Providing timely support to students in Educational Virtual Worlds

A Thesis submitted for the degree of

Masters in Research

to

Department of Computing

Faculty of Science and Engineering

Macquarie University

NSW, Australia

By Anupam Makhija

Supervisor: Professor Deborah Richards

Date of Submission: 18 Oct 2018

Acknowledgement

I would like to express my sincere gratitude to my supervisor Professor Deborah Richards for her continued guidance, support and for not accepting less than excellence for me. I am also grateful to Meredith Porte for helping me in overcoming technical obstacles I faced during my research journey. Last but not the least, I would like to thank all the students who participated in my study to achieve my goal.

LIST OF PUBLICATIONS

- Makhija A., Richards D., De Haan J., Dignum F., Jacobson M.J. (2018) The Influence of Gender, Personality, Cognitive and Affective Student Engagement on Academic Engagement in Educational Virtual Worlds. In: Penstein Rosé C. et al. (eds) Artificial Intelligence in Education. AIED 2018. Lecture Notes in Computer Science, vol 10947, 297-310. Springer, Cham

- Makhija A., Richards D., Caballé S., Conesa J. (2019) Providing Timely Support to Students in Educational Virtual Worlds. In: Xhafa F., Barolli L., Greguš M. (eds) Advances in Intelligent Networking and Collaborative Systems. INCoS 2018. Lecture Notes on Data Engineering and Communications Technologies, vol 23, 406-419. Springer, Cham

  (awarded as the best paper award of the workshops of the main conference INCoS-2018)

- Ranjbartabar H., Richards D., Makhija A., Jacobson M.J. (2018) Students' Responses to a Humanlike Approach to Elicit Emotion in an Educational Virtual World. In: Penstein Rosé C. et al. (eds) Artificial Intelligence in Education. AIED 2018. Lecture Notes in Computer Science, vol 10948, 291-295. Springer, Cham

- Richards D., Bilgin A.A., Ranjbartabar H., Makhija A. (2018) Towards Realtime Adaptation: Uncovering User Models from Experimental Data. In: Yoshida K., Lee M. (eds) Knowledge Management and Acquisition for Intelligent Systems. PKAW 2018. Lecture Notes in Computer Science, vol 11016, 46-60. Springer, Cham

- Virtual Humans to Manage Epistemic Emotions in Educational Virtual Worlds: Gender, Personality and Academic Ability Differences; Deborah Richards, Hedieh Ranjbartabar and Anupam Makhija, Accepted as an extended abstract (2 pages 325-326) paper at The 18th International Conference on Intelligent Virtual Agents (IVA-2018)

DECLARATION: This thesis contains some of the content that appears in above publications.

ABSTRACT

Educational virtual worlds (EVWs) are immersive educational tools that have the potential to engage students using interactive virtual avatars in the simulated multi-media environments. As with any learning approach, EVWs should have the capacity to adapt according to individual student needs and differences such as their personality characteristics, engagement levels, navigational behaviour, and emotional level. Based on such factors, the EVWs should be capable of providing tailored support and feedback when required. This experimental study conducted with 144 undergraduate students investigates the relationship between individual learner factors, including personality, cognitive and affective engagement, and the value of providing either tailored, generic or no hints on academic performance in an EVW. Results suggest that extravert and agreeable personality traits are positively linked to academic performance and affective and cognitive engagement levels. Contrary to our hypothesis, participants who were provided with hints did not perform better than the control group in performance. Further studies are needed to gain deeper insights to design appropriate support strategies within EVW.

Keywords: Educational Virtual World, Intelligent Virtual Agents, Personality Dimensions, Affective Engagement, Cognitive Engagement, Academic engagement, Navigational Behaviour, Epistemic Emotions.

Table of Contents

Table of Figures

Table of Tables

CHAPTER 1: INTRODUCTION

In today's advanced era of digital education technology, Educational Virtual Worlds (EVWs) is providing an educational platform with the potential of keeping learners more engaged in their learning process. Educational Virtual Worlds can help students learn new concepts more effectively in various learning domains (Dieterle, 2009; Eschenbrenner, Nah, & Siau, 2008; Warburton, 2009). EVWs provide an interactive and stimulating space featuring simulation and visualisation, with the possibility of offering a new dimension in online learning and teaching (Calongne, 2008).

An Educational Virtual World is a new generation of computer-based educational systems that provide information and social space integrating multiple tools where students can not only be active members but also be the actors and contributors in the given space (Dillenbourg, Schneider, & Synteta, 2002). EVWs allow learning in a safe and fun environment that can be supported by intelligent virtual agents (i.e. non-player characters) who can play different roles such as mentor, teacher, trainer and helper and personally interact with learners to demonstrate complex skills, provide feedback or answer their queries (Paiva et al., 2005; Rickel & Johnson, 1997).

1.1 Aims and Significance

Every student exhibits distinct personality characteristics (Chamorro-Premuzic & Furnham, 2008), engagement levels (Hoff & Lopus, 2014) and emotional responses (Shen, Wang, & Shen, 2009) that can affect their academic performance and behaviour in various educational settings. An accurate understanding of these academic predictors and individuals' needs can help tailor the EVWs and improve the learning experiences of learners within an EVW.

This research seeks to contribute to our understanding of the value of providing tailored support to students using an EVW, particularly with respect to improving their academic achievement. Furthermore, this research explores if there are differences in the responses of participants based on their personality, cognitive engagement or emotional engagement with their studies.

1.2 Research Questions and Approach

This research aims to use data about the learner to develop approaches to provide appropriate support in an EVW. We thus sought to answer:

- what learner data might be relevant in providing support in an EVW?
- what sort of support is appropriate to different learners?

To formulate low-level research questions and design an experimental study using an EVW to answer the questions, it was necessary to investigate the literature further. Thus, following the Literature Review in Chapter 2, Chapter 3 commences with the low-level research questions.

1.3 Thesis Structure

This thesis is structured as follows: Chapter 2 discusses individual factors such as engagement, personality, emotions and navigational behaviour that can affect learning within an EVW platform. Chapter 3 describes the materials and methodology used for data collection in this study. Chapter 4 analyses the results and Chapter 5 discusses the findings of the data analysis followed by the conclusion and future work. The References and Appendices are provided in the last sections of this thesis.

CHAPTER 2: LITERATURE REVIEW

This chapter reviews the literature to understand the key factors that can influence the learning in EVWs. A review of the literature can provide an understanding of aspects of the learner and different ways to provide support to students within an EVW. Given the extensive research in this area, and the focus of this thesis, only selected key factors are presented below: student engagement, personality, navigation and learner's emotional state (Cruz-Benito, Therón, García-Peñalvo, & Lucas, 2015; Rudolph P Darken & Sibert, 1996; Tian et al., 2014; Yee, Harris, Jabon, & Bailenson, 2011). Section 2.1 presents an overview of the relationship between academic success and these four factors. Section 2.2 considers different ways to provide support to the learners in an EVW. Section 2.3 discusses the current gaps.

2.1 Factors affecting academic performance

This section focuses on the connection of learning with factors such as engagement (section 2.1.1) and personality (section 2.1.2) that have been found to be significantly correlated with the academic success of the individual. Given that navigation is a key aspect of using an EVW, section 2.1.3 discusses navigation and its effect on performance. Section 2.1.4 reviews epistemic emotions due to their relevance to learner and impact on engagement and academic performance.

2.1.1 Engagement

The user engagement plays a vital role in face to face classrooms as well as in online educational environments such as EVWs (Beer, Clark, & Jones, 2010; Hoff & Lopus, 2014). Many studies have

confirmed that engaging students in active learning build the foundation of effective education. Evaluating user engagement in virtual environments can help improve the learning curve significantly (Beer et al., 2010; Fredricks, Blumenfeld, & Paris, 2004).

Past research studies have indicated that student engagement is a multidimensional construct with four primary dimensions namely academic, behavioural, emotional (psychological/affective), and cognitive engagement (Fredricks, Blumenfeld, & Paris, 2004). Figure 1 represents four engagement dimensions along with their respective context, indicators, and outcomes. These engagement dimensions can predict student involvement in educational activities and performance.

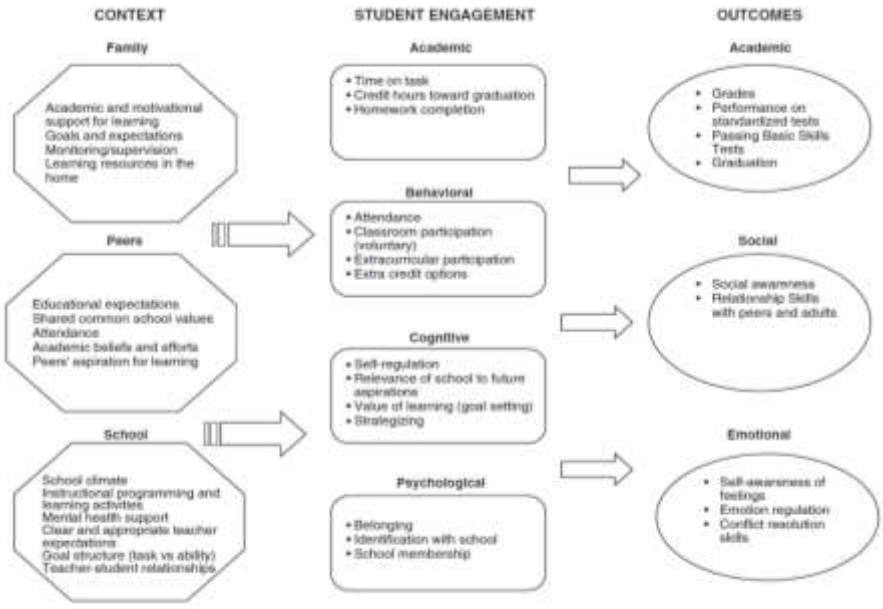

Figure 1. Engagement Subtypes, indicators and outcomes

(Appleton, Christenson, Kim, & Reschly, 2006)

Academic engagement can be measured in terms of the variables such as time spent to complete the task or grades received in the educational activity (Appleton et al., 2006). The level of behavioural engagement of a student can be represented by looking at their attendance, participation in class and extra-curricular activities (Appleton et al., 2006). The other two types (cognitive and affective) are known as internal forms of engagement. These types are as influential as academic and behavioural engagement but there is a lack of research evidence to confirm their link with academic performance (Appleton et al., 2006). Cognitive engagement involves indicators related to self-learning, future goals, and intrinsic motivation while affective engagement corresponds to the relationships with peers and teachers and feeling of identification with the

school that leads to more emotional awareness and improved conflict resolution skills, as depicted in Figure 1 (Appleton et al., 2006).

It is challenging to measure each dimension of user engagement in an accurate way because of the ambiguity in user engagement definitions (O'Brien, 2016). The ambiguity is related to multiple and sometimes conflicting definitions. According to Brandtzæg, Følstad, and Heim (2018), user's feeling of being in control of the interaction is referred to as engagement while Webster and Ahuja (2006) suggest engagement to be similar to flow with user perception of control, interest, attention, and curiosity. Various engagement measurement approaches have been investigated by researchers so far, including self-report surveys, observational checklists and rating scales and automated measurements (Bouvier, Lavoué, Sehaba, & George, 2013; Whitehill, Serpell, Lin, Foster, & Movellan, 2014). The subjective measurement approaches such as self-report and observational checklists can be used to measure cognitive and affective dimensions of engagement (Appleton et al., 2006). Appleton et al. (2006) devised a survey instrument named the Student Engagement Instrument (SEI) that includes 30 and 26 survey questions to measure the cognitive and affective engagement level of the student, respectively. The behavioural dimension of engagement can be measured using automated methods based on the data and evidence related to an objective category. The academic engagement dimension can be easily reflected by the academic results of students (Appleton et al., 2006).

2.1.2 Personality

Student personality is another significant factor that can play a role in shaping academic performance in various settings (Fosse, Buch, Säfvenbom, & Martinussen, 2015). This area of research has been attracting the attention of researchers in educational systems (Bauer & Liang, 2003; Chamorro-Premuzic & Furnham, 2003).

Past results have consistently shown different personality characteristics to be significantly correlated with academic performance (Fosse et al., 2015; Poropat, 2009). The Big Five model is a widely used instrument to measure personality dimensions. The personality traits in this model are labelled as Openness, Conscientiousness, Extraversion, Agreeableness, and Neuroticism also referred to as OCEAN (Goldberg, 1992; Gosling, Rentfrow, & Swann, 2003). The openness trait relates to the tendency to be more creative, curious, imaginative and polished. Conscientious individuals are hardworking, organised, efficient, dependable and achievement striving (Furnham & Monsen, 2009). Agreeableness represents characteristics of being more courteous, helpful,

sympathetic, thoughtful, good-natured and accommodating nature. Extravert people are more outgoing and friendly people with qualities of being sociable, adventurous, active and energetic. Neuroticism refers to the degree that defines emotional stability and calmness. People with a higher degree of neuroticism experience self-consciousness, negative emotions, and anxiety. People with a lower degree of neuroticism are composed and self-confident and have a tendency to show even temperament in most situations (Barrick & Mount, 1991; Busato, Prins, Elshout, & Hamaker, 2000).

According to a relevant past study, the conscientiousness personality dimension has been suggested to be a strong predictor of academic success for university students (Chamorro-Premuzic & Furnham, 2008). Findings from another similar study revealed that school students with conscientious personality trait demonstrated better performance in science subjects (Furnham & Monsen, 2009). Having a sense of freedom encourages highly conscientious people to be more interested in studies while individuals with a low level of emotional stability might not handle stressful conditions and that condition may affect their academic performance negatively (Furnham & Monsen, 2009). Similar results have emerged from another study that points out that people with a high level of openness and conscientiousness achieve success at the university level (Hazrati-Viari, Rad, & Torabi, 2012). In another relevant study, Eyong, David, and Umoh (2014) found a positive association of conscientiousness and agreeability traits with academic performance in one of the studies including secondary school students.

2.1.3 Navigation

A key feature of EVWs is that users are free to roam and explore allowing smooth and continuous interactions within a virtual world. User navigation and navigation paths are specific behavioural aspects in the context of learning within EVWs (Hanna, Richards, & Jacobson, 2014). Navigation can be an indicator of the level of engagement and predictor of performance (Hanna, Richards, Hitchens, & Jacobson, 2014). The navigation-related behaviour can shape the user's relationship to the virtual environment and has the capacity to impact the performance of the user within a virtual environment (Sas, O'Hare, & Reilly, 2004). In a past study, De Haan and Richards (2017) investigated the use of navigational hints for influencing the performance within EVWs. Behavioural engagement can be measured in form of data captured during a learning activity and navigation in the virtual world (De Haan & Richards, 2017; Hanna, Richards, Hitchens, et al., 2014; Hanna, Richards, & Jacobson, 2014).

