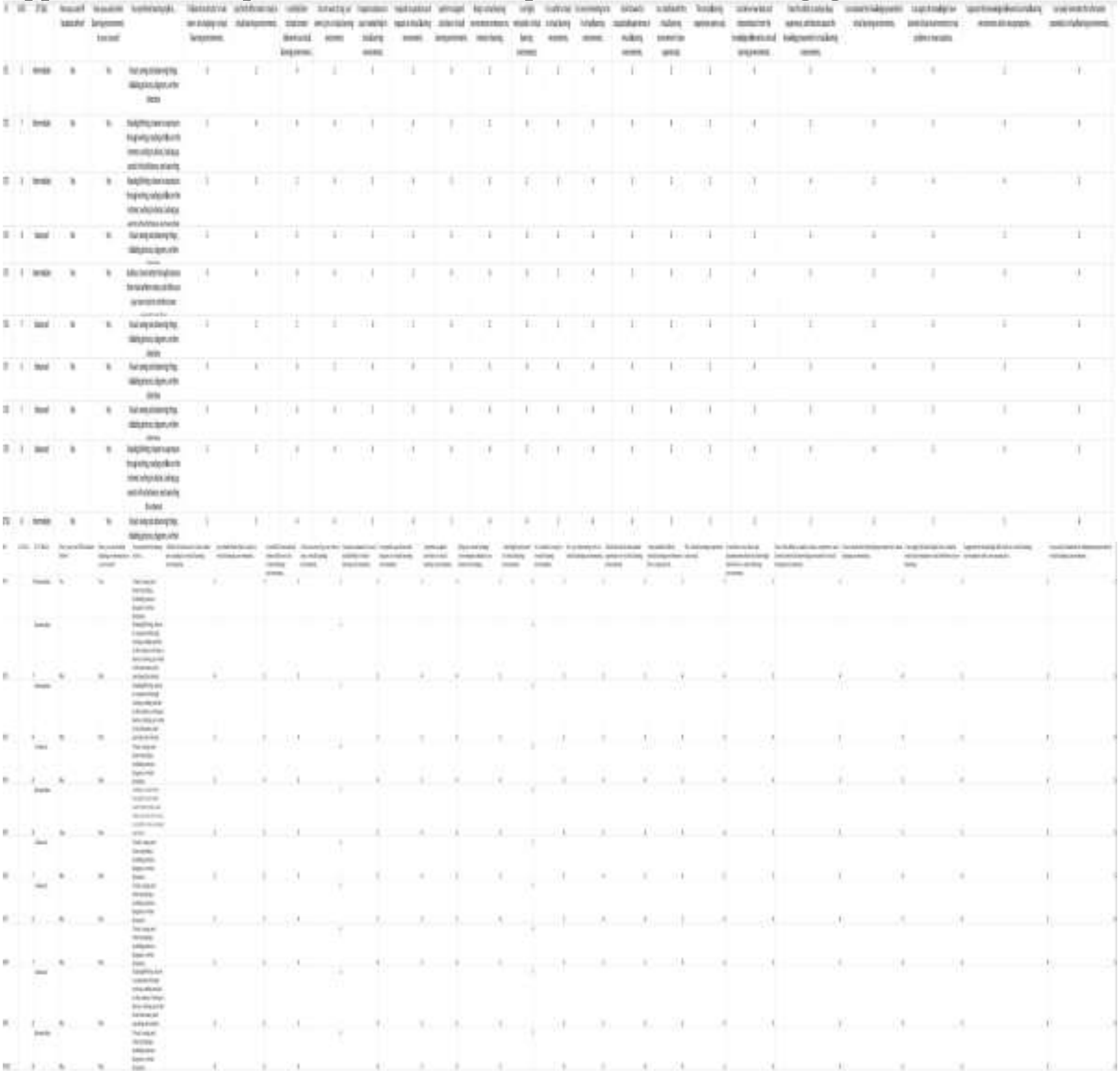
Participant's Name (printed)

____

Participant's Signature   Date

British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by *ECRTD-UK*
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Appendix 9: A Snapshot of Students' Responses in Pre/Post Engagement Questionnaire



British Journal of English Linguistics
Vol.7, No.6, pp.18-74, December 2019
Published by *ECRTD-UK*
Print ISSN: 2055-6063(Print), Online ISSN: 2055-6071(Online)

Appendix 10: A Snapshot of Students' Responses in NASA TLX (Task Load Index)

ID	tlx_Score	scale_mental	scale_physical	scale_temporal	scale_perform	scale_effort	scale_frustration	workload_mental	workload_physical	workload_temporal	workload_perform	workload_effort	workload_frustration	workloadWeighted_mental	workloadWeighted_physical	workloadWeighted_temporal	workloadWeighted_perform	workloadWeighted_effort	workloadWeighted_frustration	timeStamp
ST1	9.16	1	3	14	8	17	12	0.17	0.17	0.17	0.17	0.17	0.17	0.17	0.5	2.33	1.33	2.83	2	Sunday Jul 7 7:50:13 GMT+0300 2019
ST2	3.01	7	1	2	1	6	1	0.17	0.17	0.17	0.17	0.17	0.17	1.17	0.17	0.33	0.17	1	0.17	Monday Jul 8 7:30:06 GMT+0300 2019
ST3	12.17	21	2	21	4	9	16	0.17	0.17	0.17	0.17	0.17	0.17	3.5	0.33	3.5	0.67	1.5	2.67	Monday Jul 8 8:10:02 GMT+0300 2019
ST4	5.16	7	2	0	12	8	2	0.17	0.17	0	0.17	0.17	0.17	1.17	0.33	0	2	1.33	0.33	Monday Jul 8 8:30:03 GMT+0300 2019
ST5	9.34	8	2	19	13	7	7	0.17	0.17	0.17	0.17	0.17	0.17	1.33	0.33	3.17	2.17	1.17	1.17	Monday Jul 8 8:45:53 GMT+0300 2019
ST6	8.83	17	3	8	2	4	19	0.17	0.17	0.17	0.17	0.17	0.17	2.83	0.5	1.33	0.33	0.67	3.17	Monday Jul 8 9:10:49 GMT+0300 2019
ST7	1.17	1	1	2	0	3	0	0.17	0.17	0.17	0	0.17	0	0.17	0.17	0.33	0	0.5	0	Monday Jul 8 9:26:45 GMT+0300 2019
ST8	9.17	16	3	22	10	2	2	0.17	0.17	0.17	0.17	0.17	0.17	2.67	0.5	3.67	1.67	0.33	0.33	Monday Jul 8 10:05:35 GMT+0300 2019
ST9	9.67	9	1	18	3	13	14	0.17	0.17	0.17	0.17	0.17	0.17	1.5	0.17	3	0.5	2.17	2.33	Wed Jul 10 9:58:07 GMT+0300 2019
ST10	4.84	7	6	13	1	2	0	0.17	0.17	0.17	0.17	0.17	0	1.17	1	2.17	0.17	0.33	0	Wed Jul 10 12:40:58 GMT+0300 2019

Appendix 11: Video Recordings for participants inside VR Experience

https://drive.google.com/drive/folders/1e8urtP5k8uKRQ2AXdNur7yvtAW9ePoEk?usp=sharing