Within an EVW users control their movement in the virtual world through navigating around the world. Users must devise a strategy to reach the desired goal by scanning and seeking the environment to extract the required information using environmental cues and artificial aids such as maps. Past researchers have investigated design principles for improving navigation efficiency in virtual environments (Darken & Sibert, 1993).

According to past research, three distinct levels of metrics have been suggested to evaluate wayfinding and influence decision-making by users in the virtual world. The first metric, at the bottom level, is the measurement of users' task performance in terms of time taken, distance travelled, and the number of errors made. The second metric is based on the physical behaviour of the user during navigation such as time spent in motion or staying idle, the path taken and classification of errors made during navigation and at the third topmost level, the metrics are the decision making or cognitive rationale (Ruddle & Lessels, 2006).

2.1.4 The Learner's State: Epistemic Emotions

In the journey towards facilitating deeper learning and a better understanding of new concepts, the learners go through various positive and negative states of the mind in the form of positive and negative emotions. Previous studies indicate that it is crucial to managing negative emotions effectively to maintain the learning process at that time otherwise students might get disengaged and give up on the learning process. Virtual agents with affective capabilities can assist learning by assessing the learner's emotional state and can provide appropriate support in the form of feedback and motivation by building social connections with learners (Sabourin, Mott, & Lester, 2011). Awareness of the learner's negative affective state builds a foundation for the process of successful achievement of learning goals in online digital environments (Shen et al., 2009).

According to a theory proposed by Kort, Reilly, and Picard (2001), emotions can be divided into two types namely basic and non-basic emotions. Anger, sadness, surprise, happiness, fear, and disgust belong to the basic emotions category while confusion, boredom, engagement, curiosity and frustration are examples of non-basic emotions. Non-basic emotions are the emotions most frequently experienced by learners while learning a new concept. Emotions that are generated during the learning process are referred to as epistemic emotions. These emotions have the characteristic of being dynamic and are associated with cognitive activities (Pekrun & Linnenbrink-Garcia, 2012).

Out of all these states during learning, engagement and confusion are most likely to occur during a learning process and can help influence the learning in a positive manner while negative states like anxiety, frustration, boredom can be detrimental to learning and can hinder the learning process (Wu, Shen, & Miao, 2013). These negative learning states of mind should be detected in the early stages of learning to avoid frustration. Understanding learners' emotions allow the system to intervene to provide timely and appropriate feedback to the students to keep them motivated and provide them with the coping strategies needed to deal with these emotions (Mcquiggan, Lee, & Lester, 2007).

According to another study, certain patterns are observed in terms of the emotional transition from one state to another state, during the learning process and that has proved to be an interesting input for research in this area (D'Mello, Lehman, Pekrun, & Graesser, 2014).

While performing learning activities, as the complexity of the task increases, a learner can get into a state called stuck and may go through a phase of non-optimal experience. The feeling of 'stuck' can be defined as a feeling of being out of control along with a lack of focus, mental fatigue, and distress. This is the stage when the possibility of learners to lose motivation for learning increases (Burleson & Picard, 2004). Figure two illustrates that a learner is more engaged initially while working to achieve the superordinate goal of completing a task. Subsequently, the boredom emotion appears and causes disengagement from the task. Confusion is the next phase in this cycle that is generated because of misconceptions and unexpected feedback. Confusion converts to frustration when the learner feels stuck and cannot make further progress (D'Mello & Calvo, 2013). This transition pattern allows prediction of the next state and might help to guide how to respond to the learner's current state to move or keep them in the desired emotional state.

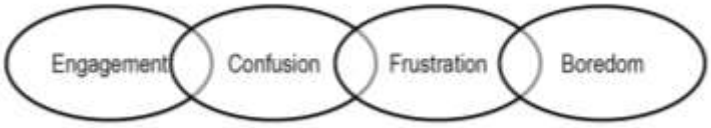

Figure 2. Emotional transition pattern during learning

(D'Mello et al., 2014)

Kapur and Bielaczyc (2012) proposed a theory of productive failure. According to this theory, students' long term learning can be maximised if they work with a goal but without any instructional support or scaffolding from the teacher in the initial phase. This effort by the student helps to challenge and engage them. A consolidation phase led by the teacher follows that allows

the student to consider any hypotheses made earlier and potentially refine their understanding and revise misconceptions. A related concept is the notion of impasse-driven learning. A study conducted by Kapur and Bielaczyc (2012) demonstrated the concept of impasse driven learning through a set of students who worked on complex mathematical problems at their own pace without teacher's instructional support but were gradually provided increased scaffolding to arrive at a final solution. The results indicated that, in their first attempt, students were not able to achieve the desired success without teacher support but they performed well in their post-test reflecting the efficacy of theory of productive failure and impasse-driven learning. The past empirical studies aligned with this theory have emphasised that there is an association between achievement of successful learning outcomes and impasse driven learning where students are allowed to understand the underlying concept without any instructional guidance initially and help is provided later only when they have adequately tried solving the problem in various ways and cannot make their way forward (VanLehn, Siler, Murray, Yamauchi, & Baggett, 2003).

Confusion is one of the significant epistemic emotions that is associated with complex learning and can help in improving learners' engagement (D'Mello et al., 2014). The research idea suggested by VanLehn et al. (2003) highlights the positive influence of impasse-driven learning suggesting that new concepts are grasped well by learners when they reach an impasse and engage themselves in cognitive activities to come out of the state of confusion. According to a prior research study conducted by VanLehn et al. (2003), confusion allows deeper learning if it is managed within the zone of optimal confusion as depicted in Figure 3. This implies that induced confusion should be resolved within certain thresholds in terms of discrepancy and duration otherwise it might lead to frustration, disengagement, and boredom, and can influence learning in a negative direction (Arguel & Lane, 2015; D'Mello et al., 2014).

Figure 3. Zone of Optimal Confusion

(Arguel & Lane, 2015)

VanLehn et al. (2003) also emphasised that students should be provided with goals to help them reach an impasse and make learning successful. There are intelligent tutoring systems (ITS) that

have the capability to interpret and sense user actions and offer flexibility according to student needs (Lima & Rosatelli, 2003). Some ITS follow the principle of impasse learning and allow students to try and take some steps before providing any help or hints but these systems fail to draw out satisfactory explanations from students (VanLehn et al., 2003).

According to a research study done by Arguel, Lockyer, Lipp, Lodge, and Kennedy (2017), during the learning process, new information is mapped into prior knowledge. In these scenarios, there is the possibility of inconsistencies to be generated either within the diverse information gained via EVW or between new information and the prior knowledge. This situation brings the learner into a state of cognitive disequilibrium or impasse that can further lead to confusion. Arguel and Lane (2015) discuss two intervention strategies where the first approach is to induce confusion by provoking task-relevant cognitive disequilibrium, for example, by providing inconsistent or contradictory information and the second approach is to manage the confusion using appropriate strategies.

To avoid the negative effects of confusion during complex learning and to help students come out of impasse or cognitive disequilibrium, strategies such as self-reporting surveys and behavioural and physiological measures have been proposed. These strategies can't be one solution to fit all but it is feasible to combine some of these strategies to provide a multimodal solution depending on the situation and individual factors such as age, personality and prior knowledge (Arguel et al., 2017).

2.2 Providing Support to Learners in an EVW

There are numerous ways to provide support to students in EVWs so that they can achieve desired learning outcomes. The techniques such as self-explanations (Chi, Leeuw, Chiu, & LaVancher, 1994), verbal or non-verbal feedback (Hattie & Timperley, 2007; Lin, Atkinson, Christopherson, Joseph, & Harrison, 2013; Ranjbartabar & Richards, 2017), providing motivation (Burleson & Picard, 2004; van der Meij, 2013), improving self-efficacy skills (Kim, 2005), use of empathy (Kim, Baylor, & Shen, 2007), use of game-based environment (Sabourin & Lester, 2014) and scaffolding (Villarica & Richards, 2014a; Wu & Looi, 2010) can be implemented to provide appropriate help to the learners to maintain learning.

The crucial aspect of providing appropriate help in form of encouragement and support, the learners' epistemic emotions should be recognised correctly and support should be provided according to the understanding of the individual learner (Kort et al., 2001). Eliciting learner's

emotions in Educational Virtual Worlds is a challenging aspect (Leony, Muñoz-Merino, Pardo, & Kloos, 2013). Some of the affect detection techniques that have been implemented so far are observation of learner's emotions by human observers (Woolf et al., 2009), self-reporting (Sabourin, Mott, & Lester, 2011), measuring physiological data (Prendinger & Ishizuka, 2005) including facial expressions (Ammar, Neji, Alimi, & Gouardères, 2010; Duo & Song, 2012; Kort et al., 2001), gaze pattern (Levitski, Radun, & Jokinen, 2012). Other approaches included use of machine learning techniques such as Decision-Tree, Support Vector and Bayesian networks (Arguel et al., 2017; Woolf et al., 2009), analysis of learning data generated during usage (Tian et al., 2014). Another study demonstrated the use of the combination of physiological signals, such as conversational cues, skin conductance, heartbeat and body temperature for eliciting emotions during learning (D'Mello & Graesser, 2010).

All of these methods to elicit the learners' affective state have their own pros and cons. Self-reporting is easy to gather but the limitation of this method is that sometimes the user is not ready to reveal their emotions. In addition, self-reporting of emotions is conducted in regular intervals of time so it does not reflect the moment by moment emotional state of the user. One significant problem concerning the techniques using physiological signals to assess the learners' affective states is that proper lab equipment needs to be set up and that makes it hard to use these techniques in all scenarios (Arguel et al., 2017).

It becomes particularly hard to detect the learner's state and respond appropriately within EVWs because of the dynamic nature of these emotions and because the learner is often remote. During face to face teaching, a human teacher can take necessary steps to resolve situations that can hinder the learning process, however in Educational Virtual Worlds, it is very likely that students can experience various states of academic emotions and there is no teacher to observe it . The result of this might be that the student can get disengaged and lose motivation to continue learning (Arguel et al., 2017). In the case of Educational Virtual World, it becomes a complex process to recognise when a learner needs help to move forward. It also misses on providing the social element that is a critical factor of successful learning (Duo & Song, 2012).

In the context of EVWs, analysis of the interaction data generated during the interaction of the user and virtual agent can help to find some patterns and to predict the learner's level of engagement with the learning task (Hanna, Richards, & Jacobson, 2014; Sas et al., 2004). It allows designing specific algorithms to provide necessary intervention (Arguel et al., 2017). While some work has been done in this direction, still there is more scope of research in this area. To

implement research strategies in this direction, the relevant data can be collected in form of time taken to complete the given task, path navigated in case of game-based virtual world, idle time and questions asked during the interaction (Hanna, Richards, Hitchens, et al., 2014; Hanna, Richards, & Jacobson, 2014).

2.3 Research Gap and Focus

There is an interplay between personality factors, engagement and academic performance however, little attention has been focused on measuring the relationship between these factors. Prior research has established the role of engagement (Carini, Kuh, & Klein, 2006; Hassaskhah, Khanzadeh, & Mohamad Zade, 2013; Hoff & Lopus, 2014) and personality factors as determinants of academic success (Lounsbury, Sundstrom, Loveland, & Gibson, 2003; Paunonen & Ashton, 2001) however, little attention has been focused on assessing the combined influence of these factors.

There are many existing EVWs such as Omosa 3D Virtual system (Richards et al., 2012), Crystal Island (Robison, Mcquiggan, & Lester, 2009) and relevant agent architectures such as FAtiMA-Fearnot AffecTIve Mind Architecture (Paiva et al., 2005), CARE-Companion-Assisted Reactive Empathizer (McQuiggan & Lester, 2007) based on emotion-based frameworks to provide help to learners. These architectures still lack the capacity to deal with significant learning emotions such as engagement, confusion, frustration, and boredom. Moreover, the current focus is on the pedagogical agent portraying emotion, not on detection of the user/learner's emotion. In addition to lack of detection, to provide appropriate help within an educational virtual world, there is need to understand what relationships may exist between demographics, personality traits, engagement level along with the emotional state that the learner is going through. Depending on the complexity level of the learning task, students can feel a lack of motivation, confusion or get stuck. At that time, the virtual agents or something in the EVW environment should encourage learners and support their learning in many ways such as by responding affectively or guiding them by providing appropriate hints or demonstrating the complex task for them (Villarica & Richards, 2014b). In a recent study, the possibility of using multi-data and multimodal approaches to elicit the emotional state of the learner within EVW was explored in order to suggest the most appropriate intervention (Makhija, Richards, Caballé, & Conesa, 2018).

To conclude, there is a need to design a more adaptive and responsive system that can provide timely help to learners in science inquiry learning environments. This research is directed towards bridging the gap by exploring whether providing tailored support based on the interaction of the

student with an EVW provides useful support to improve performance and whether personal aspects of the learner are associated with their performance or any changes in performance according to the type of help they have received.

CHAPTER 3: METHODOLOGY

The main purpose of this research study was to determine what help, in the form of hints, might be useful to aid students using an EVW and whether different individual factors influence the usefulness of those hints. This chapter presents the methodology used to answer the following research questions.

Research Questions:

1. Does the provision of generic hints influence the performance in a quiz after using an EVW?
2. Does the provision of tailored hints influence the performance in a quiz after using an EVW?
3. Is there a relationship between an individual's big five personality traits and their performance in a quiz after using an EVW?
4. Is there a relationship between and individual's level of affective engagement or cognitive engagement and their performance in a quiz after using an EVW?
5. Is there a relationship between individual factors, quiz performance and type of hints provided (generic, tailored, none)?

To be able to answer these research questions, it was necessary to capture and analyse data related to students' personality and engagement self-reported by participants along with the system interaction data from using the EVW.

3.1 Research Design

For this purpose, an online study was conducted involving an educational virtual world used by students. This study was approved by Macquarie University's Human Research Ethics Committee (See Appendix E). The students in this study were invited to participate via Psychology Sona pool and received half an hour course credit for their participation. The Psychology Sona pool is a platform provided by Macquarie University where researchers can get access to the participants for their research studies. This study builds upon a previous study that focussed on navigation paths and performance in educational virtual worlds (De Haan & Richards, 2017).

An experiment was designed involving three groups. Table 1 illustrates the formation of these groups. Each group used the EVW in two separate navigations. In the first attempt, all groups did scenario 1 without getting any hints. During the second attempt, all participants did scenario 2 but according to assigned group received no hints, general hints or hints specific to their behaviours in scenario 1. Group 1 completed scenario 2 without getting any hints, referred to as Control group (SC). Group 2 completed scenario 2 by receiving all the general hints and is referred to as AllHints (SA) group. Group 3 completed scenario 2 after receiving tailored hints based on analysis of their logfile data from scenario 1 and is referred to as TailoredHints (ST) group.

Section	Group 1	Group 2	Group 3
Part A	Demographic Questionnaire		
Part B	Scenario 1		
Part C	Scenario 1 Questionnaire		
Part D	Scenario 2 (Without Hints)	Scenario 2 (With All Hints)	Scenario 2 (With Tailored Hints)
Part E	Scenario 2 Questionnaire		
Part F	TIPI (Ten Item Personality Inventory) Survey		
Part G	SEI (Student Engagement Instrument) Survey		

Table 1. Experimental Groups and Study Procedure

3.2 Experimental Procedure & Data Collection

The research study procedure took around 30 minutes for each participant. It initiated with a demographic questionnaire with participants that included questions to extract participants' information related to gender, cultural group, age, and the number of hours they spent playing computer games every week. Followed by a demographic questionnaire, the participants navigated through the virtual world where they had two interactions with the EVW depending on the group participants belonged to, as illustrated in the experimental design section.

Following each interaction with EVW, participants performance was assessed through fifteen relevant questions based on the experience they had by observing and conversing with the virtual characters and information that was collected by them during navigation of the virtual world. The list of questions for scenario 1 and scenario 2 asked from these participants is provided in Appendix C and Appendix D respectively. These questionnaires were prepared through Qualtrics, a single platform to create an online survey along with the embedded software. This feature made it

convenient for participants to complete the survey without leaving the Qualtrics and finish it in their own time and place. After each interaction with the EVW, the participants were also asked to indicate their enjoyment rating on the scale of 1 to 5 with 1 being not enjoyable, and 5 most enjoyable, along with the reason for their answer. If participants belonged to hints groups (SA or ST), they were also asked about the usefulness of the hints if they received any.

Participants also completed the Ten Item Personality Instrument (TIPI) that is comprised of 10 questions (Gosling et al., 2003) (see Appendix B). TIPI provides a quick measurement of the Big-Five personality dimensions (extraversion, agreeableness, conscientiousness, emotional stability, and openness to experiences) (Gosling et al., 2003). This instrument has been included to analyse their performance within EVW based on their personality factor. Participants responded to TIPI statements by rating their level on Likert scale of 1-7 with 1 indicating that they never felt or acted that way, and 7 indicating that they certainly acted that way. The response choices for Ten Item Personality Measure (TIPI) questionnaire was as follows: 1-Disagree Strongly; 2-Disagree Moderately; 3-Disagree Little; 4-Neither Agree Nor Disagree; 5-Agree a Little; 6-Agree Moderately; 7-Agree Strongly.

To measure cognitive and affective engagement, the Student Engagement Instrument (SEI) (Appleton et al., 2006) was used (see Appendix A). Originally SEI is designed for use on students in middle school, so for this study purpose, it was modified for university students. For example, the word “school” is replaced with “university”. SEI includes 33 items, nineteen items to measure affective engagement and fourteen to measure cognitive engagement. Participants responded to statements by rating their level on Likert scale of 1-4 with 1 indicating that they never felt or acted that way, and 4 indicating that they certainly acted that way. Response choices for SEI questionnaire were as follows: 1-Agree Strongly; 2-Agree; 3-Disagree; 4-Agree Strongly.

The student’s behavioural engagement was captured via their interactions within the world consisting of how many virtual characters they spoke to, where they had visited, how long they were in the world, the number of navigation steps and navigational patterns. The academic engagement was captured via the number of correct answers: min-0 max-15 to the 15 questions in the questionnaires given after using the EVW. The virtual world was deliberately laid out in such a way that participants may not always know exactly what to do, or where to go next. Omosa is a fictitious world and thus the answers to the questions could not be based on prior knowledge. We used the number of correct answers to categorise participants into three groups to represent their performance in an educational virtual world based: Low-1 to 5 correct; Medium-6 to 10 correct;

High-11 or more correct. In this research, quiz scores were used to evaluate the academic dimension of user engagement. Cognitive and affective engagement was captured using the Student Engagement Inventory (SEI).

To facilitate analysis, the data were transformed and prepared. Preparation included checking response values to be valid and complete. First, data regarding participants' demographics were tabulated. Personality, Quiz and SEI scores were computed and added as new variables. Categories were created for many factors to allow crosstabulations and chi-square tests from continuous data. A score from 0-15 was calculated for each quiz. The quizzes included multiple choice questions, with one correct answer that received a score of 1. To answer the research questions and to identify potential significant relationships between the categorical variables such as gender, cultural groups, personality, and different types of engagement levels and quiz scores, we calculated paired t-test on the results. Descriptive statistics, crosstabulations, and chi-square tests were also performed using the reporting and data analysis tools in Qualtrics.

3.3 Materials

The EVW was first extended by another study related to navigational hints as discussed in section 3.1 (De Haan & Richards, 2017). For this study, we further modified the EVW to answer suggested research questions. Omosa VW is an EVW designed using unity 3D at Macquarie University for teaching biology and science inquiry concepts. It represents a fictitious island with five main locations. The five locations in Omosa island are the village, the research lab, the hunting ground, the animal counting station, and the weather station. Participants can visit all of these locations by clicking on the map provided in the virtual world and can get the required help in written text form from virtual agents on main locations. Students can ask for help regarding pre-defined scientific questions related to biological systems of the island.

Participants were allowed to navigate freely on the island. During their journey on the island, it was possible for them to converse with all characters, collect inventory items and gather relevant information through observation. All the navigational information related to participants was stored in the database on the server. The data from navigation patterns and log files helped to determine when the student needed help to continue with their learning and that help was provided through hints.

The interaction data between the participants and Omosa EVW were stored in the database in regular intervals in terms of which places they visited, which virtual character did they talk to,

what questions did they ask, what inventory items they picked up. These logfile data points allowed calculation of the type of hint the participants in the tailored hints group (ST) would receive. To support the experiment and answer the research question, the EVW and dialogues with the virtual characters were modified to capture the relevant data, provide two scenarios that guided the participant to cover the relevant context in the EVW and determine what hints were needed to the participants of tailored hints group (ST). Before starting the interaction with EVW, the participants were given the playing instructions in the form of dialog through the first character they meet in the environment.

Figure 4. Omosa Map

3.3.1 Scenarios

To ensure the efficacy of provided hints, this experiment needed to be designed to eliminate any learning effects from using the EVW twice. As participants interacted with the EVW twice, we can expect that the second time they interact, they will know more about the world because of the prior knowledge they gained in the first interaction. To eliminate this confounding factor, two scenarios (1 and 2) were designed in the virtual world with slightly different information and

pathways together, each with an associated questionnaire containing 15 questions. The questionnaires and their answers can be found in the corresponding scenario (See Appendices C and D for scenario Questionnaires).

Both scenarios were similar in nature in a way that the amount of information and distribution of information over characters was evenly distributed. Nature and difficulty level of scenario and questions were adjusted to be similar. The questions could be answered correctly based on the conversation with the characters as well as based on the observations during navigation. For example, the participants could collect some inventory items on their way in the virtual world and that helped them to figure out the correct answer for specific questions (see Appendix C-Question 1.12).

The conversations for each scenario were designed to ensure they were different in both scenarios so that participants could not learn the answer to a question in the other scenario. The scenarios varied in a way that they had different starting points and different optimal paths.

3.3.2 Hints

In this study, six hints have been used for all hints (SA) and tailored hint (ST) groups. Out of six hints, fourth, fifth and sixth hint are based on navigational behaviour of the participants and were developed for a study based on navigation paths undertaken in past (De Haan, Richards, & Dignum, 2017). First three hints were added to address specific research questions in current research. Hints have been provided in two ways to two different groups. The all hints group (SA) is provided with generic hints and tailored hints group (ST) is provided with tailored hints (see Table 2). These hints are based on the log file data indicating, for example, that participants didn't talk to characters enough or didn't pick any inventory item because of inattentiveness or didn't follow the optimal path on basis of navigation pattern.

There are six different hints that were provided to the participants in both hints groups for the study (see Table 2). The all hints (SA) group were shown all the generic hints irrespective of their behaviour and navigation pattern in the virtual world while the tailored hints (ST) group participants were given only specific hints depending on their behaviour and pattern followed during EVW navigation. The All hints group always received all six hints while the number of hints to the tailored hint group participants varied according to their interaction with EVW. The control (SC) group was not given any hints at all.

The first hint was based on the number of virtual characters visited by the participants. There are six characters (Lyina, Omewey, Charlie, Pedro, Kim, and Zafirah) located in different locations in

Omosa that can interact with the participant and provide information depending on the design of the scenario. During the first scenario, they started the island visit meeting with Lyina character and see Omewey, Charlie, and Pedro further on their way as suggested by the previous character. For the second scenario, the sequence was to go to Charlie first and then visit Pedro, Zafirah, and Omewey subsequently. If the participant missed any of the characters in the first interaction that they were supposed to see, character hint (hint 1) was shown (see Table 2) to them during second interaction.

The second hint was based on the number of inventory items collected by the participants. The collection of inventory item also helped them answer some of the questions in the questionnaire and indicated their observation skills. If the participant did not collect a single inventory item in the first interaction, inventory hint (hint 2) was shown (see Table 2) to them during second interaction.

The third hint was based on the areas covered by the participant. If they did not cover all the required areas in the first interaction, they were provided with the hint 3 (see Table 2) during second interaction

The fourth hint was based on the number of teleports that were done using the map. Teleporting using the map meant that the participant could simply click a location on the Omosa game/island map and it would transport them to that location place in the Omosa EVW. There are five main island locations including the starting position so all locations can be visited using four teleports. If the participant teleported more than eight times during first interaction, they were provided with teleport hint during their second interaction.

The fifth hint is based on the amount of backtracking a participant does. The environment is divided into zones where each location in the environment has its own zone, which is based on the metrics defined by Ruddle and Lessels (Ruddle & Lessels, 2006). A path is marked as backtracking when a zone is left and re-entered at any point during navigation. If a participant backtracked during their first play-through the backtracking, hint 4 was given at the start of the second interaction, as shown in Table 2.

The sixth hint, based on the optimal path, that was originally designed for an earlier study related to evaluating navigation paths and improving performance in EVW (De Haan & Richards, 2017). The characters in the virtual world give instructions on where to find another character. If a player follows that path exactly means the optimal path has been followed. The following of the optimal

path is calculated by taking the shortest distance to the path for each data point and if the distance is over a threshold then data point gets counted as outside the optimal path. If more than five percent of the data points are outside, the path is deemed as not the optimal path. If the participant did not follow the optimal path during the first interaction, the optimal path hint (see Table 2) was given at the start of the second interaction.

Trigger	**Tailored Hints**	**Generic Hints**
Character	*You missed some characters you needed to speak to last time, make sure you pay attention to which characters are suggested.*	*Make sure you pay attention to which characters you have to talk to.*
Inventory Item	*Your backpack is empty. Remember to collect the inventory items on the way. That will help you gain more knowledge about the Island.*	*Collect all inventory items on the way. That will help you gain more knowledge about the Island.*
Areas	*Last time you missed some important areas, while navigating the island. Try to work out which areas are important.*	*Try to cover all important areas, while navigating the island.*
Teleporting	*You teleported a lot in your last playthrough, try and figure out where you are going instead of teleporting.*	*Don't teleport a lot. Try and figure out where you are going instead of teleporting.*
Backtracking	*You went back to a place you had already been, please try to discover new parts of the island (backtracking hint)*	*Avoid going back to a place you have already been, please try to discover new parts of the island.*
Optimal Path	*You seemed to have missed some instructions last time. Pay attention to what the people have to say, and try to follow their instructions*	*Pay attention to all the instructions. Observe carefully and try to follow instructions of people on the way.*

Table 2. Detail of hints provided to the participants

3.4 Chapter Summary

This chapter discusses the research questions, hypothesis, and methodology used to answer these research questions. It also includes the research study procedure along with details of hints and specific scenarios designed for this study. The materials section discusses the EVW named Omosa that has been used in this research study.

CHAPTER 4: RESULTS

A sample of 232 undergraduate Psychology students participated in this study. At a critical timepoint during this study, an unannounced change was made by technical services for our faculty, who were hosting this study, that affected the data collection. The affected data was collected during the last week of the semester. This is a week that sees high study activity because students must complete their studies by Friday 5 pm to receive course credit. Unbeknown to us, at that time certain changes were made to the security certificate on the server, that intermittently prevented some of the students from accessing the software. The study had been running for a couple of weeks and carefully tested, so it was only through a small number of participants contacting the researchers that some issue became apparent. Qualtrics survey data were still being captured, but EVW data that were being captured on local servers was intermittently being sent to the desired database because of a server issue. It took a few days to identify and rectify the error as Qualtrics notifications were indicating successful completion of the study by the participants. Over 80 records were partially affected. The loss of some logfile data in the database that captured the logfile data meant that it was unclear what the impact might be on the tailored hints. Thus we had to discard any incomplete records. At the end, 144 records were found to be complete and useful that has been used for the analysis purpose.

While the Qualtrics randomiser equally distributed participants into each group the above mentioned intermittent technical issues caused some of the participants to stop the participation and resulted in an imbalance in the number of participants in each group. After discarding incomplete and incorrect records, a total of 37 were assigned to the 'Control' group and 57 and 50 were assigned to 'All Hints' and 'Tailored Hints' groups respectively.

4.1 Demographics

In terms of demographics, out of 144 respondents, 38 were male and 106 were female. See Table 3 for gender breakup for each experimental group. Ages ranged from 17-43 with the mean age of

these participants was 19.80 and a standard deviation of 3.81. 53 (36.81%) of participants regularly played computer games.

Group	Female	Male	Total
Control (SC)	35 70%	15 30%	50 34.72%
All Hints (SA)	46 80.70%	11 19.30%	57 39.58%
Tailored Hints (ST)	25 67.57%	12 32.43%	37 25.69%
Total	106 73.61%	38 26.39%	144 100%

Table 3. Gender distribution across treatment groups

The results regarding cultural background are shown in Figure 5. The largest cultural group represented, Oceania (which includes Australia) (26%), followed by South-East Asian (15%)

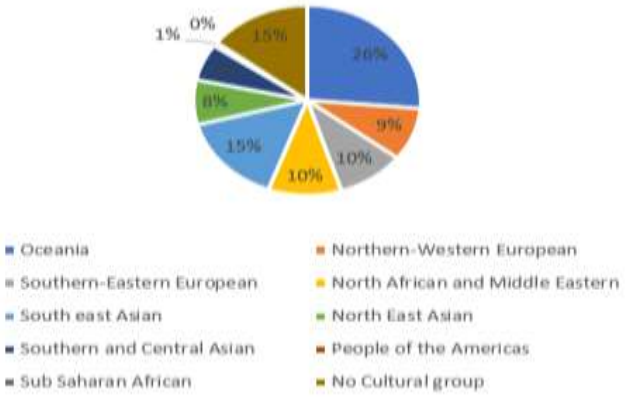

Figure 5. Number of participants in different cultural groups

4.1.1 Personality Data

For each individual, means have been calculated for each of the five personality dimensions (OCEAN) as self-reported by the participants. This data was categorised into low (L), medium (M) and high (H) for better identification of patterns and profiles. Table 4 shows the personality data categorised by gender.

Gender	Openness				Conscientious				Extraversion				Agreeableness				Neuroticism			
	L	M	H	Total	L	M	H	Total	L	M	H	Total	L	M	H	Total	L	M	H	Total
Female	5 4.72%	26 24.53%	75 70.75%	106 100%	12 11.32%	36 33.96%	58 54.72%	106 100%	32 30.19%	34 32.08%	40 37.74%	106 100%	5 4.72%	43 40.57%	58 54.72%	106 100%	25 23.58%	44 41.51%	37 34.91%	106 100%
Male	6 15.79%	14 36.84%	18 47.37%	38 100%	6 5.79%	12 31.58%	20 52.63%	38 100%	15 39.47%	7 18.42%	16 42.11%	38 100%	3 7.89%	20 52.63%	15 39.47%	38 100%	3 7.89%	12 31.58%	23 60.53%	38 100%
Total	11 7.64%	40 27.78%	93 64.58%	144 100%	18 12.50%	48 33.33%	78 54.17%	144 100%	47 32.64%	41 28.47%	56 38.89%	144 100%	8 5.56%	63 43.75%	73 50.69%	144 100%	28 19.44%	56 38.89%	60 41.67%	144 100%

Table 4. Categorised Personality Data based on Gender (L=low, M=medium, H=high)

4.1.2 Engagement Data

Affective engagement (AE) is comprised of: Teacher-Student Relationships (TSR) (mean=2.04, s.d=0.68), Peer Support at School (PSS) (mean=1.88, s.d=0.71), Family Support for Learning (FSL) (mean=2.14, s.d=0.69) while Cognitive engagement (CE) is comprised of Control and Relevance of School Work (CRSW) (mean=1.34, s.d=0.55), Future Aspirations and Goals (FG) (mean=2.24, s.d=0.71), Intrinsic Motivation (IM) (mean=1.92, s.d=0.74).

For analysis and comparison, the mean results have been categorised into low, high and medium. Based on the SEI questionnaire answers provided by all participants, their affective and cognitive engagement score was calculated. Participants were divided in to three categories of low (<2.75), Medium (2.75-3.15) and High (>=3.15) of engagement levels for both types of engagement levels. [1]

After averaging the results for the three components in each type of engagement, Table 5 presents the number of participants in each category of Affective and Cognitive engagement. Table 6 breaks down Affective and Cognitive Engagement into its components.

Level	Affective Engagement	Cognitive Engagement
Low	31 21.53%	44 30.56%
Mid	87 60.42%	80 55.56%
High	26 18.06%	20 13.89%
Total	144 100%	144 100%

Table 5. Number of participants in affective and cognitive engagement categories

Category	TSR	PSS	FSL	CRSW	FG	IM
Low	30 20.83%	46 31.94%	26 18.06%	101 70.14%	23 15.97%	45 31.25%
Mid	78 54.17%	70 48.61%	72 50.00%	37 25.69%	63 43.75%	65 45.14%
High	36 25.00%	28 19.44%	46 31.94%	6 4.17%	58 40.28%	34 23.61%
Total	144 100%	144 100%	144 100%	144 100%	144 100%	144 100%

Table 6. Number of participants in engagement subcategories. TSR=Teacher-Student Relationships, PSS=Peer Support at School, FSL=Family Support for Learning, CRSW= Control and Relevance of School Work, FG=Future Aspirations and Goals, IM= Intrinsic Motivation

[1] We were unable to find any information regarding standard cut-offs values in literature, so we decided to take values so as to determine a more balanced number of participants per groups.

4.1.3 Comparisons of demographics

To identify relationships between different demographic factors, numerous cross-tabulations were run. Table 7 shows the link between affective and cognitive engagement for all participants.

		Affective Engagement			
		low	medium	high	Total
Cognitive Engagement	low	20 9.47 45.45% 64.52%	23 26.58 52.27% 26.44%	1 7.94 2.27% 3.85%	44 100.00% 30.56%
	medium	11 17.22 13.75% 35.48%	60 48.33 75.00% 68.97%	9 14.44 11.25% 34.62%	80 100.00% 55.56%
	high	0 4.31 0.00% 0.00%	4 12.08 20.00% 4.60%	16 3.61 80.00% 61.54%	20 100.00% 13.89%
	Total	31 21.53% 100.00%	87 60.42% 100.00%	26 18.06% 100.00%	144 100% 100%

Table 7. Link between Affective and cognitive engagement categories

Chi-square tests revealed a significant relationship between game playing behaviour and agreeability personality dimension X2 (N=144) = 7.81, p= 0.02. People who do not play games were more agreeable. There were 59.34% out of 91 participants who didn't play games reported to be highly agreeable (see Table 8).

Agreeability Personality Dimension				
	Low	Medium	High	Total
Play Games	3 2.94 5.66% 37.50%	31 23.19 58.49% 49.21%	19 26.87 35.85% 26.03%	53 100% 36.81%
Doesn't Play games	5 5.06 5.49% 62.50%	32 39.81 35.16% 50.79%	54 46.13 59.34% 73.97%	91 100% 63.19%
Total	8 5.56% 100%	63 43.75% 100%	73 50.69% 100%	144 100% 100%

Table 8. Association between agreeability and game playing behaviour

While linking engagement levels with OCEAN categories, chi-square tests revealed significant associations between specific OCEAN categories and engagement levels: affective engagement and extraversion_category X2 (N=144) = 11.97, p= 0.02); affective engagement and

agreeableness_category X2 (N=144) = 9.43, p= 0.05; cognitive engagement and openness_category X2 (N=144) = 12.18, p= 0.02); cognitive engagement and extraversion_category X2 (N=144) = 14.71, p= 0.01). Out of 26 high AE participants, 57.69% belong to high extravert category and out of 20 high CE participants, 50% belong to high extraversion_category. Out of 20 high CE participants, 11 (55%) have reported to be with high openness category. The results also indicate link between conscientiousness and affective engagement X2 (N=144) = 8.79, p= 0.07) and cognitive engagement respectively X2 (N=144) = 8.75, p= 0.07).

Chi-square tests were calculated to find link between individual personality dimensions and engagement levels. These results also revealed significant associations between specific personality dimensions and engagement levels aligned with previous results regarding association between OCEAN categories and engagement levels: affective engagement and extraverted X2 (N=144) = 23.70, p= 0.02); affective engagement and dependable X2 (N=144) = 21.43, p= 0.04); affective engagement and anxious X2 (N=144) = 28.65, p= 0.00); affective engagement and calm X2 (N=144) = 25.45, p= 0.01); cognitive engagement and extraverted X2 (N=144) = 30.21, p= 0.00); cognitive engagement and calm X2 (N=144) = 25.39, p= 0.01).

Cross-tabulations showed that participants with high affective engagement most commonly (57.69%) chose the Likert scale option “agree moderately” and “agree strongly” to being extraverted while participants with low affective engagement most commonly (54.84%) chose “disagree a little” or “agree a little” to be extraverted. Results also show that 65.38% of high AE participants chose “agree moderately” and “agree strongly” to be calm and emotionally stable with high level. Chi-square tests comparing gender or the cultural group with any of the engagement or personality variables for any scenario did not reveal any significant differences.

4.2 Performance Scores and Hints

The quizzes in both scenarios had fifteen questions each. For each treatment group, the average number of correct answers were calculated along with standard deviations (see Table 9).

Group	Participants	Correct Answers				
		Quiz 1	Std Dev.	Quiz 2	Std Dev.	Avg
Control	37	7.3	3.2	7.3	2.5	7.3
All Hints	50	6.8	3.0	6.6	2.6	6.7
Tailored Hints	57	7.1	2.5	6.4	2.6	6.8

Table 9. Correct Answers for each treatment group

Participants who received all hints had a mean score of 6.8 in their first scenario and 6.6 in their second scenario. Participants who received tailored hints had a mean score of 7.1 in their first scenario and 6.4 in their second scenario. The participants in the control group, who did not receive any hints had the highest score of 7.3 in their first attempt and score of 7.3 in their second attempts as well. Results indicated no change or decrease in score for all groups from the first attempt to second attempt instead of showing improvements. Chi-square tests comparing gender, age, the cultural group with quiz scores for any scenario did not reveal any significant differences. Similarly, no relationship was observed between game playing characteristic and quiz scores.

To summarise the data for full analysis, all the participants were categorised into three groups namely low (1-5), medium (6-10), high (11-15) according to the number of correct quiz answers. The low, medium and high score split has been considered to have a balanced number of participants in each group. The categorised results for all treatment groups are shown in Table 10 for both quizzes.

Group	Quiz 1 Score				Quiz 2 Score			
	Low	Med	High	**Total**	Low	Med	High	**Total**
All Hints Group (SA)	13 26.00% 40.63%	13 26.00% 37.14%	24 48.00% 31.17%	50 100.00% 34.72%	13 26.00% 40.63%	12 24.00% 30.77%	25 50.00% 34.25%	50 100.00% 34.72%
Tailored Hints Group (ST)	10 17.54% 31.25%	16 28.07% 45.71%	31 54.39% 40.26%	57 100.00% 39.58%	13 22.81% 40.63%	19 33.33% 48.72%	25 43.86% 34.25%	57 100.00% 39.58%
Control Group (SC)	9 24.32% 28.13%	6 16.22% 17.14%	22 59.46% 28.57%	37 100.00% 25.69%	6 16.22% 18.75%	8 21.62% 20.51%	23 62.16% 31.51%	37 100.00% 25.69%
Total	32 22.22% 100.00%	35 24.31% 100.00%	77 53.47% 100.00%	144 100.00% 100.00%	32 22.22% 100.00%	39 27.08% 100.00%	73 50.69% 100.00%	144 100.00% 100.00%

Table 10. Number of participants in each category of Quiz 1 and Quiz 2 scores for each treatment group

Blue : Represents column percentage **Red:** Represents row percentage

In this study, a paired sample (or dependent) t-test has been conducted for each group (SC, SA, and ST respectively) to compare quiz score means after using both scenarios. The dependent t-test compares the means of two related groups to determine whether there is a statistically significant difference between these means. In this test, the same participants are tested more than once and present in both groups and measured on two occasions on the same dependent variable.

The t score is a ratio between the difference between two groups and the difference within the groups. A large t-score tells that the groups are different while a small t-score indicates that the groups are similar. Every t-value has a p-value that is the probability that the results from the sample data occurred by chance. The p-value calculated from t-test can be used to determine statistical significance in a hypothetical test. The significance level of 5% has been chosen for this study. If the p-value is less than 0.05 then the null hypothesis can be rejected with 95% confidence.

The performance of participants in each group was measured before and after they received their hints, to assess, whether hints improved their performance. A dependent t-test has been used here as there are two related groups for each condition. The first related group consists of the participants after using scenario 1 and prior to getting hints while the second related group consists of the same participants, but now after getting hints (all or tailored) at the end of the scenario 2.

The null hypothesis is that the pairwise difference between the two tests is equal (H0: $\mu d = 0$). The null hypothesis could be stated as H0: Participants who receive help (all hints or tailored hints) do not perform significantly better than the control group in the second scenario. However, we are interested to consider the different help strategies and thus propose and test the following two hypotheses using a one-tailed test because we are interested in the change in one direction, namely improved score performance.

H1: Participants who receive (all hints or tailored hints) perform significantly better than the control group in the second scenario.

H2: Participants who receive tailored hints perform significantly better than the all hints group in the second scenario.

The results of pairwise t-test is as follows.

- For 'Control' group, the value of t is -0.056. The value of p is 0.48. The result is not significant at $p \leq 0.05$. (Control)
- For 'All Hints' group, the value of t is -0.46. The value of p is 0.32. The result is not significant at $p \leq 0.05$.
- The Tailored Hints group, the value of t is -2.00. The value of p is 0.03. The result is significant at $p \leq 0.05$.

4.2.1 Quiz Results based on Engagement level

The categorised quiz scores were compared with levels of affective and cognitive engagement. Table 11 presents the average of quiz scores achieved by participants along with standard deviation for both scenarios along with the number of participants in each category of affective and cognitive engagement level.

Level	Affective Engagement (AE)			Cognitive Engagement (CE)		
	Low	Med	High	Low	Mid	High
Quiz1	6.23	6.90	8.46	6.25	7.33	7.60
Std. Dev	2.52	2.72	3.45	2.52	2.99	3.05
Quiz 2	5.97	6.53	8.08	6.05	6.80	7.65
Std. Dev.	2.26	2.70	2.54	2.75	2.44	2.57
Average Quiz 1 & 2	6.10	6.72	8.27	6.15	7.07	7.63

Table 11. Average Score of all participants with different level of affective and cognitive Engagement

According to these results, participants within high affective engagement category achieved the highest score for both attempts as compared to the participants in medium and low affective engagement category, however there is consistently decreasing numbers across the first attempt to second attempt. The participants with high cognitive engagement achieved the highest score in the first and second attempt and people with medium and low cognitive engagement have lower scores as compared to highly cognitively engaged participants, however that pattern is not consistent across the first attempt to the second attempt except high CE category people. Results indicate that only participants with high cognitive engagement improved their average score, though not significantly. Table 12 drills down further and compares categories of engagement with quiz score categories. It was further noted that four participants with high cognitive engagement were able to move from a medium quiz score to a high score after playing a second time, while no other groups were able to move to high quiz score.

	Level	Affective Engagement (AE)			Cognitive Engagement (CE)			
		Low	Medium	High	Low	Medium	High	**Total**
Quiz 1	Low	10 31.25%	18 56.25%	4 12.50%	15 46.88%	14 43.75%	3 9.38%	32
	Mid	5 14.29%	24 66.57%	6 17.14%	7 20.00%	22 62.86%	6 17.14%	35
	High	16 20.78%	45 58.44%	16 20.78%	22 28.57%	44 57.14%	11 14.29%	77
Quiz 2	Low	10 31.25%	19 59.38%	3 9.38%	16 50.00%	12 37.50%	4 12.50%	32
	Mid	9 23.08%	26 66.67%	4 10.26%	8 20.51%	30 76.92%	1 2.56%	39
	High	12 16.44%	42 57.53%	19 26.03%	20 27.40%	38 52.05%	15 20.55%	73
	Total	31	87	26	44	80	20	

Table 12. Number of participants within different levels of affective and cognitive Engagement

Chi-square tests comparing AE and CE categories with quiz score categories for all participants did not find any significant differences. However, when AE and CE are broken down into their components, we see some significant differences across score categories. For all groups (SC, SA, ST), the chi-square test showed statistically significant different results for FSL with the quiz score 2 $X2$ ($N=144$) = 9.73, $p= 0.05$, demonstrating statistical differences between these two variables. Chi-square tests also revealed significant associations between CE and quiz score 2 $X2$ ($N=144$) = 16.71, $p= 0.00$); IM and quiz score 2 $X2$ ($N=144$) = 12.53, $p= 0.01$).

Across all groups, there were 46 people in high FSL category and 27 (58.70%) of them got a high quiz score for scenario 2. Out of 26 participants with a high level of affective engagement, 61.54% of participants scored high for scenario 1 while this number increased to 73.08% for scenario 2. These figures emphasise the positive relationship between affective engagement and quiz score in line with the previous results.

For participants in tailored hints group (ST), the chi-square test showed statistically significant results for TSR with the quiz score 2 $X2$ ($N=57$) = 10.33, $p= 0.04$; FSL and quiz score 2 $X2$ ($N=57$) = 10.90, $p= 0.03$); IM and quiz score 2 $X2$ ($N=57$) = 12.02, $p= 0.02$). Chi-square tests do not show any statistically significant differences for engagement levels and quiz scores for control (SC) or all hints (SA) groups.

4.2.2 Quiz Results based on Personality

Table 13 presents the summary data for the personality dimensions of participants who scored high in the quizzes. Chi-square test was calculated on personality and quiz scores for both scenarios for all treatment groups. For all participants regardless of group, chi-square tests reveal significant differences in quiz performance between participants with different levels

of conscientiousness and quiz score 2 X2 (N=144) = 10.34, p= 0.04. There were 78 students with high conscientiousness and 44 of them scored high quiz score for attempt 2. The majority (54.17%) of participants scored high on the conscientiousness personality dimension; 22 (28.21%) of these achieved a score between 6 and 10 in the quiz test, while 56.41% achieved a score of 11/15 or over. These high achievers represent 60.27% of the participants achieving a high score. For their first attempt as well, this positive relationship is demonstrated with 42 (54.55%) high achievers out of 78 participants with a high level of conscientiousness. No significant differences were found between any other personality dimension and the quiz scores.

		Openness				Conscientiousness				Extravert				Agreeability				Neuroticism			
		Low	Medium	High	Total	Low	Medium	High	Total	Low	Medium	High	Total	Low	Medium	High	Total	Low	Medium	High	Total
Quiz 1 Score	Low	2 2.44 6.25 18.18	14 8.89 43.75 35.00	16 20.67 50.00 17.20	32 100.00 22.22	5 4.00 15.63 27.78	10 10.67 31.25 20.83	17 17.33 53.13 21.79	32 100.00 22.22	11 10.44 34.38 23.40	8 9.11 25.00 19.51	13 12.44 40.63 23.21	32 100.00 22.22	3 1.78 9.38 37.50	16 14.00 50.00 25.40	13 16.22 40.63 17.81	32 100.00 22.22	4 6.22 12.50 14.29	16 12.44 50.00 28.57	12 13.33 37.50 20.00	32 100.00 22.22
	Medium	3 2.67 8.57 27.27	10 9.72 28.57 25.00	22 22.60 62.86 23.66	35 100.00 24.31	3 4.38 8.57 16.67	13 11.67 37.14 27.08	19 18.96 54.29 24.36	35 100.00 24.31	8 11.42 22.86 17.02	8 9.97 22.86 19.51	19 13.61 54.29 33.93	35 100.00 24.31	3 1.94 8.57 37.50	15 15.31 42.86 23.81	17 17.74 48.57 23.29	35 100.00 24.31	8 6.81 22.86 28.57	14 13.61 40.00 25.00	13 14.58 37.14 21.67	35 100.00 24.31
	High	6 5.88 7.79 54.55	16 21.39 20.78 40.00	55 49.73 71.43 59.14	77 100.00 53.47	10 9.63 12.99 55.56	25 25.67 32.47 52.08	42 41.71 54.55 53.85	77 100.00 53.47	28 25.13 36.36 59.57	25 21.92 32.47 60.98	24 29.94 31.17 42.86	77 100.00 53.47	2 4.28 2.60 25.00	32 33.69 41.56 50.79	43 39.03 55.84 58.90	77 100.00 53.47	16 14.97 20.78 57.14	26 29.94 33.77 46.43	35 32.08 45.45 58.33	77 100.00 53.47
	Total	11 7.64 100.00	40 27.78 100.00	93 64.58 100.00	144 100.00 100.00	18 12.50 100.00	48 33.33 100.00	78 54.17 100.00	144 100.00 100.00	47 32.64 100.00	41 28.47 100.00	56 38.89 100.00	144 100.00 100.00	8 5.56 100.00	63 43.75 100.00	73 50.69 100.00	144 100.00 100.00	28 19.44 100.00	56 38.89 100.00	60 41.67 100.00	144 100.00 100.00
Quiz 2 Score	Low	2 2.44 6.25 18.18	13 8.89 40.63 32.50	17 20.67 53.13 18.28	32 100.00 22.22	2 4.00 6.25 11.11	18 10.67 56.25 37.50	12 17.33 37.50 15.38	32 100.00 22.22	8 10.44 25.00 17.02	13 9.11 40.63 31.71	11 12.44 34.38 19.64	32 100.00 22.22	0 1.78 0.00 0.00	20 14.00 62.50 31.75	12 16.22 37.50 16.44	32 100.00 22.22	7 6.22 21.88 25.00	17 12.44 53.13 30.36	8 13.33 25.00 13.33	32 100.00 22.22
	Medium	4 2.98 10.26 36.36	10 10.83 25.64 25.00	25 25.19 64.10 26.88	39 100.00 27.08	5 4.88 12.82 27.78	12 13.00 30.77 25.00	22 21.13 56.41 28.21	39 100.00 27.08	12 12.73 30.77 25.53	9 11.10 23.08 21.95	18 15.17 46.15 32.14	39 100.00 27.08	3 2.17 7.69 37.50	14 17.06 35.90 22.22	22 19.77 56.41 30.14	39 100.00 27.08	6 7.58 15.38 21.43	15 15.17 38.46 26.79	18 16.25 46.15 30.00	39 100.00 27.08
	High	5 5.58 6.85 45.45	17 20.28 23.29 42.50	51 47.15 69.86 54.84	73 100.00 50.69	11 9.13 15.07 61.11	18 24.33 24.66 37.50	44 39.54 60.27 56.41	73 100.00 50.69	27 23.83 36.99 57.45	19 20.78 26.03 46.34	27 28.39 36.99 48.21	73 100.00 50.69	5 4.06 6.85 62.50	29 31.94 39.73 46.03	39 37.01 53.42 53.42	73 100.00 50.69	15 14.19 20.55 53.57	24 28.39 32.88 42.86	34 30.42 46.58 56.67	73 100.00 50.69
	Total	11 7.64 100.00	40 27.78 100.00	93 64.58 100.00	144 100.00 100.00	18 12.50 100.00	48 33.33 100.00	78 54.17 100.00	144 100.00 100.00	47 32.64 100.00	41 28.47 100.00	56 38.89 100.00	144 100.00 100.00	8 5.56 100.00	63 43.75 100.00	73 50.69 100.00	144 100.00 100.00	28 19.44 100.00	56 38.89 100.00	60 41.67 100.00	144 100.00 100.00

Table 13. Number of participants in different levels of Quiz 1 and Quiz 2 within OCEAN categories

Green: Expected Frequencies **Blue** : Represents column percentage **Red:** Represents row percentage

36

For the control group as well, the chi-square test showed a significant difference between conscientiousness and quiz score 2 $X2$ ($N=37$) = 15.20, $p= 0.00$. Out of 18 participants with high conscientiousness, 14 (77.78%) scored high for quiz 2. No significant differences were found between other personality dimensions and the quiz score for this group. For all hints (SA) group, the chi-square test showed the significant difference between extraversion and quiz score 1 $X2$ ($N=50$) = 13.70, $p= 0.01$.

For Tailored Hints group, the chi-square test showed significant differences between agreeability and quiz score 1 $X2$ ($N=57$) = 11.81, $p= 0.02$ and quiz score 2 $X2$ ($N=57$) = 9.63, $p= 0.05$. The participants with high agreeability characteristics scored higher in quizzes. Out of 31 participants with high agreeability, 64.52% scored high for quiz 1. Similar patterns have been observed for the second quiz where 60% of highly agreeable participants scored high. No significant differences were found between other personality dimension and the quiz score for this group.

4.2.3 Analysis of Individual Quiz Question Results

The Quiz questions have been examined individually for both scenarios. This analysis was undertaken to ensure that the difficulty of questions in both scenarios was similar, to check whether certain questions or types of questions were more closely connected to a hint and if differences could account for differences between score outcomes between scenario 1 and 2 for different groups.

Table 14 and 15 show the number of correct answers per question by all participants along with the percentage for both scenarios respectively. The type of question is also included (e.g. observational (obs) questions that can be answered by making observations in the EVW; or questions related to inventory (inv) items that can be collected on the way; or related to a particular character or information provided by a particular character). Table 14 and 15 show the percentage of correct answers for both scenarios for all the groups.

Scenario 1 Questions (see Appendix C)															
Qn	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	1.10	1.11	1.12	1.13	1.14	1.15
Control (37)	20 54.05	11 29.73	28 75.68	22 59.46	20 54.05	21 56.76	17 45.95	18 48.65	20 54.05	10 27.03	16 43.24	17 45.95	15 40.54	21 56.76	14 37.84
All Hints (50)	22 44.00	7 14.00	32 64.00	24 48.00	27 54.00	27 54.00	23 46.00	25 50.00	32 64.00	10 20.00	18 36.00	25 50.00	18 36.00	28 56.00	20 40.00
Tailored Hints (57)	25 43.86	13 22.81	39 68.42	25 43.86	30 52.63	32 56.14	28 49.12	35 61.40	40 70.18	11 19.30	24 42.11	25 43.86	27 47.37	32 56.14	19 33.33
Type	Lyi	Oldm	Oldm	Oldm	Charl	Charl	Charl	Pedr	Pedr	Pedr	Obs	Obs	Obs	Obs	Inv

Table 14. Performance of Individual Questions per Group for Scenario 1

Scenario 2 Questions (see Appendix D)															
Qn	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	2.10	2.11	2.12	2.13	2.14	2.15
Control (37)	15 40.54	11 29.73	20 54.05	25 67.57	18 48.65	14 37.84	10 27.03	11 29.73	18 48.65	6 16.22	31 83.78	24 64.86	22 59.46	31 83.78	13 35.14
All Hints (50)	15 30.00	18 36.00	26 52.00	28 56.00	24 48.00	13 26.00	16 32.00	15 30.00	21 42.00	13 26.00	40 80.00	25 50.00	20 40.00	39 78.00	16 32.00
Tailored Hints (57)	19 33.33	20 35.09	19 33.33	33 57.89	19 33.33	21 36.84	14 24.56	13 22.81	26 45.61	11 19.30	49 85.96	27 47.37	24 42.11	49 85.96	21 36.84
Type	Charl	Charl	Kim	Zaf	Zaf	Zaf	Oldm	Oldm	Oldm	Oldm	obs	obs	obs	obs	inv

Table 15. Performance of Individual Questions per Group for Scenario 2

LEGEND:

Black : Represents number of correct questions **Red:** Represents percentage of correct questions

Source Lyi=Lyina; Oldm=OldMan; Charl=Charlie; Zaf=Zafirah; Ped=Pedro; Obs=Observation; Inv=Inventory

For scenario 1, where no hints were given to any of the group, Questions 1.3 and 1.9 are most commonly correctly answered for hints groups (SA and ST) and Questions 1.3 and 1.14 are the best-answered questions for the control group. Question 1.2 is the worst performing question for all the groups (see Table 14). Question 1.2 was *"What role does Omeweye play in Omosa village?"* with the answer *"Settle Disputes between Omosans"*. This information could be obtained from 'oldman' character.

For scenario 2, questions 2.11 and 2.14 are the best performing questions and 2.10 is the worst answered question (see Table 15). Question 2.11 was "What kind of house do villagers have in

Omosa? " with answer "huts". Question 2.14 was "What do the animals near the cliff look like? " with the answer "They look like antelopes with their long horns". Both of these questions are based on observation. The least correct question 2.10 in scenario 2 was "What is the meaning of Omeweye in Omosian language". This information could also be found from 'oldman ' character.

Table 16 shows the consolidated result in terms of how many percentages of answers were correct in all experimental groups respectively. It demonstrates that for the control group (ST), there is not much change in terms of correct answers but for hints groups, change is in a negative direction consistent with earlier results.

Group	**Scenario 1** **% Correct answers**	**Scenario 2** **% Correct answers**
All Participants	46.90	44.58
Control (SC)	48.65	48.47
All Hints (SA)	45.07	43.87
Tailored Hints (ST)	47.37	42.69

Table 16. Percentage correct answers for each group for both scenarios

For all participants, 46.90% answers were given correctly as compared to the second scenario where 44.58% answers were correct. Most participants in all groups performed better in the first scenario as compared to the second scenario however this difference was more for hints groups than the control group. It has been discussed further in the thesis (see Section 5.4).

4.3 Qualitative Results: Comments by participants

The participants from all hints group received all the generic hints provided in Table 2, before starting their navigation within the virtual world. The participants from tailored hints group received specific hints from the list of tailored hints provided in Table 2, depending on how they performed in their first interaction with the EVW. That resulted in more hints to be presented to the participants in all hints group as compared to the tailored hints group for most of the time. Participants in hints groups were requested to leave comments regarding the helpfulness of the hints. At the end of the study, all hints (SA) and tailored hints (ST) group participant answered a question "Please explain why you thought the hints were helpful or not" in order to determine whether hints were helpful or not from their perspective. Analysis of these comments shows that 18 (36%) participants from all hints (SA) group found the hints to be helpful while 25 (50%) participants from the same group stated they were not helpful and remaining participants didn't realise that the hints were given. One participant chose not to provide any comment and entered

'NA' to answer this question. From ST group, 23 (40.35%) participants found the hints to be helpful and 29 (50.87%) participants found them to be not helpful. Remaining five students didn't notice the hints. For both generic/all and tailored hints, reasons given for being not helpful include various factors that have been categorised and summarised (see Table 17) along with a number of participants in each category. The participants who found the hints useful stated the reasons as clarity, guidance, and navigational support.

Group	Helpful/Not Helpful	No of participants	Examples
All Hints (SA)	**Helpful**	**18 36%**	"Hints were helpful since they signalled what aspects of the game to focus on rather than roaming around for no particular reason", "They helped me understand and look for objects and information more effectively."
	Not Helpful	**25 50%**	
	Reasons		
	Insufficient information	16	"They just told me I did well the first time. There was no hints on what I should do to make the game work better", "There were not enough hints to make this helpful"
	Not clear	5	"There weren't especially helpful, just a few different tips were given however overall it was ambiguous", "Hints were not clear and were based on the previous game"
	Hard to remember	3	"I read them but forgot all about them a few minutes in... I must have been too immersed in the lore ;)", "information is difficult to retain"
	Not needed	1	"I already knew that I needed to talk to everyone and find the hidden notes", "The hints were depicting the details that could be self-found, they did not add any additional benefit."
	Did not notice	**7**	"didn't realise", "there were none"
Total		**50**	
Tailored Hints (ST)	**Helpful**	**23 40.35%**	"They gave me clues to finding more information", "Helped to grasp what was going on"
	Not Helpful	**29 50.87%**	
	Reasons		
	Insufficient information	17	"There were not many hints written", "They were extremely limited"
	Not clear	4	"The hints were quite vague and did not give any indication on where to go next", "They were vague"
	Unclear purpose of game	3	"Didn't understand the purpose of what I was trying to do", "I didn't realise the objective of the game"
	Not needed	2	"The hints were common sense", "Could already figure that out on my own"
	Confusing Task	3	"I couldn't navigate either way it was really weird", "I found the overall game difficult to navigate and understand"
	Did not notice	**5**	"Don't remember any hints", "They did not show"
Total:		**57**	

Table 17. Detailed analysis of participants' comments for all hints (ST) and tailored hints (ST) group

4.4 Chapter Summary

This chapter presents the results of this study. It provided the details of demographics of the participants in terms of their gender, age, game playing behaviour and cultural group. This section also analysed the relationships between other factors like engagement, personality, and performance of participants. Analysis of Individual questions and comments written by participants regarding the usefulness of hints have been carried out as well in order to find the link between the provision of hints and performance of participants in terms of correct answers.

CHAPTER 5: DISCUSSION

This chapter discusses the results presented in the previous chapter. Firstly, we revisit the proposed hypotheses (repeated below) and the results of the paired t-tests.

H_1: Participants who receive (all hints or tailored hints) perform significantly better than the control group in the second scenario.

H_2: Participants who receive tailored hints perform significantly better than the all hints group in the second scenario.

The pairwise t-test revealed a significant difference in the means of quiz 1 and quiz 2 for the tailored hints group (p-value = 0.03), but not for the control (p-value=0.48), or all hints groups (p-value=0.32). This might lead to wrongly conclude that H_1 is partially accepted (because tailored hints group Quiz 2 score was significantly different to their Quiz 1 score, but there is no significant difference for all hints group) and H2 is accepted. However, on analysis of the differences, we see that quiz performance was significantly worse for the tailored hints group, not better. The results showed that participants in the control group performed best in terms of correct answers for Scenario 1 while all hints (SA) group performed worst for scenario 1 and 'Tailored hints' (ST) performed worst for scenario 2. The results indicate that overall all participants performed worse for the second attempt (7.03 vs 6.69). Based on the results, both hypotheses must be rejected.

Overall, providing hints whether generic or tailored, did not improve the results of the participants significantly within an EVW. The cross-groups comparisons of demographic differences between control and hints groups were informative in a way that these comparisons highlighted gaps in our current understanding of the effect of factors such as personality, cognitive engagement and affective engagement on academic engagement and performance. These areas are in need of more investigation.

5.1 Providing hints to students with different personality and engagement levels

Regarding the influence of personality, our cohort had a majority of people with high openness, conscientiousness, and agreeable personality dimension and most of these categories attained a high score for scenario 1 and scenario 2. Past research results have also demonstrated that conscientious personality dimension is the most important determinant to affect the level of academic achievement (Chamorro-Premuzic & Furnham, 2008) and openness (Hazrati-Viari et al., 2012) has also been a successful predictor of performance in educational settings. People with openness personality dimension tend to perform better in subsequent tasks because of their creative and curious nature. Conscientious individuals have traits of being more organised, disciplined and hardworking (O'Connor & Paunonen, 2007). For this reason, it was expected that people with high open and conscientious personality dimension would achieve a higher score within EVW irrespective of hints provided.

Our results indicated a positive relationship between conscientiousness and performance, consistent with another relevant study (Hazrati-Viari et al., 2012). It appears that they learnt from their experience from using the EVW for the first time. This improvement was true whether participants received hints or not. However, no links were established between an openness personality trait and academic performance in this research. Results reflected a positive association between agreeability dimension and the performance score for tailored hints (ST) group. People who are agreeable tend to be more pleasant and accommodating in nature and can interact with fellow students with the cooperation and learn better in group situations (Eyong et al., 2014). It can be considered that the cooperative behaviour of these participants with virtual characters was one of the factors for improved learning and achieve better scores in quizzes as a result. Findings also revealed a positive link between extraverted category and performance for all hints (SA) group as demonstrated by prior research (Furnham & Monsen, 2009). Extravert trait relates to sociability, assertiveness and active behaviour. They tend to explore more and try to seek help more actively (Bidjerano & Dai, 2007). This might be the reason that participants with a high level of extravert trait performed well within EVW. The results did not show any significant results regarding the link between performance and other personality dimensions.

Student engagement can be influenced by teachers, school, family, curriculum, school, community, resources, etc. (Appleton et al., 2006). By using the Appleton instrument, the research sought to capture these factors that influence engagement. The literature suggests a strong positive correlation between student engagement and student achievement (Carini et al., 2006;

Fredricks et al., 2004; Hassaskhah et al., 2013; Hoff & Lopus, 2014). The measures of student engagement indicate individuals' intrinsic involvement with their learning and their learning processes. In the virtual world, the participants started their journey at one location where they were guided by a virtual character to move ahead in a certain direction and also suggested to collect the inventory items on their way by observing and exploring their surroundings. It was expected that the participant with high engagement level would pay more attention to the conversation retaining most of the information and would follow through the instructions to go to the next destination as suggested by the previous character. It was expected that an engaged person would be more attentive to the surroundings within the environment and perform better by answering most of the quiz questions correctly. Results show that participants with a higher level of affective and cognitive engagement scored higher in quizzes.

Consistent with previous study results (Makhija, Richards, De Haan, Dignum, & Jacobson, 2018), the results from this study also confirm a positive relationship between affective engagement and cognitive engagement. The students who possess a high level of affective engagement, also possess a high level of cognitive engagement as well and vice versa. This pattern is reflected in a similar manner for control groups and hint groups as well, with statistically significant results. These results also show a link between an extravert personality trait and affective engagement. Results suggest that extraverted people are more affectively engaged. A similar pattern has been observed between cognitive engagement and extravert dimension.

Other factors such as gender, gaming hours, and cultural background were also considered to determine their influence on performance, however, no specific differences were observed in this study from a performance perspective for any of these factors as quiz scores did not show any significant differences.

5.2 All and Tailored Hints Comments

In answer to the helpfulness of the hints by participants in all hints (SA) group and tailored hints (ST) groups, approximately half of the people from both groups found them to be not helpful. Reasons for these hints to be not useful mainly included insufficient or unclear information provided by these hints. Few participants found the environment to be very easy to navigate while few others found the environment to be challenging and did not consider hints to be sufficient enough to guide them in the right direction. Through these comments, participants have expressed their desire to get specific hints that can provide more information and can motivate them for experiential learning in education virtual world. According to a small number of

participants in all hints (SA) group, hints were hard to remember, however, no one from tailored hint (ST) group provided this statement. It could be attributed to the fact that as the number of hints provided for all hints (SA) group was more than the number of hints provided to the tailored hints (ST) group, it might be harder for participants to consider and remember all hints while navigating the virtual world. This suggests that providing clear and specific hints might provide an improved learning experience in EVWs.

The navigational hints used in this study have been derived from hints from a past study (De Haan & Richards, 2017) and new hints were added to address specific research questions in the current study. The comments left by participants in this study also indicated that hints need to be more specific and should be designed with sufficient content to be able to provide guidance in the right direction. This is consistent with a past study that suggested that hints need to be more specific and should be formulated in a way that attains maximum benefit in terms of learning outcomes within EVWs (De Haan & Richards, 2017). Results also suggested that generic hints might not always be suitable for all the participants and it might be a valid strategy to provide specific hints within EVW based on student's individual factors including personality, engagement level, navigational behaviour, and emotional state.

5.3 Individual question performance based on treatment group

Hints were of strategic and conceptual nature. In that direction, the possibility of mapping hints to the grouping of types of individual questions was also explored. If students paid attention to these hints, then questions related to that hint should have shown better results than for the questions that did not require a hint, a hint was not provided, or the student did not pay attention to. However, It was not possible to make a link between quiz questions, types of hints followed/ignored or any other factors that might influence performance.

As expected, during the first scenario similar patterns of question answering were observed for all treatment groups. Both of the best performing questions for all groups in scenario 2 were based on observation. Even though these were the best-answered questions for all groups, the percentage was highest for tailored hints (ST) group. The reason might be that these participants got relevant tailored hints and paid more attention to their environment in the second scenario.

From both the scenarios, information regarding worst performing questions could be obtained from 'oldman' character. This character was hard to locate in the virtual environment and needed

careful attention and navigational skill. There is a possibility that many participants skipped this character and could not get the relevant information. Further analysis is needed in this aspect.

5.4 Answering the research questions

In answer to the first and second research questions, It cannot be suggested that providing help to the participants in the form of generic or tailored hints improves quiz performance or that one kind of help (i.e. generic or tailored hints) is better than the other. Overall, the participants that were in the hint groups did worse than those in the control group for both quizzes. The average quiz 1 score for the control group did not change for the quiz 2 score, while the average went down for both hint groups. This indicates that our intervention of adding hints has worsened performance and in case of tailored hints performed significantly declined. This suggests that hints do make a difference, but if not designed well, they can make it worse for students. These results may suggest that rather than determining what hints to give students, in line with impasse learning (VanLehn et al., 2003), we should allow students to request help when they feel a need for it and provide support within that context. For example, if a student reports they feel lost or don't know whom to speak to, we can provide the appropriate hint at that time.

These results suggest that there is a need for further investigation of other factors (such as personality, engagement or other demographics) as well while providing hints. While our results are inconclusive and in some cases opposite to what we had expected, these results have provided us with an opportunity to explore more deeply the reasons behind our results and to consider alternative interpretations and extract new possible insights from data collection to get new leads for understanding individual factors that might influence the learning in EVW. The unexpected answers to the first two research questions concerning providing all/generic or tailored hints also further motivate finding answers to the third and fourth research questions that look at the relationship between scores, personality and engagement levels, respectively, and the fifth research question that looks for connections between these factors and the scores.

In answer to the third research question, our results show some evidence of relationships between the conscientiousness personality dimension with performance in the second round for all participants and for participants in the control group (SC). A link between agreeability characteristics and performance was also found for participants of tailored hint (ST) group for both scenarios. Our analysis of the five personality dimensions with the second quiz score revealed significant differences for all hints (SA) group with extravert dimension. Thus, we can suggest that different personality types benefit more from hints or help than others.

In answer to the fourth research question, our results show some evidence of relationships between the engagement categories (i.e. Family Support for Learning (FSL), Teacher-Student Relationships (TSR) and intrinsic motivation (IM)) with performance in second scenario for all participants and participants in Tailored hints group (ST). For example, our data showed that some participants with high cognitive engagement were able to improve their score in the second quiz even though the number of participants achieving a high score reduced for the other cognitive engagement groups and all of the affective engagement groups. This suggests that these students problem-solving skills and/or motivation to improve their performance was stronger than for others.

In answer to the fifth research question, our results show some evidence of relationships between the engagement categories and specific personality characteristics. Extravert people have shown the characteristics of being more affectively and cognitively engaged. There is also evidence of a positive association of affective and cognitive engagement levels with agreeability trait.

5.5 Chapter Summary

While our hypotheses regarding the effect of hints on performance were not supported, our results confirmed our expectations based on the literature that participants that are high in conscientiousness, extravertedness and agreeableness scored higher in quizzes than their counterparts who are low in these characteristics. Also, our results confirmed our expectation that participants who had a higher level of affective and cognitive engagement, scored better in quizzes as compared to those with a lower level of these dimensions.

CHAPTER 6: CONCLUSION

The motivation for this research study is to increase understanding of the learner and their performance in EVWs, to gain insights for effective EVW learning design and how the EVW should adapt according to specific student features and needs. This research attempts to investigate whether providing hints (generic/all or tailored) and other factors such as personality characteristics and engagement (Cognitive and Affective) influence the students' performance in an EVW. The academic performance of the participants was measured through the quiz scores conducted after interaction with the EVW. Hints were provided to the participants based on various aspects as discussed in the methodology section. A previous study was used to derive appropriate navigation hints that captured navigation data, identified navigation path patterns and analysed navigation paths to identify their relationship to quiz score (De Haan & Richards,

2017). This study significantly extends that initial work by utilising more logfile data items to derive tailored hints, evaluating different hint strategies and analysing the results according to distinctive individual differences.

6.1 Summary

This research study investigated performance behaviours exhibited by 144 first year psychology students as they participated in a scientific inquiry-based project delivered by an Educational Virtual World (EVW). Participants' data including gender, cultural group, personality type, and engagement profile were collected. Participants used the EVW twice. After each use, they answered a quiz about what they learnt in the world. After the first usage, participants were randomly allocated to receive no hints, all generic hints or tailored hints. Tailored hints were created based on their navigation patterns and other log file data from the first use of the EVW. This research examined the relationship between students' engagement score, personality type and performance they achieved after getting generic and tailored hints while navigating through the virtual world. Specific positive correlations have been found in this study. The quiz scores were found to be positively correlated with engagement dimensions. The study found that engagement behaviour and personality traits are correlated with the learning outcome but providing hints did not demonstrate expected improvements.

6.2 Limitations and Future Work

This study recruited psychology students as we had access to these participants. Due to the technical problems related to our university servers at a critical point in the data collection, we had unreliable data service and needed to discard incomplete data. Larger sample sizes and different populations might show different results.

Due to worsened quiz performance after receiving hints, the results offer limited guidance for improving the design of EVWs. Further studies using alternative designs may shed light on why this happened. Further understanding could possibly be gained through qualitative analysis of the comments regarding the usefulness of hints by participants. These comments can be analysed further by taking individual factors into consideration and align hints to specific personalities and engagement levels accordingly. EVWs should be designed to provide more specific guidance for students with different levels of academic achievement and their individual behaviour and potentially hints should only be offered when sought. These recommendations have not been tested in this study.

To provide a controlled study, two structured scenarios were created. The scenarios might not have been sufficiently motivating or directed to encourage interaction and remembering the information provided. Productive failure (Kapur, 2008) and impasse learning (VanLehn et al., 2003) suggest that students should be provided with a challenge at the start of learning new concepts, rather than a highly structured or guided task. A challenge or "quest" might have drawn out more differences in results according to individual personality and engagement differences. We had chosen not to implement an open-ended exploratory-type quest in order to control the experiment and contain it within a time block that would attract and retain participants.

As presented in the literature review, this project reviewed the importance of the learner's state, particularly with respect to the epistemic emotions experienced by the user. However, the capture of the learner's emotional state and appropriate tailored responses to their state added further complexity that was beyond the scope of a study that could be conducted in a nine-month time frame. This direction was left as future work.

Looking to the future, by assessing the real-time performance in virtual world, and considering factors such as personality characteristics and level of cognitive and affective engagement, students can be given appropriate help in term of practical hints such as which direction they should head to, what action they should take, where they should pay attention to, etc. Characters can also interact according to the personality of the learner and predict their behaviour during navigation. From a future perspective, the collected data can serve as a base to predict the behaviour pattern of individuals such as to determine whether a person can backtrack or not. Future work can also consider recognising student's affective state to evaluate the effective strategies to provide assistance. Further studies are needed to confirm the importance and value of tailoring hints, though intuitively and also supported by impasse learning concepts, offering help before it is needed is usually unwanted and unhelpful. All the suggested strategies will help enhance the adaptability of virtual worlds to facilitate effective and efficient learning outcomes by providing tailored support and present them as a valuable asset for students' learning. In a future research endeavour, multimodal input can be used to understand the learners' states effectively during complex learning and appropriate help strategies can be suggested matching with the individual's need.

Further research with larger cohorts and other non-psychology students would strengthen the generality of the results. The findings of this study stress the significance of understanding the learner better in order to design more efficient virtual learning platforms. As a future work, school

student populations can be included for further examination of links studied in this study. New research could be based on these findings to specify more clearly what helps to improve performance in an EVW and in what way help can be provided to students depending on their personality traits, engagement level and affective state that the learner is going through at that moment, to improve their performance. A few other navigational factors such as time spent, idle time in the world, the path taken to reach the destination, errors made, etc, can be considered in future studies. The factors such as engagement and personality of individuals have also been considered in the data analysis to find their link with the academic performance in this research. Further data capture and analysis are needed to explore the combined effect of these factors to allow successful learning in EVWs.

6.3 Final Remarks

Virtual environments present a promising platform for education and training purposes by enabling students to learn in a three-dimensional environment. By sensing the learner's state, a virtual agent can provide tailored help to the students in the virtual world. A combination of artificial intelligent technologies needs to be employed to take full advantages of its potential while interacting with the students. Analysis of past performance and other predictor variables such as personality, affective and cognitive engagement could provide a better perspective of the probable academic performance of students in the virtual world and can help in providing appropriate hints prior to the navigation in the virtual world.

A deeper and accurate understanding of learner features and behaviours in EVWs can facilitate better artificially intelligent technologies to improve the learning effectiveness of EVWs. This research found some differences in performance depending on the way hints were provided in either all or tailored form along with the influence of personality and engagement level and the other factors such as gender, culture, and game playing behaviour. The information regarding personality and engagement level were captured outside the EVW through self-reported questionnaires. Help was provided based on the attempt made in a previous playthrough, not the current one. In the future, this information can be captured in real time and passed on to the EVW in order to adapt it using the rules uncovered from the current experimental datasets. Students' emotions, personality, and other individual behaviours can be measured via interaction with IVAs and this collected data can be used to provide tailored responses to each student based on their personal characteristics. This naturalistic way to collect data could help inform the responses of IVAs and can help enhance the learning outcomes in EVWs.

The motivation behind this study is inspired by the fact that awareness of the learner's behaviour, engagement level, knowledge of their personality trait and their performance in EVWs can help gain better insights and help achieve an effective and efficient design of EVW that would adapt according to the needs of individual student to facilitate better learning in virtual environment.

Furthermore, this research aimed to determine an effective support strategy for the learners by considering their engagement scores and personality traits and responding appropriately to engage the learner within an EVW and bring them back to the optimal learning path. Emotions are a crucial component of learning complex skills. As discussed in the literature review, EVWs should be designed with the capability to assess when the student gets stuck or disengaged (Burleson & Picard, 2004; Villarica & Richards, 2014a) while learning and be able to direct the student in the right direction to maintain the positive learning. Furthermore, understanding of individual characteristics and identification of the emotional state of users can improve the effectiveness and efficiency of Educational Virtual World (Kort et al., 2001; Shen et al., 2009) and adapt them according to individual needs of the user. The purpose of this study was to determine whether logfile data captured while playing in an EVW could be used to provide useful tailored hints and to see whether some types of learners benefit more from such hints. Our results invite further investigation into understanding what factors are important in designing EVW that can be flexible enough to adapt for the user according to their emotional needs and personal characteristics and provide them a tailored experience and enhanced learning.

7. REFERENCES

Ammar, M. B., Neji, M., Alimi, A. M., & Gouardères, G. (2010). The affective tutoring system. *Expert Systems with Applications, 37*(4), 3013-3023.

Appleton, J. J., Christenson, S. L., Kim, D., & Reschly, A. L. (2006). Measuring cognitive and psychological engagement: Validation of the Student Engagement Instrument. *Journal of School Psychology, 44*(5), 427-445.

Arguel, A., & Lane, R. (2015). Fostering deep understanding in geography by inducing and managing confusion: an online learning approach.

Arguel, A., Lockyer, L., Lipp, O. V., Lodge, J. M., & Kennedy, G. (2017). Inside Out: Detecting Learners' Confusion to Improve Interactive Digital Learning Environments. *Journal of Educational Computing Research, 55*(4), 526-551.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel psychology, 44*(1), 1-26.

Bauer, K. W., & Liang, Q. (2003). The effect of personality and precollege characteristics on first-year activities and academic performance. *Journal of College Student Development, 44*(3), 277-290.

Beer, C., Clark, K., & Jones, D. (2010). Indicators of engagement. *Curriculum, technology & transformation for an unknown future. Proceedings ascilite Sydney*, 75-86.

Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences, 17*(1), 69-81.

Bouvier, P., Lavoué, E., Sehaba, K., & George, S. (2013). *Identifying Learner's Engagement in Learning Games: A Qualitative Approach based on Learner's Traces of Interaction.* Paper presented at the 5th International Conference on Computer Supported Education (CSEDU 2013).

Brandtzæg, P. B., Følstad, A., & Heim, J. (2018). Enjoyment: lessons from Karasek. In *Funology 2* (pp. 331-341): Springer.

Burleson, W., & Picard, R. W. (2004). *Affective agents: Sustaining motivation to learn through failure and a state of stuck.* Paper presented at the Workshop on Social and Emotional Intelligence in Learning Environments.

Busato, V. V., Prins, F. J., Elshout, J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and individual differences, 29*(6), 1057-1068.

Calongne, C. M. (2008). Educational frontiers: Learning in a virtual world. *Educause review, 43*(5), 36-38.

Carini, R. M., Kuh, G. D., & Klein, S. P. (2006). Student engagement and student learning: Testing the linkages. *Research in higher education, 47*(1), 1-32.

Chamorro-Premuzic, T., & Furnham, A. (2003). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in personality, 37*(4), 319-338.

Chamorro-Premuzic, T., & Furnham, A. (2008). Personality, intelligence and approaches to learning as predictors of academic performance. *Personality and individual differences, 44*(7), 1596-1603.

Chi, M. T., Leeuw, N., Chiu, M. H., & LaVancher, C. (1994). Eliciting self-explanations improves understanding. *Cognitive science, 18*(3), 439-477.

Cruz-Benito, J., Therón, R., García-Peñalvo, F. J., & Lucas, E. P. (2015). Discovering usage behaviors and engagement in an Educational Virtual World. *Computers in Human behavior, 47*, 18-25.

D'Mello, S., & Calvo, R. A. (2013). *Beyond the basic emotions: what should affective computing compute?* Paper presented at the CHI'13 Extended Abstracts on Human Factors in Computing Systems.

D'Mello, S. K., & Graesser, A. (2010). Multimodal semi-automated affect detection from conversational cues, gross body language, and facial features. *User Modeling and User-Adapted Interaction, 20*(2), 147-187.

D'Mello, S., Lehman, B., Pekrun, R., & Graesser, A. (2014). Confusion can be beneficial for learning. *Learning and Instruction, 29*, 153-170.

Darken, R. P., & Sibert, J. L. (1993). *A toolset for navigation in virtual environments.* Paper presented at the Proceedings of the 6th annual ACM symposium on User interface software and technology.

Darken, R. P., & Sibert, J. L. (1996). Navigating large virtual spaces. *International Journal of Human-Computer Interaction, 8*(1), 49-71.

De Haan, J., & Richards, D. (2017). Navigation Paths and Performance in Educational Virtual Worlds. *Navigation*.

De Haan, J., Richards, D., & Dignum, F. (2017). *Aiding learning efficiency in virtual worlds.* Paper presented at the Virtual System & Multimedia (VSMM), 2017 23rd International Conference on.

Dieterle, E. (2009). Multi-user virtual environments for teaching and learning. In *Encyclopedia of Multimedia Technology and Networking, Second Edition* (pp. 1033-1041): IGI Global.

Duo, S., & Song, L. X. (2012). An e-learning system based on affective computing. *Physics Procedia, 24*, 1893-1898.

Eschenbrenner, B., Nah, F. F.-H., & Siau, K. (2008). 3-D virtual worlds in education: Applications, benefits, issues, and opportunities. *Journal of Database Management (JDM), 19*(4), 91-110.

Eyong, E. I., David, B. E., & Umoh, A. J. (2014). The Influence of Personality Trait on the Academic Performance of Secondary School Students in Cross River State, Nigeria. *IOSR Journal Of Humanities And Social Science, 19*(3), 12-19.

Fosse, T. H., Buch, R., Säfvenbom, R., & Martinussen, M. (2015). The impact of personality and self-efficacy on academic and military performance: The mediating role of self-efficacy. *Journal of Military Studies, 6*(1), 47-65.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of educational research, 74*(1), 59-109.

Furnham, A., & Monsen, J. (2009). Personality traits and intelligence predict academic school grades. *Learning and Individual Differences, 19*(1), 28-33.

Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological assessment, 4*(1), 26.

Gosling, S. D., Rentfrow, P. J., & Swann, W. B. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in personality, 37*(6), 504-528.

Hanna, N., Richards, D., Hitchens, M., & Jacobson, M. J. (2014). *Towards Quantifying Player's Involvement in 3D Games Based-on Player Types.* Paper presented at the Proceedings of the 2014 Conference on Interactive Entertainment.

Hanna, N., Richards, D., & Jacobson, M. J. (2014). *Academic Performance in a 3D Virtual Learning Environment: Different Learning Types vs. Different Class Types.* Paper presented at the Pacific Rim Knowledge Acquisition Workshop.

Hassaskhah, J., Khanzadeh, A. A., & Mohamad Zade, S. (2013). The Relationship between Internal Forms of Engagement (Cognitive-Affective) and Academic Success across Years of Study. *Issues in Language Teaching, 1*(2), 251-272.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112.

Hazrati-Viari, A., Rad, A. T., & Torabi, S. S. (2012). The effect of personality traits on academic performance: The mediating role of academic motivation. *Procedia-Social and Behavioral Sciences, 32*, 367-371.

Hoff, J., & Lopus, J. S. (2014). *Does student engagement affect student achievement in high school economics classes.* Paper presented at the the annual meetings of the Allied Social Science Association, Philadelphia, PA.

Kapur, M., & Bielaczyc, K. (2012). Designing for productive failure. *Journal of the Learning Sciences, 21*(1), 45-83.

Kim, Y. (2005). *Empathetic virtual peers enhanced learner interest and self-efficacy.* Paper presented at the Workshop on Motivation and Affect in Educational Software, in conjunction with the 12th International Conference on Artificial Intelligence in Education.

Kim, Y., Baylor, A. L., & Shen, E. (2007). Pedagogical agents as learning companions: the impact of agent emotion and gender. *Journal of Computer Assisted Learning, 23*(3), 220-234.

Kort, B., Reilly, R., & Picard, R. W. (2001). *An affective model of interplay between emotions and learning: Reengineering educational pedagogy-building a learning companion.* Paper presented at the Advanced Learning Technologies, 2001. Proceedings. IEEE International Conference on.

Leony, D., Muñoz-Merino, P. J., Pardo, A., & Kloos, C. D. (2013). Provision of awareness of learners' emotions through visualizations in a computer interaction-based environment. *Expert Systems with Applications, 40*(13), 5093-5100.

Levitski, A., Radun, J., & Jokinen, K. (2012). *Visual interaction and conversational activity.* Paper presented at the Proceedings of the 4th Workshop on Eye Gaze in Intelligent Human Machine Interaction.

Lima, R., & Rosatelli, M. (2003). *An intelligent tutoring system to a virtual environment for teaching and learning.* Paper presented at the IX Workshop de Informática na Escola, Campinas. Anais do XXIII Congresso da Sociedade Brasileira de Computação.

Lin, L., Atkinson, R. K., Christopherson, R. M., Joseph, S. S., & Harrison, C. J. (2013). Animated agents and learning: Does the type of verbal feedback they provide matter? *Computers & Education, 67*, 239-249.

Lounsbury, J. W., Sundstrom, E., Loveland, J. M., & Gibson, L. W. (2003). Intelligence,"Big Five" personality traits, and work drive as predictors of course grade. *Personality and individual differences, 35*(6), 1231-1239.

Makhija, A., Richards, D., Caballé, S., & Conesa, J. (2018). *Providing Timely Support to Students in Educational Virtual Worlds.* Paper presented at the International Conference on Intelligent Networking and Collaborative Systems.

Makhija, A., Richards, D., De Haan, J., Dignum, F., & Jacobson, M. J. (2018). *The Influence of Gender, Personality, Cognitive and Affective Student Engagement on Academic Engagement in Educational Virtual Worlds.* Paper presented at the International Conference on Artificial Intelligence in Education.

Mcquiggan, S. W., Lee, S., & Lester, J. C. (2007). *Early prediction of student frustration.* Paper presented at the International Conference on Affective Computing and Intelligent Interaction.

McQuiggan, S. W., & Lester, J. C. (2007). Modeling and evaluating empathy in embodied companion agents. *International Journal of Human-Computer Studies, 65*(4), 348-360.

O'Brien, H. (2016). Theoretical perspectives on user engagement. In *Why Engagement Matters* (pp. 1-26): Springer.

O'Connor, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and individual differences, 43*(5), 971-990.

Paiva, A., Dias, J., Sobral, D., Aylett, R., Woods, S., Hall, L., & Zoll, C. (2005). Learning by feeling: Evoking empathy with synthetic characters. *Applied Artificial Intelligence, 19*(3-4), 235-266.

Paunonen, S. V., & Ashton, M. C. (2001). Big Five predictors of academic achievement. *Journal of Research in personality, 35*(1), 78-90.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In *Handbook of research on student engagement* (pp. 259-282): Springer.

Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological bulletin, 135*(2), 322.

Prendinger, H., & Ishizuka, M. (2005). The empathic companion: A character-based interface that addresses users'affective states. *Applied Artificial Intelligence, 19*(3-4), 267-285.

Ranjbartabar, H., & Richards, D. (2017). *Student Designed Virtual Teacher Feedback.* Paper presented at the Proceedings of the 9th International Conference on Computer and Automation Engineering.

Richards, D., Jacobson, M. J., Taylor, M., Newstead, A., Taylor, C., Porte, J., . . . Hanna, N. (2012). *Learning to be scientists via a virtual field trip.* Paper presented at the Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 3.

Robison, J. L., Mcquiggan, S. W., & Lester, J. C. (2009). *Modeling Task-Based vs. Affect-based Feedback Behavior in Pedagogical Agents: An Inductive Approach.* Paper presented at the AIED.

Ruddle, R. A., & Lessels, S. (2006). Three levels of metric for evaluating wayfinding. *Presence: Teleoperators and Virtual Environments, 15*(6), 637-654.

Sabourin, J., Mott, B., & Lester, J. (2011). *Computational models of affect and empathy for pedagogical virtual agents.* Paper presented at the Standards in emotion modeling, Lorentz Center International Center for workshops in the Sciences.

Sabourin, J. L., & Lester, J. C. (2014). Affect and engagement in Game-BasedLearning environments. *IEEE Transactions on Affective Computing, 5*(1), 45-56.

Sas, C., O'Hare, G., & Reilly, R. (2004). *A performance analysis of movement patterns.* Paper presented at the International Conference on Computational Science.

Shen, L., Wang, M., & Shen, R. (2009). Affective e-learning: Using" emotional" data to improve learning in pervasive learning environment. *Journal of Educational Technology & Society, 12*(2), 176.

Tian, F., Gao, P., Li, L., Zhang, W., Liang, H., Qian, Y., & Zhao, R. (2014). Recognizing and regulating e-learners' emotions based on interactive Chinese texts in e-learning systems. *Knowledge-Based Systems, 55*, 148-164.

Van der Meij, H. (2013). Motivating agents in software tutorials. *Computers in Human behavior, 29*(3), 845-857.

VanLehn, K., Siler, S., Murray, C., Yamauchi, T., & Baggett, W. B. (2003). Why do only some events cause learning during human tutoring? *Cognition and Instruction, 21*(3), 209-249.

Villarica, R., & Richards, D. (2014a). *Educational Scaffolding for Students Stuck in a Virtual World*.

Villarica, R., & Richards, D. (2014b). *Intelligent and empathic agent to support student learning in virtual worlds.* Paper presented at the Proceedings of the 2014 Conference on Interactive Entertainment.

Warburton, S. (2009). Second Life in higher education: Assessing the potential for and the barriers to deploying virtual worlds in learning and teaching. *British Journal of Educational Technology, 40*(3), 414-426.

Webster, J., & Ahuja, J. S. (2006). Enhancing the design of web navigation systems: The influence of user disorientation on engagement and performance. *Mis Quarterly*, 661-678.

Whitehill, J., Serpell, Z., Lin, Y.-C., Foster, A., & Movellan, J. R. (2014). The faces of engagement: Automatic recognition of student engagementfrom facial expressions. *IEEE Transactions on Affective Computing, 5*(1), 86-98.

Woolf, B., Burleson, W., Arroyo, I., Dragon, T., Cooper, D., & Picard, R. (2009). Affect-aware tutors: recognising and responding to student affect. *International Journal of Learning Technology, 4*(3-4), 129-164.

Wu, L., & Looi, C.-K. (2010). *Agent prompts: scaffolding students for productive reflection in an intelligent learning environment.* Paper presented at the Intelligent Tutoring Systems.

Wu, Q., Shen, Z., & Miao, C. (2013). *Stimulating students' curiosity with a companion agent in virtual learning environments.* Paper presented at the EdMedia: World Conference on Educational Media and Technology.

Yee, N., Harris, H., Jabon, M., & Bailenson, J. N. (2011). The expression of personality in virtual worlds. *Social Psychological and Personality Science, 2*(1), 5-12.

8. APPENDICES : QUESTIONNAIRES

A. Student Engagement Instrument (SEI)

All questions are answered on a 4 point Likert scale as follows:

1-Strongly Agree *2-Agree* *3-Disagree* *4-Strongly Disagree*

1. Overall, staff members at my university treat students fairly.
2. Staff members at my school listen to the students.
3. At my university, lecturers care about students.
4. My lecturers are there for me when I need them.
5. The university rules are fair.
6. Overall, my lecturers are open and honest with me.
7. I enjoy talking to the lecturers here.
8. I feel safe at the university.
9. Most lecturers at my school are interested in me as a person, not just as a student.
10. The exams in my classes do a good job of measuring what I'm able to do.
11. Most of what is important to know you learn in school and university.
12. The grades in my classes do a good job of measuring what I'm able to do.
13. What I'm learning in my classes will be important in my future.
14. After finishing my coursework I check it over to see if it's correct.
15. When I do coursework I check to see whether I understand what I'm doing.
16. Studying is fun because I get better at something.
17. When I do well at university it's because I work hard.
18. I feel like I have a say about what happens to me at the university.
19. Others students at the university care about me.
20. Students at my university are there for me when I need them.
21. Other students here like me the way I am.
22. I enjoy talking to the students here.
23. Students here respect what I have to say.
24. I have some friends at the university.
25. Going to university is important for achieving my future goals.
26. My education will create many future opportunities for me.
27. I am hopeful about my future.
28. My family are there for me when I need them.
29. When I have problems at the university my family are willing to help me.
30. When something good happens at the university, my family want to know about it.
31. My family want me to keep trying when things are tough at university.
32. I'll learn, but only if my family gives me a reward.
33. I'll learn, but only if the lecturer gives me a reward.

B. Ten Item Personality Inventory (TIPI)

Here are a number of personality traits that may or may not apply to you. Please write a number next to each statement to indicate the extent to which you agree or disagree with that statement. You should rate the extent to which the pair of traits applies to you, even if one characteristic applies more strongly than the other.

1-Disagree Strongly *2-Disagree Moderately* *3- Disagree Little* *4-Neither Agree Nor Disagree*
5-Agree a Little *6-Agree Moderately* *7-Agree Strongly*

I see myself as:

1. Extraverted, enthusiastic.
2. Critical, quarrelsome.
3. Dependable, self-disciplined.
4. Anxious, easily upset.
5. Open to new experiences, complex.
6. Reserved, quiet.
7. Sympathetic, warm.
8. Disorganized, careless.
9. Calm, emotionally stable.
10. Conventional, uncreative.

C. Scenario 1

1.1 What is the occupation of the woman in the village?

- *Hunter*
- Cook
- Village chief
- Biologist

1.2 What role does Omeweye play in Omosa village?

- *Settle Disputes between Omosans*
- Provides advice to villagers regarding farming
- Watches for predator animals in the farm
- Predicts weather conditions

1.3 Why do Omosans burn off some land areas?

- *To prevent fires from ravaging lands.*
- To repel bad elements
- To cook their meal
- To scare away wild animals

1.4 What do the villagers do to encourage growth of plants?

- *Fire-stick farming*
- compost
- irrigation
- shifting cultivation

1.5 What reason has the ecologist found about slow growing of plants?

- *Dry climate*
- Insufficient rain
- massive flooding
- Infertile Soil

1.6 What kind of data is analysed by the ecologist?

- *Data from the fields*
- Data about animal behaviour
- Data from village hall
- Data about Omosa residents

1.7 What did the ecologist find out about the plant life?

- *There are more fire-loving plants*
- There are more rain-loving plants
- There are more drought-loving plants
- There are more compost-loving plants

1.8 What has the animal behaviourist found out about the predatory animals?

- *The predators are hunting more aggressively.*
- The predators are starting to eat more plants.
- The predators are hunting in packs more.
- The predators have learned how to swim.

1.9 How do prey animals protect themselves from predators?

- *By moving in tighter pack*
- By attacking predators in groups
- By hiding in the caves
- By hiding in the plants

1.10 Animals are dying out at the island. According to the biologist , what is reason behind that?

- *Their habitat is being disturbed*
- The weather conditions have changed
- There are too many predators
- Visitors have brought diseases from overseas

1.11 What kind of plants could be found between the farm and the ecology lab?

- *Burnt trees.*
- Lush trees with fruit.
- Low brushes.
- There are no plants between the farm and ecology lab.

1.12 Where could you find the tree Ring Data?

- *In the building near the cliff.*
- In the village hall.
- On top of the cliff.
- In the forest.

1.13 What is the biologist Location in the world?

- *On the cliffs*
- Near the farm
- In the village
- In the red research facility

1.14 What do the villagers keep in townhall?

- *The stuffed heads of hunted animals*
- Paintings of the island
- Traditional rugs that the villagers make
- Fruits and seeds

1.15 What are predator animals called at Omosa?

- *Tooru*
- Yernt
- Toru
- Leopard

D. Scenario 2

2.1 How long is the history of fire-stick farming on Omosa?

- *At least 40000 years*
- At least 50000 Years
- At least 10000 Years
- At least 1000 years

2.2 What pattern has been observed by Ecologist regarding fire loving plants?

- *Increase in pollen from fire-loving plants*

- Decrease in growth of fire-loving plants
- Decrease in pollen from fire-loving plants
- Increase in growth of fire-loving plants

2.3 How has the behaviour of animals changed over the years?

- *The animals travel more to find food.*
- The animals travel less to conserve energy.
- The animal packs are less dense to find more food.
- The animal packs are more dense to protect against predators.

2.4 What is the Omosa weather pattern called?

- *Great Omosan Oscillation*
- Grand Omosan Oscillation
- Great Omosan System
- Great Tropical System

2.5 What has the climatologist, in the red research facility, found out?

- *That the last wet season was home to a massive flooding of the hunting grounds.*
- That the last dry season was the cause of a mass migration of birds.
- That the unchanging season has made living conditions for animals harder.
- That the rapidly changing seasons have caused more tornadoes.

2.6 What is the weather pattern in Omosa?

- *Long wet seasons and long dry seasons alternate.*
- Wet seasons with short dry seasons.
- Dry seasons with short wet seasons.
- Always dry weather.

2.7 The old man near the farm says that the hunting practices changed over time. How did the hunting practices change?

- *People hunt in places where it used to be forbidden to hunt*
- People are forbidden to hunt in places where they used to hunt
- People use traps to hunt animals, but used only bows in the past
- People only use bows to hunt, but used to hunt with traps

2.8 Why does the farm near the old man not have any crops on it?

- *There has just been a harvest.*
- The farm flooded over a little while ago.
- The harvest failed because of a plant disease.
- The farm is only used in the winter.

2.9 What ritual do the villagers practice?

- *The villagers dance around the fire in a festive ritual*
- The villagers sit around the fire and tell stories
- The villagers go hunting in the night once a year
- The villagers organise sports once a year

2.10 What is meaning of Omeweye in Omosian language

- *Ancient one*
- Wise one
- Old one

- Elder one

2.11 What kind of house do villagers have in Omosa?

- *Huts*
- Igloo
- Cottage
- Concrete

2.12 In the research centre where you started

- *On the cliffs*
- In the village
- In the big red research facility

2.13 Where could you find the book with field notes of the flora and fauna of the island?

- *In the village hall.*
- Next to the old man.
- In the building where you started.
- On top of the cliff.

2.14 What do the animals near the cliff look like?

- *They look like antelopes with their long horns.*
- They look like zebras with their striped skin.
- They look like bears.
- They look like horses with green scaly skin.

2.15 What are prey animals called at Omosa?

- *Yernt*
- Tooru
- Toru
- Leopard